December 2019

DETECTION AND INFERENCE IN GRAVITATIONAL WAVE ASTRONOMY

Steven Daniel Reyes
Syracuse University

ABSTRACT

We explore the detection and astrophysical modeling of gravitational waves detected by the Advanced Laser Interferometer Gravitational wave Observatory (LIGO) and Virgo. We discuss the techniques used in the `PyCBC` search pipeline to discover the first gravitational wave detection GW150914, and estimate the statistical significance of GW150914, and the marginal trigger LVT151012. During Advanced LIGO's first observing run there were no detections of mergers from binary neutron star and neutron star-black hole binaries. We use Bayesian inference to place upper limits on the rate of coalescence of these binaries. We use developments made in the `PyCBC` search pipeline during Advanced LIGO and Virgo's second observing run to re-analyze Advanced LIGO's first observing run and re-estimate the statistical significance of LVT151012. We present sufficient evidence to claim LVT151012 as a gravitational wave event. In Advanced LIGO and Virgo's 2nd observing run a gravitational wave due to the merger of two binary neutron stars, known as GW170817, was discovered. We develop tools for Bayesian hypothesis testing so that we can investigate the interior dynamics of neutron stars using the GW170817 signal. Finally, we use Bayesian parameter estimation from `PyCBC` with tools of Bayesian hypothesis testing to investigate the presence of nonlinear tidal dynamics from a pressure – gravity mode instability in GW170817. We find that significant waveform degeneracies allow the effect of nonlinear tides to be compatible with the data at the level of nonsignificance (Bayes factor of unity). We also investigate further constraints on these nonlinear tides.

DETECTION AND INFERENCE IN GRAVITATIONAL WAVE ASTRONOMY

By

Steven Reyes

B.A., University of Chicago

Dissertation
Submitted in Partial Fulfillment of the Requirements
for the degree of
Doctor of Philosophy in Physics

Syracuse University
December 2019

ACKNOWLEDGEMENTS

I would like to first thank my advisor, Duncan Brown, for his support and mentorship throughout my PhD. Learning to think scientifically is a challenging endeavor and I am grateful for Duncan's guidance through the entire process. His experience and suggestions have been a constant help in attaining my research goals.

I am grateful to the entire Syracuse University Gravitational wave group for their unique perspectives and for their diverse expertise. Science, I believe, is best done in the context of diversity, community, and passion, and I believe the group deeply embodies this ethic in a way that few places could rival.

I am grateful to Peter Saulson and Stefan Ballmer for their helpful perspectives, critiques, and advice in the progress and development of the research I have conducted here at Syracuse University.

I would like to thank my coworkers and officemates: Thomas Vo, Jaysin Lord, TJ Massinger, Soumi De, Chris Biwer, Lorena Magaña Zertuche, Swetha Bhagwat, Daniel Finstad, Varun Srivastava, Chaitanya Afle, Derek Davis, Daniel Vander-Hyde, Nick Didio, Gaby Serna, Fabian Magaña-Sandoval, Ari Pederson, Eric Muñiz, and my virtual-officemate Alex Urban. Thank you for your friendship and constant encouragement. You all have been great fun to be around.

I would like to thank my committee members, Natalie Russo, Denver Whittington, Stefan Ballmer, Jay Hubisz, Walter Freeman, and Duncan Brown, for taking time out of their schedules to serve on my thesis defense committee.

To the other (under)graduate students at Syracuse University who have befriended me over the past years, I would like to thank you all as well. I hope you all continue to be excellent.

To my parents, my siblings, my cousins, my aunts, my uncles, and my entire family, you've given me every opportunity possible to become anything that I could ever wish to be. Thank you for everything.

Contents

List of Tables

List of Figures

Preface

Chapters 2 and Chapter 3 represent work that stems from my participation in the LIGO Scientific Collaboration (LSC). The work presented in this thesis does not reflect the scientific opinion of the LSC and it was not reviewed by the collaboration.

The content of Chapter 2 is taken from
Abbott, Benjamin P et al. "GW150914: First results from the search for binary black hole coalescence with Advanced LIGO." In:Physical Review D93.12 (2016), p. 122003.

The content of Chapter 3 is taken from
Benjamin P. Abbott et al., "Upper Limits on the Rates of Binary Neutron Star and NeutronStarblack Hole Mergers From Advanced Ligos First Observing run." In:Astrophys.J.832.2 (2016), p. L21.doi:10.3847/2041-8205/832/2/L21.

The content of Chapter 4 is taken from
Nitz, A. H., Capano, C., Nielsen, A. B., et al., "1-OGC: The first open gravitational-wave catalog of binary mergers from analysis of public Advanced LIGO data". In: (2019). Astrophys. J.872.195, doi: 10.3847/1538-4357/ab0108

The content of Chapter 6 is taken from
Steven Reyes and Duncan A. Brown., "Constraints on non-linear tides due top-gmodecoupling from the neutron-star merger GW170817". In: (2018). arXiv:1808.07013[astro-ph.HE]

To my family.

Chapter 1

The Beginning of Gravitational Wave Astronomy

On September 14, 2015 at 09:50:45 UTC the Advanced Laser Interferometer Gravitational wave Observatory (LIGO) detected a signal from the binary black hole merger GW150914 [21]. The initial detection of the event was made by low-latency searches for generic gravitational-wave transients [22]. LIGO reported the results of a matched-filter search using relativistic models of compact binary coalescence waveforms that recovered GW150914 with a false alarm rate less than $5 \times 10^{-6}\,\mathrm{yr}^{-1}$, establishing it as the first direct detection of gravitational waves from the merging of two black holes.

In LIGO's second observing run, the Livingston, LA and Hanford, WA observatories were joined by a third gravitational wave detector, Virgo. This gravitational wave network detected the gravitational wave signal from two merging binary neutron stars, GW170817 [23]. The signal, GW170817, was detected with a combined signal-to-noise ratio of 32.4 with a false alarm rate less than 10^6 years. The total mass of the binary system was estimated as $\sim 2.7\ M_{\odot}$ and at a luminosity distance of ~ 40 Mpc. The association with the gamma-ray burst GRB 170817A, detected by Fermi-GBM 1.7 seconds after the coalescence, confirms that GW170817 involved the merging of a binary neutron star and provides the first direct evidence of a link between these mergers with a neutron star and short, hard gamma-ray bursts [24]. Additional identifications of electromagnetic transient counterparts in the same location further supports the interpretation of this event as a neutron star merger [25]. This unprecedented joint gravitational and electromagnetic observation has provided incredible opportunities and insight into astrophysics, dense matter, gravitation, and cosmology.

In chapter 2, we introduce the PyCBC offline search analysis that was instrumental in the discoveries of GW150914 and GW170817. We describe the analysis at the time of LIGO's first observing run. The PyCBC search is a compact binary coalescence search [26, 27, 28, 29, 30, 31, 32, 33] that targets gravitational waves from binary neutron stars, binary black holes, and neutron star–black hole binaries, using matched filtering [34] with waveforms predicted by general relativity [35, 36, 37, 38, 39, 40, 41, 42, 43]. The PyCBC analysis correlates the detector data with template waveforms that model the expected signal. The analysis identifies candidate events that are detected at both observatories consistent with the 10 ms inter-site propagation time. Events are assigned a detection-statistic value that ranks their likelihood of being a gravitational-wave signal. This detection statistic is compared to the estimated detector noise background to determine the probability that a candidate event is due to detector noise. The probability that a gravitational wave candidates is due to detector noise is evaluated for the candidate event with the largest detection statistic. In the case that this probability is lower than $\sim 10^{-7}$, we remove the candidate event from the background analysis and recalculate the probability that the other gravitational wave candidates are due to detector noise. In LIGO's first observing run two gravitational waves from binary black hole mergers were discovered, GW150914 and GW151226.

During the first observing run there were no discoveries of gravitational waves from compact binaries that contained a neutron star [44]. Conditional on the non-detection of these signals, the LIGO and Virgo collaboration searches established upper limits on the rate of mergers of these signals. The non-detection of mergers from binary neutron stars and neutron star-black hole binaries during LIGO's first observing run had important implications on plausible astrophysical formation channels for these binaries, and on whether mergers of binaries containing neutron stars could still be considered plausible mechanisms for unexplained astrophysical phenomenon such as short, hard gamma-ray bursts. We describe the analysis techniques used to set the estimated upper limit merger rates for binary systems that contain a neutron star. We also presented estimates for future rate estimations for the subsequent second and third observing runs.

Since the publication of the results from LIGO's first observing run there was considerable development of gravitational wave astrophysical analysis techniques that permitted increased sensitivity in the PyCBC search analysis [45, 46, 47]. LIGO's second observing run which ran between November 30, 2016 and ended on August 25, 2017 and also involved the Advanced Virgo (Virgo) from August 1, 2017, onward, presented a useful testbed for these techniques. At the same time, LIGO made the gravitational wave strain data needed for analysis publicly available for the entire first-observing run the GW Open Science Center [48]. In Chapter 4, we present the results of a re-analysis of the publicly available data using the PyCBC search and we publish a full catalog of candidate gravitational wave events. We call this open catalog 1-OGC. The search was successful in re-estimating the statistical significance of LVT151012, which went from a marginal event to having a 97.6% probability of being of astrophysical origin. Thus we designate LVT151012 as GW151012.

In Chapter 5 of this dissertation we introduce advanced methods and tools for conducting Bayesian statistical analyses on gravitational wave data. In particular, we focus on Bayesian hypothesis testing and the advancements in many Markov-Chain Monte Carlo techniques for conducting Bayesian hypothesis testing. We introduce three distinct approaches for hypothesis testing, two based on a parallel tempering technique [49, 50, 51], and one based on testing nested models [52, 53]. We discuss how these techniques apply to gravitational wave astronomy.

In Chapter 6 we apply these Bayesian hypothesis testing tools to explore astroseismology using the binary neutron star merger GW170817. Recent studies have estimated the star's tidal deformability and placed constraints on the equation of state of the neutron stars [54, 55, 56, 57, 19, 58, 59, 60, 61, 62]. We explore a suggestion of [63] that the star's tidal deformation can induce nonresonant and nonlinear daughter wave excitations in p- and g-modes of the neutron stars via a quasi-static instability. This instability would remove energy from a binary system and possibly affect the phase evolution of the gravitational waves radiated during the inspiral. Ref. [64] claimed that the instability can rapidly drive modes to significant energies before the merger of the binary. The details of the instability saturation are unknown and so the size of the effect of the p-g mode coupling on the gravitational-waveform is not known [64]. We conduct parameter estimation on the GW170817 signal using parameters modeling the p-g mode instability. We report a Bayes factor of unity indicating a nonsignificant result. We find that modeling GW170817 with nonlinear tidal parameters create degeneracies in the other intrinsic parameters of the binary.

Chapter 2

GW150914 and the PyCBC Offline Search Analysis

2.1 Introduction

We report the results of a matched-filter search using relativistic models of compact binary coalescence waveforms that recovered GW150914 as the most significant event during the coincident observations between the two LIGO detectors from September 12 to October 20, 2015. This is a subset of the data from Advanced LIGO's first observational period that ended on January 12, 2016.

The binary coalescence search targets gravitational-wave emission from compact-object binaries with individual masses from $1\,\mathrm{M}_\odot$ to $99\,\mathrm{M}_\odot$, total mass less than $100\,\mathrm{M}_\odot$ and dimensionless spins up to 0.99. The search was performed using two independently implemented analyses, referred to as `PyCBC` [65, 66, 67] and `GstLAL` [68, 69, 70]. These analyses use a common set of template waveforms [71, 72, 73], but differ in their implementations of matched filtering [74, 75], their use of detector data-quality information [76], the techniques used to mitigate the effect of non-Gaussian noise transients in the detector [77, 68], and the methods for estimating the noise background of the search [66, 78]. In this dissertation we will focus on the analysis done by `PyCBC`.

GW150914 was observed in both LIGO detectors [79] within the 10 ms inter-site propagation time, with a combined matched-filter signal to noise ratio (SNR) of 24. The search reported a false alarm rate estimated to be less than 1 event per

203 000 years, equivalent to a statistical significance greater than $5.1\,\sigma$. The basic features of the GW150914 signal point to it being produced by the coalescence of two black holes [21]. The best-fit template parameters from the search are consistent with detailed parameter estimation that identifies GW150914 as a near-equal mass black hole binary system with source-frame masses $36^{+5}_{-4}\,\mathrm{M}_{\odot}$ and $29^{+4}_{-4}\,\mathrm{M}_{\odot}$ at the 90% credible level [2].

The second most significant candidate event in the observation period (referred to as LVT151012) was reported on October 12, 2015 at 09:54:43 UTC with a combined matched-filter SNR of 9.6. The search reported a false alarm rate of 1 per 2.3 years and a corresponding p-value of 0.02 for this candidate event. Detector characterization studies have not identified an instrumental or environmental artifact as causing this candidate event [76]. However, its p-value is not sufficiently low to confidently claim this candidate event as a signal. Detailed waveform analysis of this candidate event indicates that it is also a binary black hole merger with source frame masses $23^{+18}_{-5}\,\mathrm{M}_{\odot}$ and $13^{+4}_{-5}\,\mathrm{M}_{\odot}$, if it is of astrophysical origin.

This chapter is organized as follows: Sec. 2.2 gives an overview of the compact binary coalescence search and the methods used. Sec. 2.3 describes the construction and configuration of the analysis used in the search. Sec. 2.4 presents the results of the search, and follow-up of the two most significant candidate events, GW150914 and LVT151012.

2.2 Search Description

The binary coalescence search [26, 27, 28, 29, 30, 31, 32, 33] reported here targets gravitational waves from compact binary coalescences using matched filtering [34] with waveforms predicted by general relativity. These binary systems include binary neutron stars, binary black holes, and neutron star–black hole binaries The `PyCBC` analysis correlates the detector data with template waveforms that model the expected signal. The analysis identifies candidate events that are detected at both observatories consistent with the 10 ms inter-site propagation time. Events are assigned a detection statistic value that ranks their likelihood of being a gravitational-wave signal. This detection statistic is compared to the estimated detector noise background to determine the probability that detector noise could generate a candidate event with the same, or greater detection statistic.

We report on a search using coincident observations between the two Advanced LIGO detectors [80] in Hanford, Washington (H1) and in Livingston, Louisiana (L1) from September 12 to October $20, 2015$. During these 38.6 days, the detectors were in coincident operation for a total of 18.4 days. Unstable instrumental operation and hardware failures affected 20.7 hours of these coincident observations. These data are discarded and the remaining 17.5 days are used as input to the analyses [76]. The `PyCBC` analysis reduces this time further by imposing a minimum length over which the detectors must be operating stably. The approach of the `PyCBC` pipeline is described in Sec. 2.3. After applying this cut, the `PyCBC` analysis searched 16 days of coincident data. To prevent bias in the results, the configuration and tuning of the analyses were determined using data taken prior to September 12, 2015.

The gravitational waveform $h(t)$ depends on the chirp mass of the binary, $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$ [81, 82], the symmetric mass ratio $\eta = (m_1 m_2)/(m_1 + m_2)^2$ [83], and the angular momentum of the compact objects $\chi_{1,2} = c\vec{S}_{1,2}/Gm_{1,2}^2$ [84, 85] (the compact object's dimensionless spin), where $\vec{S}_{1,2}$ is the angular momentum of the compact objects. The effect of spin on the waveform depends also on the ratio between the component objects' masses. Parameters which affect the overall amplitude and phase of the signal as observed in the detector are maximized over in the matched-filter search, but can be recovered through full parameter estimation analysis [2]. The search parameter space is therefore defined by the limits placed on the compact objects' masses and spins. The minimum component masses of the search are determined by the lowest expected neutron star mass, which we assume to be $1\,\mathrm{M}_\odot$ [86]. There is no known maximum black hole mass [87], however we limit this search to binaries with a total mass less than $M = m_1 + m_2 \leq 100\,\mathrm{M}_\odot$. The LIGO detectors are sensitive to higher mass binaries, however we do not report on these searches in this chapter.

For binary component objects with masses less than $2\,\mathrm{M}_\odot$, we limit the magnitude of the component object's spin to 0.05, the spin of the fastest known pulsar in a double neutron star system [88]. At current detector sensitivity, this is sufficient to detect gravitational-wave signals from mergers of binaries with neutron star components having spins up to 0.4, the spin of the fastest-spinning millisecond pulsar [89]. Observations of X-ray binaries indicate that astrophysical black holes may have near extremal spins [90]. For binary components with masses larger than $2M_\odot$, we limit the spin magnitude to less than 0.9895. This is set by our ability to generate valid template waveforms at higher spins [71]. Figure 1 shows the boundaries of the search parameter space in the component-mass plane. We will investigate binary systems with neutron stars in Chapters 3 and 6.

Since the parameters of signals are not known in advance, each detector's output is filtered against a discrete bank of templates that span the search target space [27, 91, 92, 93, 94]. The placement of templates depends on the shape of the power spectrum of the detector noise. Both analyses use a low-frequency cutoff of 30 Hz for the search. The average noise power spectral density of the LIGO detectors was measured over the period September 12 to September 26, 2015. The harmonic mean of these noise spectra from the two detectors was used to place a single template bank that was used for the duration of the search [95, 66]. The templates are placed using a combination of geometric and stochastic methods [96, 97, 69, 73] such that the loss in matched-filter SNR caused by its discrete nature is $\lesssim 3\%$. Approximately 250,000 template waveforms are used to cover this parameter space, as shown in Fig. 1. The performance of the template bank is tested numerically by simulating binary black hole waveforms and determining the fraction of the total possible matched-filter SNR recovered for each simulated signal (the fitting factor) [98]. Figure 2 shows the resulting distribution of fitting factors obtained over the observation period. The loss in matched-filter SNR is less than 3% for more than 99% of the 10^5 simulated signals.

In addition to possible gravitational-wave signals, the detector strain contains a stationary noise background that primarily arises from photon shot noise at high frequencies and seismic noise at low frequencies. In the mid-frequency range, detector commissioning has not yet reached the point where test mass thermal noise dominates, and the noise at mid frequencies is poorly understood [79, 76, 99]. The detector strain data also exhibits non-stationarity and non-Gaussian noise transients that arise from a variety of instrumental or environmental mechanisms. The measured strain $s(t)$ is the sum of possible gravitational-wave signals $h(t)$ and the different types of detector noise $n(t)$.

To monitor environmental disturbances and their influence on the detectors, each observatory site is equipped with an array of sensors [100]. Auxiliary instrumental channels also record the interferometer's operating point and the state of the detector's control systems. Many noise transients have distinct signatures, visible in environmental or auxiliary data channels that are not sensitive to gravitational waves. When a noise source with known physical coupling between these channels and the detector strain data is active, a data-quality veto is created that is used to exclude these data from the search [76]. In the PyCBC analysis, these data quality vetoes are applied after filtering. A total of 2 hours is removed from the analysis by data quality vetoes. Despite these detector characterization investigations, the data still contains non-stationary and non-Gaussian noise which can affect the astrophysical sensitivity of the search. The PyCBC analysis implements methods to identify loud, short-duration noise transients and remove them from the strain data before filtering.

The PyCBC analysis calculates the matched-filter SNR for each template and each detector's data [74, 101]. In the PyCBC analysis, sources with total mass less than $4\,\mathrm{M}_{\odot}$ are modeled by computing the inspiral waveform accurate to 3.5 post-Newtonian order [83, 102, 103]. To model systems with total mass larger than $4\,\mathrm{M}_{\odot}$, we use templates based on the effective-one-body (EOB) formalism [38], which combines results from the Post-Newtonian approach [83, 103] with results from black hole perturbation theory and numerical relativity [71, 104] to model the complete inspiral, merger and ringdown waveform. The waveform models used assume that the spins of the merging objects are aligned with the orbital angular momentum. The analysis then identifies maxima of the matched-filter SNR (triggers) over the signal time of arrival.

To suppress large SNR values caused by non-Gaussian detector noise, the PyCBC analysis calculates additional signal consistency tests to quantify the agreement between the data and the template. The PyCBC analysis calculates a chi-squared statistic to test whether the data in different frequency bands are consistent with the matching template [77]. The value of the chi-squared statistic is used to compute a new detection statistic for each maxima. This detection statistic is called the re-weighted SNR or newSNR.

The PyCBC analysis enforces coincidence between detectors by selecting trigger pairs that occur within a 15 ms window and come from the same template. The 15 ms window is determined by the 10 ms inter-site propagation time plus 5 ms for uncertainty in arrival time of weak signals. The PyCBC analyses discards any triggers that occur during the time of data-quality vetoes prior to computing coincidence. The remaining coincident events are ranked based on the quadrature sum of the re-weighted SNR from both detectors [66].

The statistical significance of a candidate event is determined by the search background. This is the rate at which detector noise produces events with a detection statistic value equal to or higher than the candidate event (the false alarm rate). Estimating this background is challenging because the detector noise is non-stationary and non-Gaussian, so its properties must be empirically determined. The background estimation is also difficult because it is not possible to shield the detector from gravitational waves to directly measure a signal-free background.

To measure the statistical significance of candidate events, the PyCBC analysis artificially shifts the timestamps of one detector's triggers by an offset that is large compared to the inter-site propagation time, and a new set of coincident events is produced based on this time-shifted data set. For instrumental noise that is uncorrelated between detectors this is an effective way to estimate the background. To account for the search background noise varying across the target signal space, candidate and background events are divided into three search classes based on template length. To account for having searched multiple classes, the measured statistical significance is decreased by a trials factor equal to the number of classes [105]. This is is considered a conservative correction factor for Frequentist p-values.

The p-value of a candidate event evaluates the probability that detector noise could generate a detection statistic at the same level or greater than the candidate event's detection statistic. If this p-value falls below a certain pre-determined threshold the candidate is considered a potential gravitational wave signal.

2.3 PyCBC Detection Statistic and Statistical Significance Evaluation

The PyCBC analysis [65, 66, 67] uses fundamentally the same methods [106, 74, 77, 107, 108, 109, 110, 111, 112, 113] as those used to search for gravitational waves from compact binaries in the initial LIGO and Virgo detector era [114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125], with the improvements described in Refs. [65, 66].

In this Section, we describe the configuration of the PyCBC analysis used in this search. To prevent bias in the search result, the configuration of the analysis was determined using data taken prior to the observation period searched. When GW150914 was discovered by the low-latency transient searches [21], all tuning of the PyCBC analysis was frozen to ensure that the reported p-values are unbiased. No information from the low-latency transient search is used in this analysis.

Of the 17.5 days of data that are used as input to the analysis, the PyCBC analysis discards times for which either of the LIGO detectors is in their observation state for less than 2064 s; shorter intervals are considered to be unstable detector operation by the PyCBC analysis and are removed from the observation time. After discarding time removed by data-quality vetoes and periods when detector operation is considered unstable the observation time remaining is 16 days.

For each template $h(t)$ and for the strain data from a single detector $s(t)$, the analysis calculates the square of the matched-filter SNR defined by [74]

$$\rho^2(t) \equiv \frac{1}{\langle h|h\rangle} \left|\langle s|h\rangle(t)\right|^2 , \tag{2.1}$$

where the correlation is defined by

$$\langle s|h\rangle(t) = 4 \int_0^\infty \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)} e^{2\pi i f t} \,\mathrm{d}f , \tag{2.2}$$

where $\tilde{s}(f)$ is the Fourier transform of the time domain quantity $s(t)$ given by

$$\tilde{s}(f) = \int_{-\infty}^{\infty} s(t) e^{-2\pi i f t} \,\mathrm{d}t. \tag{2.3}$$

The quantity $S_n(|f|)$ is the one-sided average power spectral density of the detector noise, which is re-calculated every 2048 s (in contrast to the fixed spectrum used in template bank construction). Calculation of the matched-filter SNR in the frequency domain allows the use of the computationally efficient Fast Fourier Transform [126, 127]. The square of the matched-filter SNR in Eq. (2.1) is normalized by

$$\langle h|h\rangle = 4 \int_0^\infty \frac{\tilde{h}(f)\tilde{h}^*(f)}{S_n(f)} \,\mathrm{d}f , \tag{2.4}$$

so that its mean value is 2, if $s(t)$ contains only stationary noise [128].

Non-Gaussian noise transients in the detector can produce extended periods of elevated matched-filter SNR that increase the search background [66]. To mitigate this, a time-frequency excess power (burst) search [129] is used to identify high-amplitude, short-duration transients that are not flagged by data-quality vetoes. If the burst search generates a trigger with a burst SNR exceeding 300, the `PyCBC` analysis vetoes these data by zeroing out 0.5s of $s(t)$ centered on the time of the trigger. The data is smoothly rolled off using a Tukey window during the 0.25 s before and after the vetoed data. The threshold of 300 is chosen to be significantly higher than the burst SNR obtained from plausible binary signals. For comparison, the burst SNR of GW150914 in the excess power search is ~ 10. A total of 450 burst-transient vetoes are produced in the two detectors, resulting in 225 s of data removed from the search. A time-frequency spectrogram of the data at the time of each burst-transient veto was inspected to ensure that none of these windows contained the signature of an extremely loud binary coalescence.

The analysis places a threshold of 5.5 on the single-detector matched-filter SNR and identifies maxima of $\rho(t)$ with respect to the time of arrival of the signal. For each maximum we calculate a chi-squared statistic to determine whether the data in several different frequency bands are consistent with the matching template [77]. Given a specific number of frequency bands k, the value of the reduced χ^2, denoted as χ_r^2, is given by

$$\chi_r^2 = \frac{k}{2k-2}\frac{1}{\langle h|h\rangle}\sum_{i=1}^{k}\left|\langle s|h_i\rangle - \frac{\langle s|h\rangle}{k}\right|^2, \tag{2.5}$$

where h_i is the sub-template corresponding to the i-th frequency band. A reduced chi-squared statistic is defined as the chi-squared statistic divided by the number of degrees of freedom. Values of χ_r^2 near unity indicate that the signal is consistent with a coalescence. To suppress triggers from noise transients with large matched-filter SNR, $\rho(t)$ is re-weighted by [123, 108]

$$\hat{\rho} = \begin{cases} \rho \Big/ [(1+(\chi_r^2)^3)/2]^{\frac{1}{6}}, & \text{if } \chi_r^2 > 1, \\ \rho, & \text{if } \chi_r^2 \leq 1. \end{cases} \tag{2.6}$$

Triggers that have a re-weighted SNR $\hat{\rho} < 5$ or that occur during times subject to data-quality vetoes are discarded.

The template waveforms span a wide region of time-frequency parameter space and the susceptibility of the analysis to a particular type of noise transient can vary across the search space. This is demonstrated in Fig. 3 which shows the cumulative number of noise triggers as a function of re-weighted SNR for Advanced LIGO engineering run data taken between September 2nd and September 9th, 2015. The response of the template bank to noise transients is well characterized by the gravitational-wave frequency at the template's peak amplitude, f_{peak}. Waveforms with a lower peak frequency have fewer cycles in the detector's most sensitive frequency band from 30–2000 Hz [79, 99], and so are less easily distinguished from noise transients by the re-weighted SNR.

The number of bins in the χ^2 test is a tunable parameter in the analysis [66]. Previous searches used a fixed number of bins [130] with the most recent Initial LIGO and Virgo searches using $p = 16$ bins for all templates [123, 124]. Investigations on data from LIGO's sixth science run [131, 124] showed that better noise rejection is achieved with a template-dependent number of bins. The left two panels of Fig. 3 show the cumulative number of noise triggers with $k = 16$ bins used in the χ^2 test. Empirically, we find that choosing the number of bins according to

$$k = \lfloor 0.4(f_{\mathrm{peak}}/\mathrm{Hz})^{2/3} \rfloor \tag{2.7}$$

gives better suppression of noise transients in Advanced LIGO data, as shown in the right panels of Fig. 3. Here we use the notation $\lfloor \ldots \rfloor$ to denote the floor function.

The `PyCBC` analysis enforces signal coincidence between detectors by selecting trigger pairs that occur within a 15 ms window and come from the same template. We rank coincident events based on the quadrature sum $\hat{\rho}_c$ of the $\hat{\rho}$ from both detectors [66]. The final step of the analysis is to cluster the coincident events, by selecting those with the largest value of $\hat{\rho}_c$ in each time window of 10 s. Any other events in the same time window are discarded. This ensures that a loud signal or transient noise artifact gives rise to at most one candidate event [66].

The statistical significance of a candidate event is determined by the rate at which detector noise produces events with a detection statistic value equal to or higher than that of the candidate event. To measure this, the analysis creates a "background data set" by artificially shifting the timestamps of one detector's triggers by many multiples of 0.1 s and computing a new set of coincident events. Since the time offset used is always larger than the time-coincidence window, coincident signals do not contribute to this background. Under the assumption that noise is not correlated between the detectors [76], this method provides an unbiased estimate of the noise background of the analysis.

To account for the noise background varying across the target signal space, candidate and background events are divided into different search classes based on template length. Based on empirical tuning using Advanced LIGO engineering run data taken between September 2nd and September 9th, 2015, we divide the template space into three classes according to: (i) $\mathcal{M} < 1.74\,\mathrm{M}_\odot$; (ii) $\mathcal{M} \geq 1.74\,\mathrm{M}_\odot$ and $f_{\mathrm{peak}} \geq 220\,\mathrm{Hz}$; (iii) $\mathcal{M} \geq 1.74\,\mathrm{M}_\odot$ and $f_{\mathrm{peak}} < 220\,\mathrm{Hz}$. The statistical significance of candidate events is measured against the background from the same class. For each candidate event, we compute the p. This is the probability of finding one or more noise background events in the observation time with a detection statistic value above that of the candidate event, given by [66, 73]

$$p(\hat{\rho}_c) \equiv Pr(\geq 1 \text{ noise event above } \hat{\rho}_c \,|\, T, T_b) = 1 - \exp\left[-T\frac{1 + n_b(\hat{\rho}_c)}{T_b}\right], \qquad (2.8)$$

where T is the observation time of the search, T_b is the background time, and $n_b(\hat{\rho}_c)$ is the number of noise background triggers above the candidate event's re-weighted SNR $\hat{\rho}_c$.

Eq. (2.8) is derived assuming Poisson statistics for the counts of time-shifted background events, and for the count of coincident noise events in the search [66, 73]. This assumption requires that different time-shifted analyses (i.e. with different relative shifts between detectors) give independent realizations of a counting experiment for noise background events. We expect different time shifts to yield independent event counts since the 0.1 s offset time is greater than the 10 ms gravitational-wave travel time between the sites plus the ~ 1 ms autocorrelation length of the templates.

If a candidate event's detection statistic value is larger than that of any noise background event, as is the case for GW150914, then the `PyCBC` analysis places an upper bound on the candidate's p-value. After discarding time removed by data-quality vetoes and periods when the detector is in stable operation for less than 2064 seconds, the total observation time remaining is $T = 16$ days. Repeating the time-shift procedure $\sim 10^7$ times on these data produces a noise background analysis time equivalent to $T_b = 608\,000$ years. Thus, the smallest p-value that can be estimated in this analysis is approximately 7×10^{-8}. Since we treat the search parameter space as 3 independent classes, each of which may generate a false positive result, this value should be multiplied by a trials factor or look-elsewhere effect [105] of 3, resulting in a minimum measurable $p = 2 \times 10^{-7}$. The results of the `PyCBC` analysis are described in Sec. 2.4.

2.4 Search Results

GW150914 was observed on September 14, 2015 at 09:50:45 UTC as the most significant event by both analyses. The individual detector triggers from GW150914 occurred within the 10 ms inter-site propagation time with a combined matched-filter SNR of 24. The `PyCBC` analysis finds a matched-filter SNR for the individual detector triggers in the Hanford detector ($\rho_{\mathrm{H1}} = 20$) and the Livingston detector ($\rho_{\mathrm{L1}} = 13$). GW150914 was found with a template with component masses $47.9\,\mathrm{M}_\odot$ and $36.6\,\mathrm{M}_\odot$. The effective spin of the best-matching template is $\chi_{\mathrm{eff}} = (c/G)(\mathbf{S_1}/m_1 + \mathbf{S_2}/m_2) \cdot (\hat{\mathbf{L}}/M) = 0.2$, where $\mathbf{S_{1,2}}$ are the spins of the compact objects and $\hat{\mathbf{L}}$ is the direction of the binary's orbital angular momentum. Due to the discrete nature of the template bank, follow-up parameter estimation is required to accurately determine the best fit masses and spins of the binary's components [1, 2].

The frequency at peak amplitude of the best-matching template is $f_{\mathrm{peak}} = 144\,\mathrm{Hz}$, placing it in noise-background class (iii) of the `PyCBC` analysis. Figure 4 shows the result of the `PyCBC` analysis for this search class. In the time-shift analysis used to create the noise background estimate, a signal may contribute events to the background through random coincidences of the signal in one detector with noise in the other detector [73]. This can be seen in the background histogram shown by the black line. The tail is due to coincidence between the single-detector triggers from GW150914 and noise in the other detector. If a loud signal is in fact present, these random time-shifted coincidences contribute to an overestimate of the noise background and a more conservative assessment of the significance of an event. Figure 4 shows that GW150914 has a re-weighted SNR $\hat{\rho}_c = 23.6$, greater than all background events in its class. This value is also greater than all background in the other two classes. As a result, we can only place an upper bound on the false alarm rate, as described in Sec. 2.3. This bound is equal to the number of classes divided by the background time T_b. With 3 classes and $T_b = 608\,000$ years, we find the false alarm rate of GW150914 to be less than $5 \times 10^{-6}\,\mathrm{yr}^{-1}$. With an observing time of 384 hr, the $p\ < 2 \times 10^{-7}$.

We can convert this p-value to single-sided zero-mean, unit-variance Gaussian standard deviations using $z = -\sqrt{2}\,\mathrm{erf}^{-1}\left[1 - 2(1-p)\right]$, where erf^{-1} is the inverse error function. Here z denotes the z-score or standard score for the single-sided Gaussian standard deviation. The `PyCBC` analysis measures the significance of GW150914 as greater than $5.1\,\sigma$.

The difference in time of arrival between the Livingston and Hanford detectors from the individual triggers in the `PyCBC` analysis is 7.1 ms, consistent with the time delay of $6.9^{+0.5}_{-0.4}$ ms recovered by parameter estimation [2]. Figure 5 shows the matched-filter SNR ρ, the χ^2-statistic, and the re-weighted SNR $\hat{\rho}$ for the best-matching template over a period of ± 5 ms around the time of GW150914 (we take the `PyCBC` trigger time in L1 as a reference). The matched-filter SNR peaks in both detectors at the time of the event and the value of the reduced chi-squared statistic is $\chi^2_{\mathrm{H1}} = 1$ in the Hanford analysis and $\chi^2_{\mathrm{L1}} = 0.7$ in the Livingston analysis at the time of the event. This indicates a high match between the template and the data in both detectors. The re-weighted SNR of the individual detector triggers of $\hat{\rho}_{\mathrm{H1}} = 19.5$ and $\hat{\rho}_{\mathrm{L1}} = 13.3$ are larger than that of any other single-detector triggers in the analysis; therefore the significance measurement of $5.1\,\sigma$ set using the 0.1 s time shifts is a conservative bound on the p-value of GW150914.

The `PyCBC` analysis has shown that the probability of measuring a detection statistic as great or greater than GW150914's detection statistic due to a random coincidence of detector noise is extremely small. We therefore conclude that GW150914 is a gravitational-wave signal. To measure the signal parameters, we use parameter estimation methods that assume the presence of a coherent coalescing binary signal in the data from both detectors [1, 2]. Two waveform models were used which included inspiral, merger and ring-down portions of the signal: one which includes spin components aligned with orbital angular momentum [132, 104] and one which includes the dominant modulation of the signal due to orbital precession caused by mis-aligned spins [133, 134]. The parameter estimates are described by a continuous probability density function over the source parameters. We conclude that GW150914 is a nearly equal mass black-hole binary system of source-frame masses $36^{+5}_{-4}\ \mathrm{M}_{\odot}$ and $29^{+4}_{-4}\ \mathrm{M}_{\odot}$ (median and 90% credible range). The spin magnitude of the primary black hole is constrained to be less than 0.7 with 90% probability. The most stringent constraint on the spins of the two black holes is on the effective spin parameter $\chi_{\mathrm{eff}} = -0.06^{+0.17}_{-0.18}$.

The parameters of the best-fit template are consistent with these values, given the discrete nature of the template bank. We estimate GW150914 to be at a luminosity distance of 410^{+160}_{-180} Mpc, which corresponds to a redshift $0.09^{+0.03}_{-0.04}$. Full details of the source parameters for GW150914 are given in Ref. [2] and summarized in Table 1.

When an event is confidently identified as a real gravitational wave signal, as for GW150914, the background used to determine the significance of other events is re-estimated without the contribution of this event. This is the background distribution shown as purple lines in Fig. 4. Both analyses reported a candidate event on October 12, 2015 at 09:54:43 UTC as the second-loudest event in the observation period, which we refer to as LVT151012. This candidate event has a combined matched-filter SNR of 9.6. The `PyCBC` analysis reported a false alarm rate of 1 per 2.3 years and a corresponding p-value of 0.02 for this event. Detector characterization studies have not identified an instrumental or environmental artifact as causing this candidate event [76], however its p-value is not sufficiently low to confidently claim the event as a signal. It is significant enough to warrant follow-up, however. The results of signal parameter estimation, shown in Table 1, indicate that if LVT151012 is of astrophysical origin, then the source would be a stellar-mass binary black hole system with source-frame component masses $23^{+18}_{-5}\,\mathrm{M}_\odot$ and $13^{+4}_{-5}\,\mathrm{M}_\odot$. The effective spin would be $\chi_{\mathrm{eff}} = 0.0^{+0.3}_{-0.2}$ and the distance 1100^{+500}_{-500} Mpc.

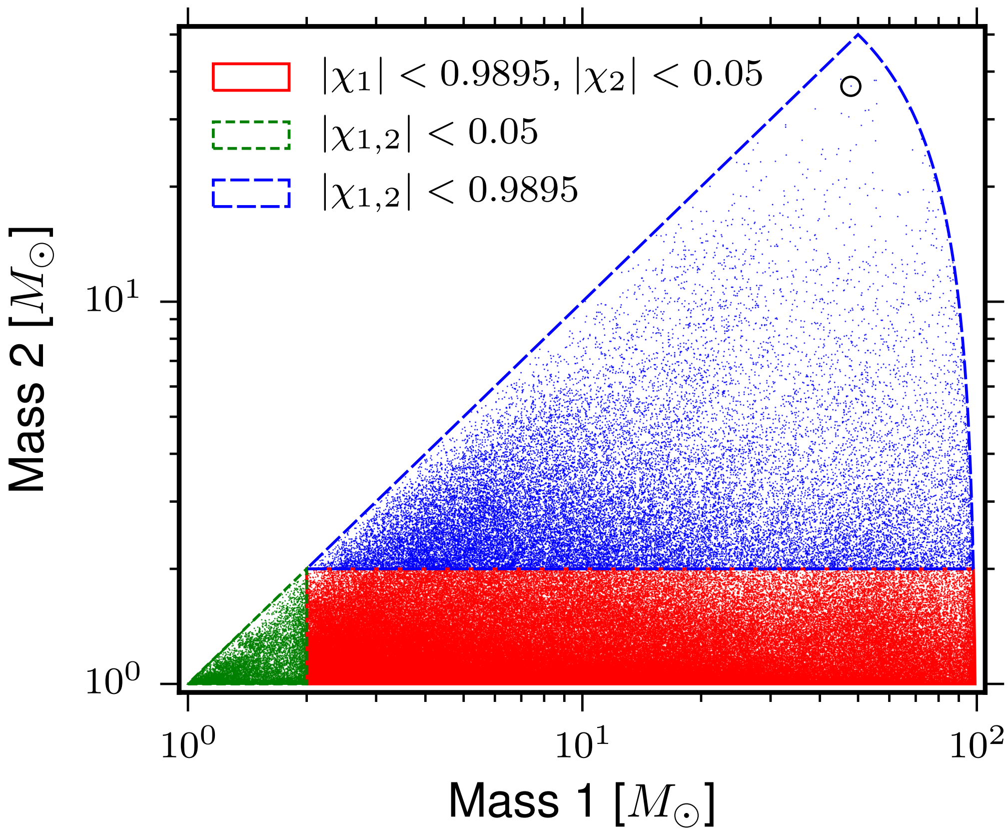


Figure 1: The four-dimensional search parameter space covered by the template bank shown projected into the component-mass plane, using the convention $m_1 > m_2$. The lines bound mass regions with different limits on the dimensionless aligned-spin parameters χ_1 and χ_2. Each point indicates the position of a template in the bank. The circle highlights the template that best matches GW150914. This does not coincide with the best-fit parameters due to the discrete nature of the template bank.

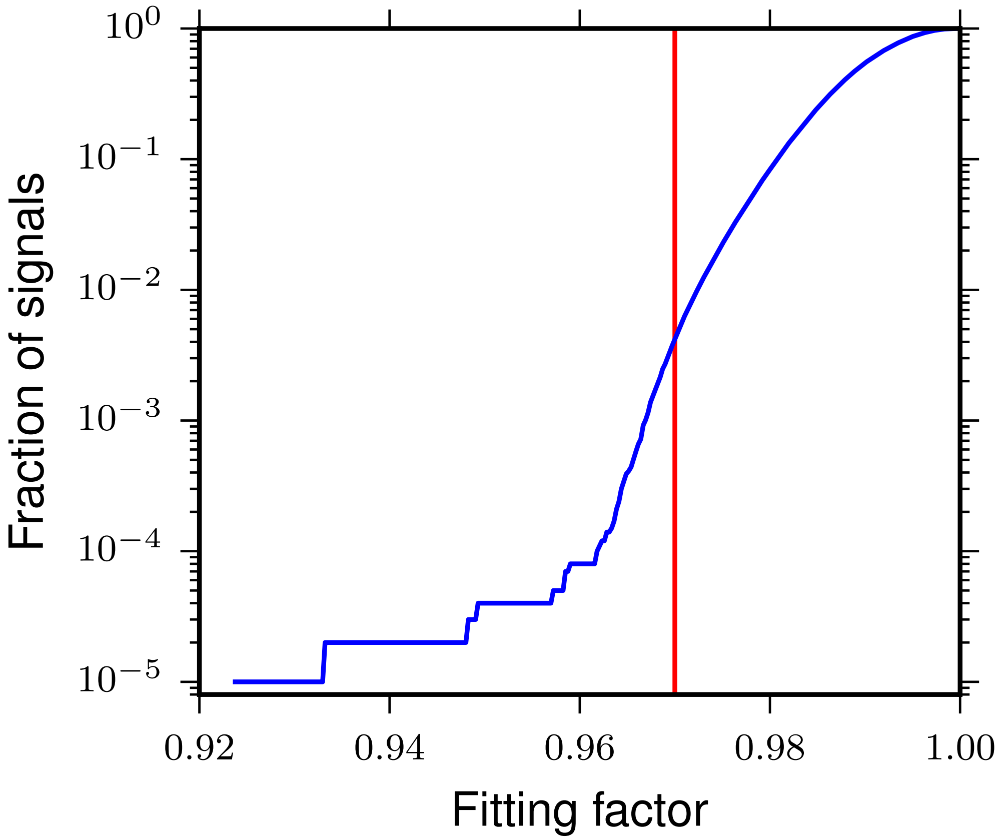
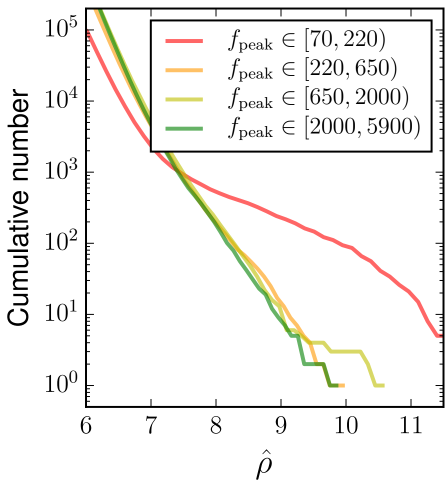
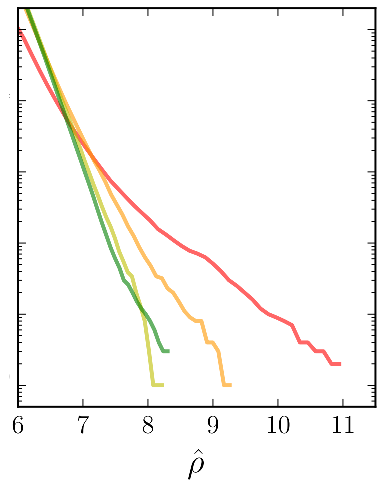


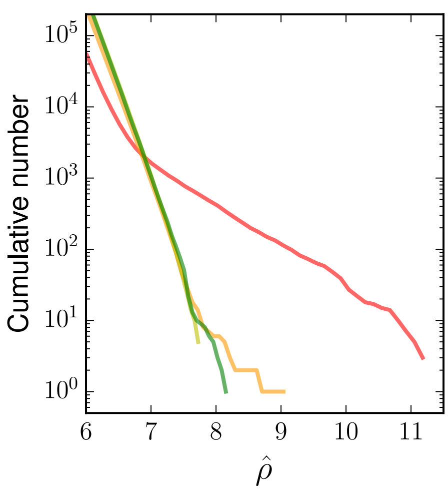
Figure 2: Cumulative distribution of fitting factors obtained with the template bank for a population of simulated aligned-spin binary black hole signals. The horizontal red line denotes that less than 1% of simulated signals have a matched-filter SNR loss greater than 3%. This demonstrates that the template bank has good coverage of the target search space.



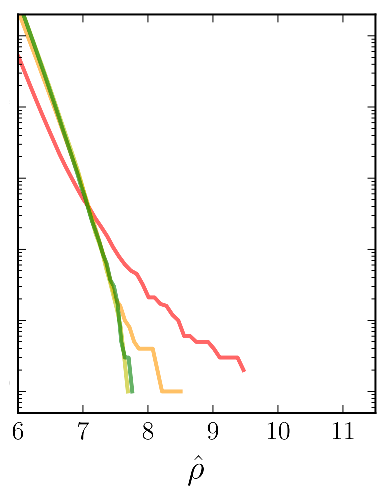
(a) H1, 16 χ^2 bins



(b) H1, optimized χ^2 bins



(c) L1, 16 χ^2 bins



(d) L1, optimized χ^2 bins

Figure 3: Distributions of noise triggers over re-weighted SNR $\hat{\rho}$, for Advanced LIGO engineering run data taken between September 2nd and September 9th, 2015. Each line shows triggers from templates within a given range of gravitational-wave frequency at maximum strain amplitude, $f_{\rm peak}$. Left: Triggers obtained from H1, L1 data respectively, using a fixed number of $p = 16$ frequency bands for the χ^2 test. Right: Triggers obtained with the number of frequency bands determined by the function $p = \lfloor 0.4 (f_{\rm peak}/\mathrm{Hz})^{2/3} \rfloor$. Note that while noise distributions are suppressed over the whole template bank with the optimized choice of p, the suppression is strongest for templates with lower $f_{\rm peak}$ values. Templates that have a $f_{\rm peak} < 220\,\mathrm{Hz}$ produce a large tail of noise triggers with high re-weighted SNR even with the improved χ^2-squared test tuning, thus we separate these templates from the rest of the bank when calculating the noise background.

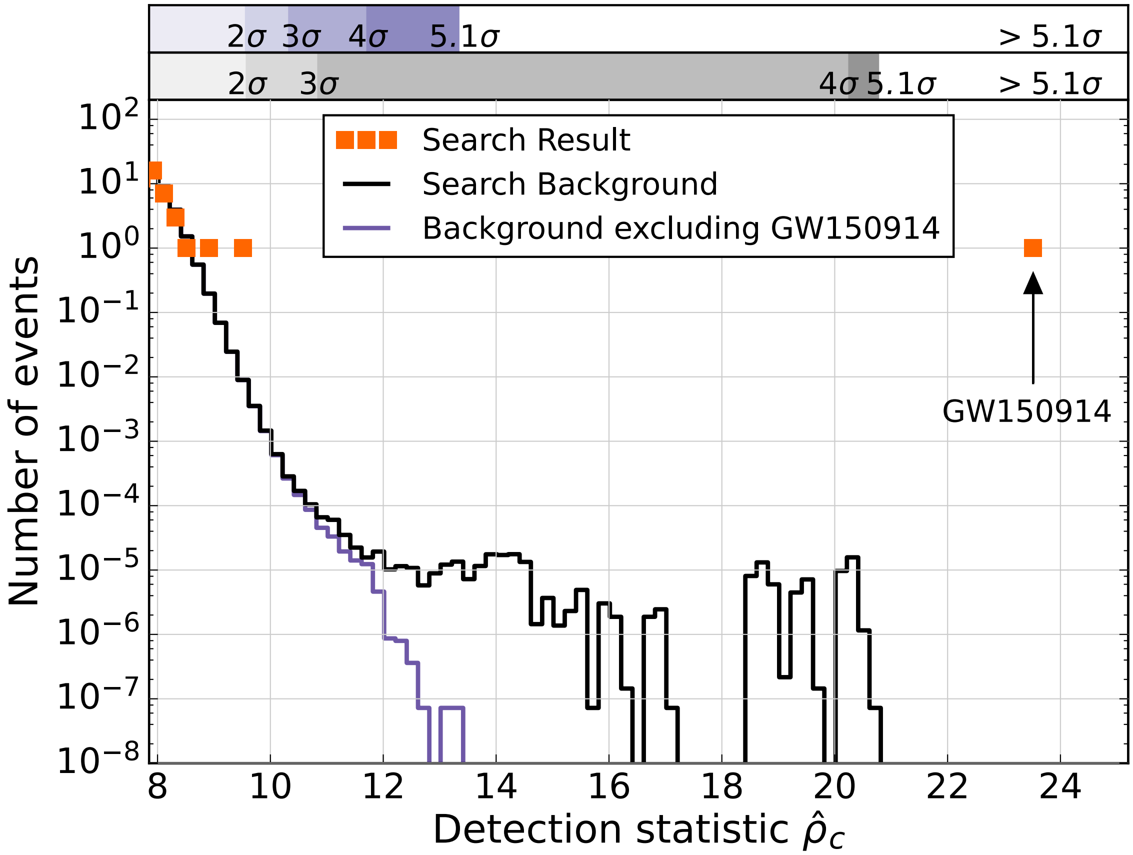


Figure 4: The classification scale for estimating the statististical significance of gravitational wave candidate events from the `PyCBC` analysis. The histogram shows the number of candidate events (orange) found in the analysis. The number of background events due to noise in the search class where GW150914 was found (black) as a function of the search detection statistic and with a bin width of $\Delta\hat{\rho}_c = 0.2$. The statistical significance of GW150914 is greater than 5.1 σ. The scales immediately above the histogram give the statistical significance of an event measured at a particular detection statistic against the noise backgrounds in units of Gaussian standard deviations. The black background histogram shows the result of the time-shift method to estimate the noise background in the observation period. The tail in the black-line background of the binary coalescence search is due to random coincidences of GW150914 in one detector with noise in the other detector. The statistical significance of GW150914 is measured against the upper gray scale. The purple background histogram is the background excluding coincidences involving GW150914 and it is the background to be used to assess the statistical significance of the second loudest event, LVT151012. The statistical significance of LVT151012 is measured against the upper purple scale.

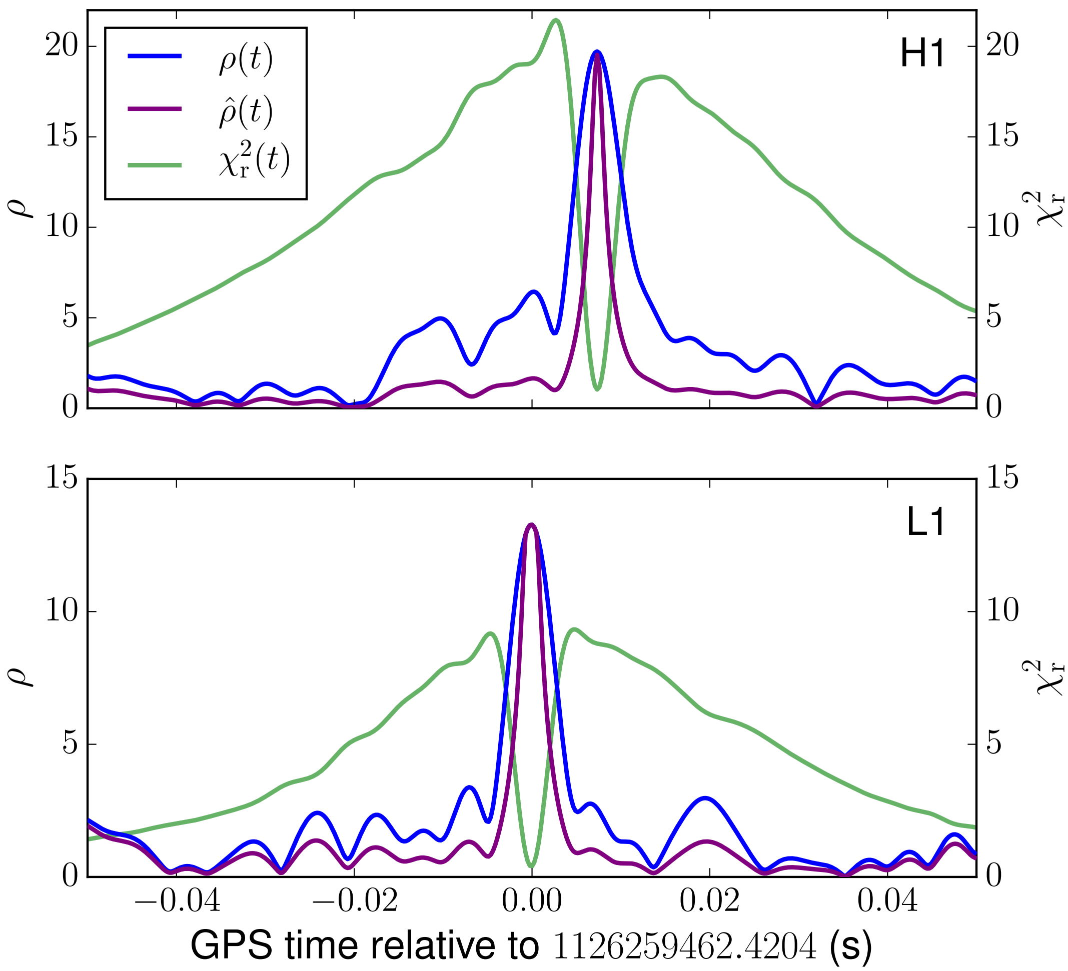


Figure 5: `PyCBC` matched-filter SNR (blue), re-weighted SNR (purple) and χ^2 (green) versus time of the best-matching template at the time of GW150914. The top plot shows the Hanford detector; bottom, Livingston. The SNR peaks at the event time of GW150914. The χ^2 consistency statistic tends towards unity at the event time for signals that match the expected signal morphology of a gravitational wave. The re-weighted SNR is peaks at the event time of of GW150914.

Event	Time (UTC)	FAR (yr^{-1})	p-value	$\mathcal{M}$ ($\mathrm{M}_\odot$)	χ_eff	D_L (Mpc)
GW150914	14 September 2015 09:50:45	$< 5 \times 10^{-6}$	$< 2 \times 10^{-7} (> 5.1\,\sigma)$	28^{+2}_{-2}	$-0.06^{+0.17}_{-0.18}$	410^{+160}_{-180}
LVT151012	12 October 2015 09:54:43	0.44	0.02 $(2.1\,\sigma)$	15^{+1}_{-1}	$0.0^{+0.3}_{-0.2}$	1100^{+500}_{-500}

Table 1: Parameters of the two most significant events. The false alarm rate (FAR) and p-value given here were determined by the `PyCBC` pipeline. The source-frame chirp mass $\mathcal{M}$, component masses $m_{1,2}$, effective spin χ_eff, and luminosity distance D_L are determined using a parameter estimation method that assumes the presence of a coherent compact binary coalescence signal starting at 20 Hz in the data [1]. The results are computed by averaging the posteriors for two model waveforms. Quoted uncertainties are 90% credible intervals that include statistical errors and systematic errors from averaging the results of different waveform models. Further parameter estimates of GW150914 are presented in Ref. [2]. In chapter 4 we will investigate how improvements to the `PyCBC` pipeline can improve the statistical significance estimate of LVT151012.

Chapter 3

Upper Limits on the Estimated Rate of Mergers of Binary Systems with a Neutron Star

3.1 Introduction

As described in Chapter 2, during Advanced LIGO's first observing run (O1), two high-mass binary black hole (BBH) events were identified with high confidence ($> 5\sigma$): GW150914 [135] and GW151226 [136]. In both cases the component masses are confidently constrained to be above the $3.2 M_{\odot}$ upper mass limit of NS set by theoretical considerations [137, 138]. The details of these observations, investigations about the properties of the observed BBH mergers, and the astrophysical implications are explored in [138, 139, 140, 141, 17?].

The search methods that successfully observed these BBH mergers also target other types of compact binary coalescences, specifically the inspiral and merger of binary neutron stars (BNS) systems and neutron star-black hole (NSBH) systems. Such systems were considered among the most promising candidates for an observation in O1. For example, a simple calculation prior to the start of O1 predicted 0.0005 - 4 detections of BNS signals during O1 [16].

In this chapter we report on the search for BNS and NSBH mergers in O1. We have searched for BNS systems with component masses $\in [1, 3] M_{\odot}$, component dimensionless spins < 0.05 and spin orientations aligned or anti-aligned with the orbital angular momentum. We have searched for NSBH systems with neutron star mass $\in [1, 3]M_\odot$, BH mass $\in [2, 99]M_\odot$ neutron star dimensionless spin magnitude < 0.05, BH dimensionless spin magnitude < 0.99 and both spins aligned or anti-aligned with the orbital angular momentum. No observation of either BNS or NSBH mergers was made in O1. We explore the astrophysical implications of this result, placing upper limits on the rates of such merger events in the local Universe that are roughly an order of magnitude smaller than those obtained with data from Initial LIGO and Initial Virgo [142, 143, 123]. We compare these updated rate limits to current predictions of BNS and NSBH merger rates and explore how the non-detection of BNS and NSBH systems in O1 can be used to explore possible constraints of the opening angle of the radiation cone of short gamma-ray bursts(GRB), assuming that short GRB progenitors are BNS or NSBH mergers.

The layout of this chapter is as follows. In Section 3.2 we describe the motivation for our search parameter space. In Section 3.3 we briefly describe the search methodology, then describe the results of the search in Section 3.4. We then discuss the constraints that can be placed on the rates of BNS and NSBH mergers in Section 3.5 and the astrophysical implications of the rates in Section 3.6.

3.2 Source considerations

There are currently thousands of known NSs, most detected as pulsars [144, 145]. Of these, ~ 70 are found in binary systems and allow estimates of the NS mass [146, 147, 148]. Published mass estimates range from $1.0 \pm 0.17\,M_\odot$ [149] to $2.74 \pm 0.21\,M_\odot$ [150] although there is some uncertainty in some of these measurements. Considering only precise mass measurements from these observations one can set a lower bound on the maximum possible neutron star mass of $2.01 \pm 0.04\,M_\odot$ [151] and theoretical considerations set an upper bound on the maximum possible neutron star mass of 2.9–3.2 $M_\odot$ [137, 152]. The standard formation scenario of core-collapse supernovae restricts the birth masses of neutron stars to be above 1.1–1.6 $M_\odot$ [153, 147, 154].

Eight candidate BNS systems allow mass measurements for individual components, giving a much narrower mass distribution [89]. Masses are reported between $1.0\,M_\odot$ and $1.49\,M_\odot$ [146, 148], and are consistent with an underlying mass distribution of $(1.35 \pm 0.13)\,M_\odot$ [155]. These observational measurements assume masses are greater than $0.9 M_{\odot}$.

The fastest spinning pulsar observed so far rotates with a frequency of 716 Hz [156]. This corresponds to a dimensionless spin $\chi = c|\vec{S}|/Gm^2$ of roughly 0.4, where m is the object's mass and $\vec{S}$ is the angular momentum.[1] Such rapid rotation rates likely require the NS to have been spun up through mass-transfer from its companion. The fastest spinning pulsar in a confirmed BNS system has a spin frequency of 44 Hz [157], implying that dimensionless spins for NS in BNS systems are ≤ 0.04 [97]. However, recycled NS can have larger spins, and the potential BNS pulsar J1807-2500B [158] has a period of 4.19 ms, giving a dimensionless spin of up to ~ 0.2.[2]

Given these considerations, we search for BNS systems with both masses $\in [1, 3] M_{\odot}$ and component dimensionless spins < 0.05. We have found that BNS systems with spins < 0.4 are generally still recovered well even though they are not explicitly covered by our search space. Increasing the search space to include BNS systems with spins < 0.4 was found to not improve overall search sensitivity [159].

NSBH systems are thought to be efficiently formed in one of two ways: either through the stellar evolution of field binaries or through dynamical capture of a NS by a BH [160, 161, 162, 163]. Though no NSBH systems are known to exist, one likely progenitor has been observed, Cyg X-3 [164].

Measurements of galactic stellar mass BH in X-ray binaries yield BH masses $5 \leq M_{\mathrm{BH}}/M_{\odot} \leq 24$ [165, 166, 167, 168]. Extragalactic high-mass X-ray binaries, such as IC10 X-1 and NGC300 X-1 suggest BH masses of $20 - 30\, M_{\odot}$. Advanced LIGO has observed two definitive BBH systems and constrained the masses of the 4 component BH to $36^{+5}_{-4}, 29^{+4}_{-4}, 14^{+8}_{-4}$ and $7.5^{+2.3}_{-2.3}\, M_{\odot}$, respectively, and the masses of the two resulting BH to 62^{+4}_{-4} and $21^{+6}_{-2}\, M_{\odot}$. In addition if one assumes that the candidate BBH merger LVT151012 was of astrophysical origin than its component BH had masses constrained to 23^{+16}_{-6} and 13^{+4}_{-5} with a resulting BH mass of 35^{+14}_{-4}. There is an apparent gap of BH in the mass range 3–5 $M_{\odot}$, which has been ascribed to the supernova explosion mechanism [169, 170]. However, BH formed from stellar evolution may exist with masses down to $2\, M_{\odot}$, especially if they are formed from matter accreted onto neutron stars [171]. Population synthesis models typically allow

[1]Assuming a mass of $1.4 M_{\odot}$ and a moment of inertia $= J/\Omega$ of $1.5 \times 10^{45}\,\mathrm{g\,cm^2}$; the exact moment of inertia is dependent on the unknown NS equation-of-state [147].

[2]Calculated with a pulsar mass of $1.37 M_{\odot}$ and a high moment of inertia, $2 \times 10^{45}\,\mathrm{g\,cm^2}$.

for stellar-mass BH up to $\sim$ 80–100 $M_\odot$ [170, 172, 173]; stellar BH with mass above 100 $M_\odot$ are also conceivable however [87, 14].

X-ray observations of accreting BH indicate a broad distribution of BH spin [174, 175, 176, 177, 178, 179, 180, 86]. Some BH observed in X-ray binaries have very large dimensionless spins (e.g Cygnus X-1 at > 0.95 [181, 182]), while others could have much lower spins (~ 0.1) [183]. Measured BH spins in high-mass X-ray binary systems tend to have large values (> 0.85), and these systems are more likely to be progenitors of NSBH binaries [90]. Isolated BH spins are only constrained by the relativistic Kerr bound $\chi \leq 1$ [184]. LIGO's observations of merging binary BH systems yield weak constraints on component spins [138, 136, 17]. The microquasar XTE J1550-564 [185] and population synthesis models [186] indicate small spin-orbit misalignment in field binaries. Dynamically formed NSBH systems, in contrast, are expected to have no correlation between the spins and the orbit.

We search for NSBH systems with NS mass $\in [1, 3]M_\odot$, NS dimensionless spins < 0.05, BH mass $\in [2, 99]M_\odot$ and BH spin magnitude < 0.99. Current search techniques are restricted to waveform models where the spins are (anti-)aligned with the orbit [187, 66], although methods to extend this to generic spins are being explored [188]. Nevertheless, aligned-spin searches have been shown to have good sensitivity to systems with generic spin orientations in O1 [189, 188]. An additional search for BBH systems with total mass greater than 100 $M_\odot$ is also being performed, the results of which will be reported in a future publication.

3.3 Search Description

To observe compact binary coalescences in data taken from Advanced LIGO we use matched-filtering against models of compact binary merger GW signals [190]. As the emitted GW signal varies significantly over the range of masses and spins in the BNS and NSBH parameter space, the matched-filtering process must be repeated over a large set of filter waveforms, or "template bank" [92]. The ranges of masses considered in the searches are shown in Figure 6. Statistical significance of potential events are produced by the `PyCBC` search pipeline as outlined in chapter 2.

BNS and NSBH mergers are prime candidates not only for observation with GW facilities, but also for coincident observation with EM observatories [191, 192, 193, 194, 195, 196, 197, 198, 199, 8]. We have a long history of working with the Fermi, Swift and IPN GRB teams to perform sub-threshold searches of GW data in a narrow window around the time of observed GRB [200, 201, 202, 203]. Such a search is currently being performed on O1 data and will be reported in a forthcoming publication. In O1 we also aimed to rapidly alert EM partners if a GW observation was made [204]. Therefore it was critical for us to run "online" searches to identify potential BNS or NSBH mergers within a timescale of minutes after the data is taken, to give EM partners the best chance to perform a coincident observation.

Nevertheless, analyses running with minute latency do not have access to full data-characterization studies, which can take weeks to perform, or to data with the most complete knowledge about calibration and associated uncertainties. Additionally, in rare instances, online analyses may fail to analyse stretches of data due to computational failure. Therefore it is also important to have an "offline" search, which performs the most sensitive search possible for BNS and NSBH sources. We give here a brief description of both the offline and online searches, referring to other works to give more details when relevant.

3.3.1 `PyCBC` Offline Search

The offline CBC search of the O1 data set consists of two independently-implemented matched-filter analyses: `GstLAL` [187] and `PyCBC` [66]. Full details of the `PyCBC` offline search pipeline are described in chapter 2.

In contrast to the online search, the offline search uses data produced with smaller calibration errors [205], uses complete information about the instrumental data quality [206] and ensures that all available data is analysed. The offline search in O1 forms a single search targeting BNS, NSBH, and BBH systems. The waveform filters cover systems with individual component masses ranging from 1 to 99 $M_\odot$, total mass constrained to less than 100 $M_\odot$ (see Figure 6), and component dimensionless spins up to $\pm$ 0.05 for components with mass less than 2 $M_\odot$ and $\pm$ 0.99 otherwise [17, 207]. Waveform filters with total mass less than 4 $M_\odot$ (chirp mass less than $1.73 M_\odot$[3]) for `PyCBC` (`GstLAL`) are modeled with the inspiral-only, post-Newtonian, frequency-domain approximant "TaylorF2" [208, 209, 210, 211, 212]. At larger masses it becomes

[3]The "chirp mass" is the combination of the two component masses that LIGO is most sensitive to, given by $\mathcal{M} = (m_1 m_2)^{3/5}(m_1 + m_2)^{-1/5}$, where m_i denotes the two component masses

important to also include the merger and ringdown components of the waveform. There a reduced-order model of the effective-one-body waveform calibrated against numerical relativity is used [71, 72].

3.3.2 Dataset

Advanced LIGO's first observing run occurred between September 12, 2015 and January 12, 2016 and consists of data from the two LIGO observatories in Hanford, WA and Livingston, LA. The LIGO detectors were running stably with roughly 40% coincident operation, and had been commissioned to roughly a third of the design sensitivity by the time of the start of O1 [213]. During this observing run the final offline dataset consisted of 76.7 days of analyzable data from the Hanford observatory, and 65.8 days of data from the Livingston observatory. We analyze only times during which *both* observatories took analyzable data, which is 49.0 days. Characterization studies of the analysable data found 0.5 days of coincident data during which time there was some identified instrumental problem—known to introduce excess noise—in at least one of the interferometers [206]. These times are removed before assessing the significance of events in the remaining analysis time. Some additional time is not analysed because of restrictions on the minimal length of data segments and because of data lost at the start and end of those segments [214, 17]. These requirements are slightly different between the two offline analyses, `PyCBC` and `GstLAL` . The `PyCBC` pipeline analysed 46.1 days of data.

3.4 Search Results

The offline search, targeting BBH as well as BNS and NSBH mergers, identified two signals with $> 5\sigma$ confidence in the O1 dataset [135, 136]. A third signal was also identified with 1.7σ confidence [17, 214]. Subsequent parameter inference on all three of these events has determined that, to very high confidence, they were not produced by a BNS or NSBH merger [138, 17]. No other events are significant with respect to the noise background in the offline search [17], and we therefore state that no BNS or NSBH mergers were observed.

3.5 Rates

3.5.1 Calculating upper limits

Given no evidence for BNS or NSBH coalescences during O1, we seek to place an upper limit on the astrophysical rate of such events. The expected number of observed events Λ in a given analysis can be related to the astrophysical rate of coalescences for a given source R by

$$\Lambda = R\langle VT\rangle. \tag{3.1}$$

Here, $\langle VT\rangle$ is the space-time volume that the detectors are sensitive to—averaged over space, observation time, and the parameters of the source population of interest. The likelihood for finding zero observations in the data s follows the Poisson distribution for zero events $p(\mathbf{s}\,|\,\Lambda) = e^{-\Lambda}$. We use the notation of $\mathcal{L}(\mathbf{s}\,|\,\Lambda)$ for likelihoods. Bayes' theorem then gives the posterior for Λ

$$p(\Lambda\,|\,\mathbf{s}) \propto \pi(\Lambda)e^{-\Lambda}, \tag{3.2}$$

where $p(\Lambda)$ is the prior on Λ, which we will denote as $\pi(\Lambda)$. We also switch notation for posteriors to $\mathcal{P}(\Lambda\,|\,\mathbf{s})$.

Searches of Initial LIGO and Initial Virgo data used a uniform prior on Λ [123] but included prior information from previous searches. For the O1 BBH search, however, a Jeffreys prior of $\pi(\Lambda) \propto 1/\sqrt{\Lambda}$ for the Poisson likelihood was used [215, 139, 17]. A Jeffreys prior has the convenient property that the resulting posterior is invariant under a change in parametrization. However, for consistency with past BNS and NSBH results we will primarily use a uniform prior, and note that a Jeffreys prior generally predicts a rate upper limit that is $\sim 40\%$ smaller. We do not include additional prior information because the sensitive $\langle VT\rangle$ from all previous runs is an order of magnitude smaller than that of O1. We estimate $\langle VT\rangle$ by adding a large number of simulated waveforms sampled from an astrophysical population into the data. These simulated signals are recovered with an estimate of the FAR using the offline analyses. Monte-Carlo integration methods are then utilized to estimate the sensitive volume to which the detectors can recover gravitational-wave signals below a chosen FAR threshold, which in this paper we will choose to be $0.01\mathrm{yr}^{-1}$. This threshold is low enough that only signals that are likely to be true events are counted as found, and we note that varying this threshold in the range 0.0001–1 yr^{-1} only changes the calculated $\langle VT \rangle$ by about $\pm 20\%$.

Calibration uncertainties lead to a difference between the amplitude of simulated waveforms and the amplitude of real waveforms with the same luminosity distance d_L. During O1, the 1σ uncertainty in the strain amplitude was 6%, resulting in an 18% uncertainty in the measured $\langle VT \rangle$. Results presented here also assume that injected waveforms are accurate representations of astrophysical sources. We use a time-domain, aligned-spin, post-Newtonian point-particle approximant to model BNS injections [216], and a time-domain, effective-one-body waveform calibrated against numerical relativity to model NSBH injections [217, 71]. Waveform differences between these models and the offline search templates are therefore included in the calculated $\langle VT \rangle$. The injected NSBH waveform model is not calibrated at high mass ratios ($m_1/m_2 > 8$), so there is some additional modeling uncertainty for large-mass NSBH systems. The true sensitive volume $\langle VT \rangle$ will also be smaller if the effect of tides in BNS or NSBH mergers is extreme. However, for most scenarios the effects of waveform modeling will be smaller than the effects of calibration errors and the choice of prior discussed above.

The posterior on Λ (Eq. 3.2) can be reexpressed as a joint posterior on the astrophysical rate R and the sensitive volume $\langle VT \rangle$

$$\mathcal{P}(R, \langle VT \rangle \,|\, \mathbf{s}) \propto \pi(R, \langle VT \rangle)\, e^{-R\langle VT \rangle}. \tag{3.3}$$

The new prior can be expanded as $\pi(R, \langle VT \rangle) = \pi(R \,|\, \langle VT \rangle)\pi(\langle VT \rangle)$. For $\pi(R \,|\, \langle VT \rangle)$, we will either use a uniform prior on R or a prior proportional to the Jeffreys prior $1/\sqrt{R\langle VT \rangle}$. As with [139**?** , 17], we use a log-normal prior on $\langle VT \rangle$

$$\pi(\langle VT \rangle) = \ln \mathcal{N}(\mu, \sigma^2), \tag{3.4}$$

where μ is the calculated value of $\ln\langle VT \rangle$ and σ represents the fractional uncertainty in $\langle VT \rangle$. Below, we will use an uncertainty of $\sigma = 18\%$ due mainly to calibration errors.

Finally, a posterior for the rate is obtained by marginalizing over $\langle VT \rangle$

$$\mathcal{P}(R \,|\, \mathbf{s}) = \int d\langle VT \rangle \, \mathcal{P}(R, \langle VT \rangle \,|\, \mathbf{s}). \tag{3.5}$$

The upper limit R_c on the rate with confidence c is then given by the solution to

$$\int_0^{R_c} dR\, \mathcal{P}(R\,|\,\mathbf{s}) = c. \tag{3.6}$$

For reference, we note that in the limit of zero uncertainty in $\langle VT\rangle$, the uniform prior for $\pi(R|\langle VT\rangle)$ gives a rate upper limit of

$$R_c = \frac{-\ln(1-c)}{\langle VT\rangle}, \tag{3.7}$$

corresponding to $R_{90\%} = 2.303/\langle VT\rangle$ for a 90% confidence upper limit [218]. For a Jeffreys prior on $\pi(R|\langle VT\rangle)$, this upper limit is

$$R_c = \frac{[\mathrm{erf}^{-1}(c)]^2}{\langle VT\rangle}, \tag{3.8}$$

corresponding to $R_{90\%} = 1.353/\langle VT\rangle$ for a 90% confidence upper limit.

3.5.2 BNS rate limits

Motivated by considerations in Section 3.2, we begin by considering a population of BNS sources with a narrow range of component masses sampled from the normal distribution $\mathcal{N}(1.35M_\odot, (0.13M_\odot)^2)$ and truncated to remove samples outside the range $[1, 3]M_\odot$. We consider both a "low spin" BNS population, where spins are distributed with uniform dimensionless spin magnitude $\in [0, 0.05]$ and isotropic direction, and a "high spin" BNS population with a uniform dimensionless spin magnitude $\in [0, 0.4]$ and isotropic direction. Our population uses an isotropic distribution of sky location and source orientation and chooses distances assuming a uniform distribution in volume. These simulations are modeled using a post-Newtonian waveform model, expanded using the "TaylorT4" formalism [216]. From this population we compute the space-time volume that Advanced LIGO was sensitive to during the O1 observing run. Results are shown for the measured $\langle VT\rangle$ in Table 2 using a detection threshold of FAR $= 0.01\,\mathrm{yr}^{-1}$. Because the template bank for the searches use only aligned-spin BNS templates with component spins up to 0.05, the `PyCBC` pipeline is 4% more sensitive to the low-spin population than to the high-spin population. The difference in $\langle VT\rangle$ between the two analyses is no larger than 5%, which is consistent with the difference in time analyzed in the two analyses. In addition, the calculated $\langle VT\rangle$ has a Monte Carlo integration uncertainty of $\sim 1.5\%$ due to the finite number of injection samples.

Using the measured $\langle VT \rangle$, the rate posterior and upper limit can be calculated from Eqs. 3.5 and 3.6 respectively. The posterior and upper limits are shown in Figure 7 and depend sensitively on the choice of uniform versus Jeffreys prior for $\Lambda = R\langle VT \rangle$. However, they depend only weakly on the spin distribution of the BNS population and on the width σ of the uncertainty in $\langle VT \rangle$. For the conservative uniform prior on Λ and an uncertainty in $\langle VT \rangle$ due to calibration errors of 18%, we find the 90% confidence upper limit on the rate of BNS mergers to be 12,100 $\mathrm{Gpc}^{-3}\ \mathrm{yr}^{-1}$ for low spin and 12,600 $\mathrm{Gpc}^{-3}\ \mathrm{yr}^{-1}$ for high spin using the values of $\langle VT \rangle$ calculated with `PyCBC`. These numbers can be compared to the upper limit computed from analysis of Initial LIGO and Initial Virgo data [123]. There, the upper limit for $1.35 - 1.35 M_\odot$ non-spinning BNS mergers is given as 130,000 $\mathrm{Gpc}^{-3}\ \mathrm{yr}^{-1}$. The O1 upper limit is more than an order of magnitude lower than this previous upper limit.

To allow for uncertainties in the mass distribution of BNS systems we also derive 90% confidence upper limits as a function of the NS component masses. To do this we construct a population of software injections with component masses sampled uniformly in the range $[1, 3]M_\odot$, and an isotropic distribution of component spins with magnitudes uniformly distributed in $[0, 0.05]$. We then bin the BNS injections by mass, and calculate $\langle VT \rangle$ and the associated 90% confidence rate upper limit for each bin. The 90% rate upper limit for the conservative uniform prior on Λ as a function of component masses is shown in Figure 8 for `PyCBC`. The fractional difference between the `PyCBC` and `GstLAL` results range from 1% to 16%.

3.5.3 NSBH rate limits

Given the absence of known NSBH systems and uncertainty in the BH mass, we evaluate the rate upper limit for a range of BH masses. We use three masses that span the likely range of BH masses: $5M_\odot$, $10M_\odot$, and $30M_\odot$. For the NS mass, we use the canonical value of $1.4M_\odot$. We assume a distribution of BH spin magnitudes uniform in $[0, 1]$ and NS spin magnitudes uniform in $[0, 0.04]$. For these three mass pairs, we compute upper limits for an isotropic spin distribution on both bodies, and for a case where both spins are aligned or anti-aligned with the orbital angular momentum (with equal probability of aligned vs anti-aligned). Our NSBH population uses an isotropic distribution of sky location and source orientation and chooses distances assuming a uniform distribution in volume. Waveforms are modeled using the spin-precessing, effective-one-body model calibrated against numerical relativity waveforms described in [71, 219].

The measured $\langle VT \rangle$ for a FAR threshold of 0.01yr^{-1} is given in Table 3 for PyCBC. The uncertainty in the Monte Carlo integration of $\langle VT \rangle$ is 1.5%–2%. The corresponding 90% confidence upper limits are also given using the conservative uniform prior on Λ and an 18% uncertainty in $\langle VT \rangle$. Analysis-specific differences in the limits range from 1% to 20%, comparable or less than other uncertainties such as calibration. These results can be compared to the upper limits found for initial LIGO and Virgo for a population of $1.35M_{\odot}$–$5M_{\odot}$ NSBH binaries with isotropic spin of 36,000 $\text{Gpc}^{-3}\ \text{yr}^{-1}$at 90% confidence [123]. As with the BNS case, this is an improvement in the upper limit of over an order of magnitude.

We also plot the 50% and 90% confidence upper limits from PyCBC and GstLAL as a function of mass in Figure 9 for the uniform prior. The search is less sensitive to isotropic spins than to (anti-)aligned spins due to two factors. First, the volume-averaged signal power is larger for a population of (anti-)aligned spin systems than for isotropic-spin systems. Second, the search uses a template bank of (anti-)aligned spin systems, and thus loses sensitivity when searching for systems with significantly misaligned spins. As a result, the rate upper limits are less constraining for the isotropic spin distribution than for the (anti-)aligned spin case.

3.6 Astrophysical Interpretation

We can compare our upper limits with rate predictions for compact object mergers involving NS, shown for BNS in Figure 10 and for NSBH in Figure 11. A wide range of predictions derived from population synthesis and from binary pulsar observations were reviewed in 2010 to produce rate estimates for canonical $1.4\,M_{\odot}$ NS and $10\,M_{\odot}$ BH [6]. We additionally include some more recent estimates from population synthesis for both NSBH and BNS [15, 220, 14] and binary pulsar observations for BNS [7].

We also compare our upper limits for NSBH and BNS systems to beaming-corrected estimates of short GRB rates in the local universe. Short GRB are considered likely to be produced by the merger of compact binaries that include NS, i.e.

BNS or NSBH systems [198]. The rate of short GRB can predict the rate of progenitor mergers [10, 11, 9, 8]. For NSBH, systems with small BH masses are considered more likely to be able to produce short GRB (e.g. [221, 222, 223]), so we compare to our $5M_{\odot}$–$1.4M_{\odot}$ NSBH rate constraint. The observation of a kilonova is also considered to be an indicator of a binary merger [196], and an estimated kilonova rate gives an additional lower bound on compact binary mergers [12]. The discovery of a BNS merger in O2 [54] in association with a gamma-ray burst GRB 170817A [224] and a kilonova (AT 2017gfo) [25, 225] have provided additional information regarding the jet profile and beaming angle [224, 226]. This joint detection has improved constraints on beaming angles [226], and when additional BNS and NSBH are discovered, constraints on this beaming angle and the rate of mergers of these binaries will improve [227].

Finally, some recent work has used the idea that mergers involving NS are the primary astrophysical source of r-process elements [228, 229] to constrain the rate of such mergers from nucleosynthesis [230, 13], and we include rates from [13] for comparison. With the discovery of GW170817 and GRB 170817A the inferred merger rate of BNS and GRB together with inferred ejected mass strongly suggest that BNS mergers are the prime sites of heavy r-process nucleosynthesis [231, 232].

The limits from O1 are not in tension with current astrophysical models. Scaling the O1 results to current expectations for advanced LIGO's next two observing runs, O2 and O3 [16], suggests that significant constraints or observations of BNS or NSBH mergers are possible during these observing runs. The observation of GW170817 and GRB 170817A in O2 further improve upon these constraints [226]. Below we follow a simple beaming angle constraint possible using the limits on BNS and NSBH merger rates from O1. This approach is agnostic of the joint discovery of GW170817, GRB 170817A, and AT 2017gfo.

Assuming that short GRB are produced by BNS or NSBH, but without using beaming angle estimates, we can constrain the beaming angle of the jet of gamma rays emitted from these GRB by comparing the rates of BNS/NSBH mergers and the rates of short GRB [233]. For simplicity, we assume here that all short GRB are associated with BNS or NSBH mergers; the true fraction will depend on the emission mechanism and jet profile. The short GRB rate R_{GRB}, the merger rate R_{merger}, and the beaming angle θ_j are then related by

$$\cos\theta_j = 1 - \frac{R_{\mathrm{GRB}}}{R_{\mathrm{merger}}} \tag{3.9}$$

We take $R_{GRB} = 10^{+20}_{-7}\mathrm{Gpc}^{-3}\ \mathrm{yr}^{-1}$ [10, 234]. Figure 12 shows the resulting GRB beaming lower limits for the 90% BNS and NSBH rate upper limits. With our assumption that all short GRBs are produced by a single progenitor class, the constraint is tighter for NSBH with larger BH mass. Observed GRB beaming angles are in the range of $3 - 25°$ [235, 8, 236, 237, 238, 239, 240]. Compared to the lower limit derived from our non-detection, these GRB beaming observations start to confine the fraction of GRBs that can be produced by higher-mass NSBH as progenitor systems. The work of [226] improves upon this given the joint observation of GW170817 and GRB 170817A, finding that the beaming angle can be constrained between $\theta_j \in (2°.88, 14°.15)$. Future constraints could also come from additional GRB and BNS or NSBH joint detections during O2, O3, and beyond [241, 242, 243, 226, 244, 245].

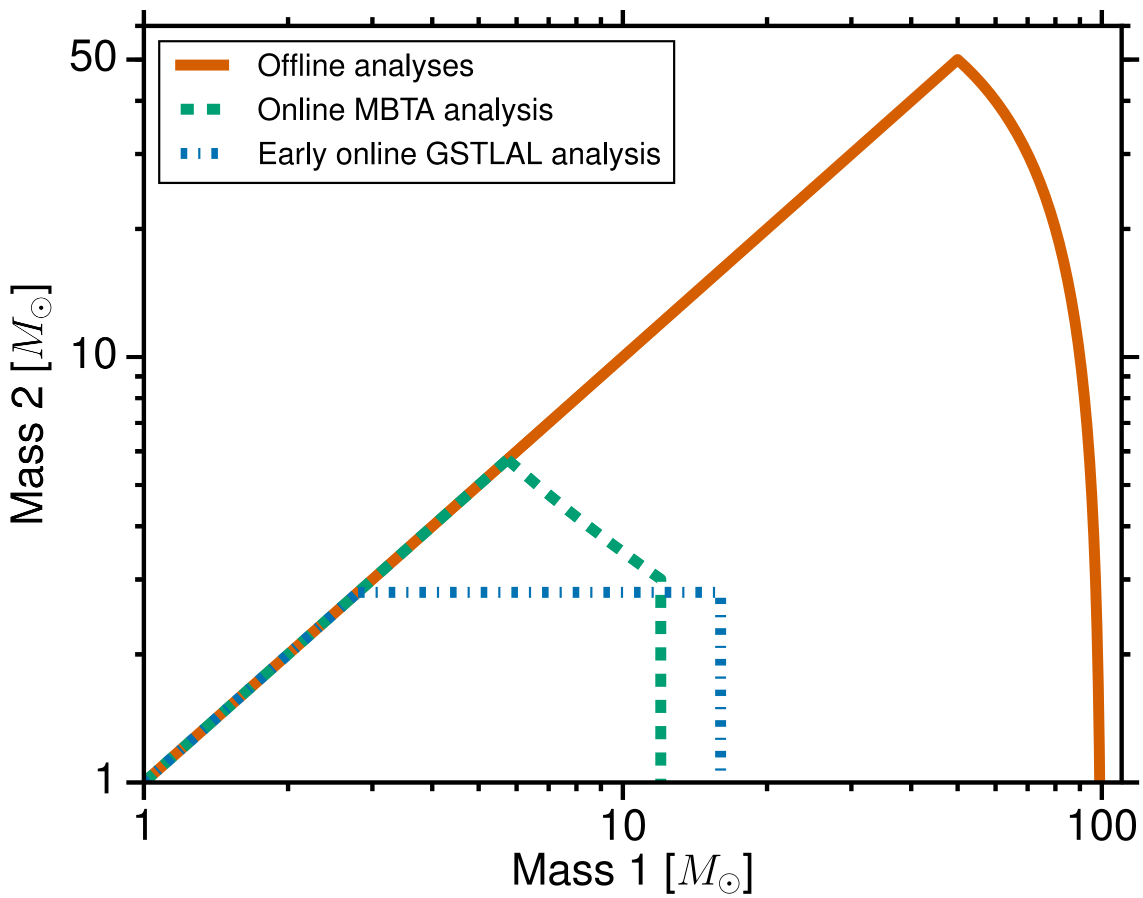


Figure 6: The range of template mass parameters considered for the three different template banks used in the search. The offline analyses, `PyCBC` and `GstLAL` used the largest bank up to total masses of $100M_\odot$. The online `GstLAL` analysis used the larger bank after December 23, 2015. The online `mbta` bank covered primary masses below $12M_\odot$ and chirp masses[3] below $5M_\odot$. The early online `GstLAL` bank up to December 23, 2015, covered primary masses up to $16M_\odot$ and secondary masses up to $2.8M_\odot$. The spin ranges are not shown here but are discussed in the text.

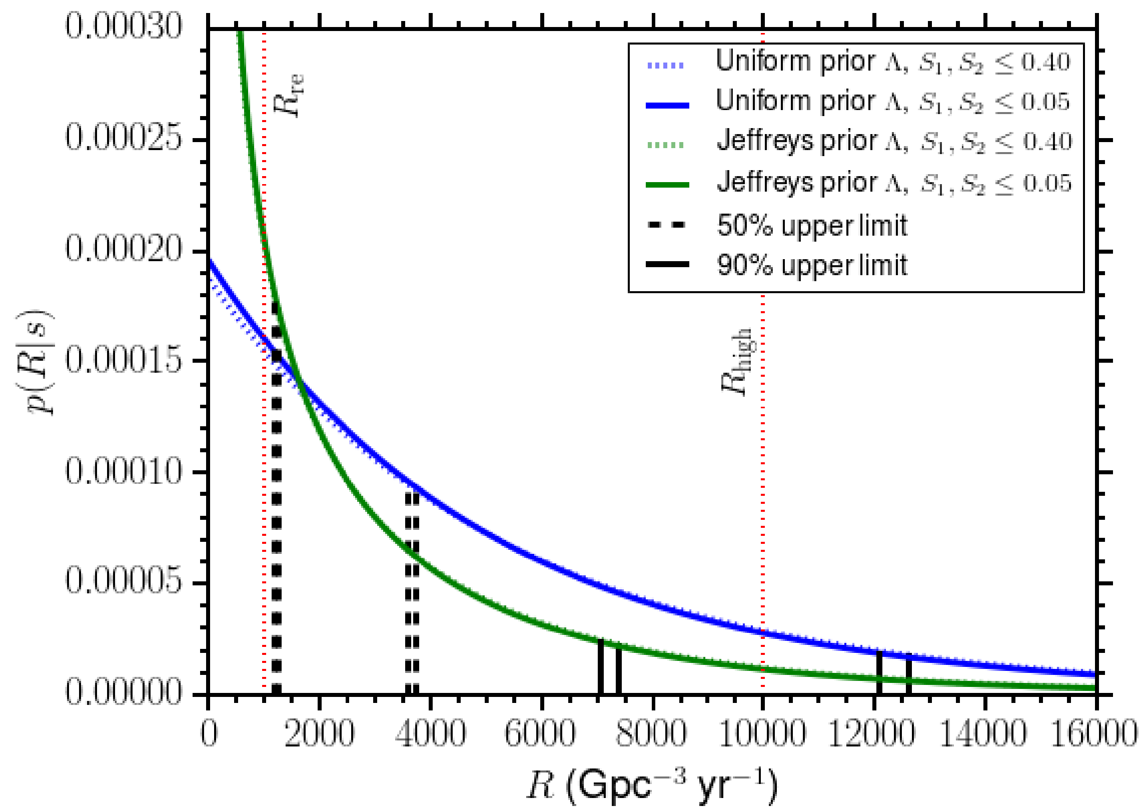


Figure 7: Posterior density on the rate of BNS mergers calculated using the `PyCBC` analysis. Blue curves represent a uniform prior on the Poisson parameter $\Lambda = R\langle VT\rangle$, while green curves represent a Jeffreys prior on Λ. The solid (low spin population) and dotted (high spin population) posteriors almost overlap. The vertical dashed and solid lines represent the 50% and 90% confidence upper limits respectively for each choice of prior on Λ. For each pair of vertical lines, the left line is the upper limit for the low spin population and the right line is the upper limit for the high spin population. Also shown are the realistic $R_{\rm re}$ and high end $R_{\rm high}$ of the expected BNS merger rates identified in [6].

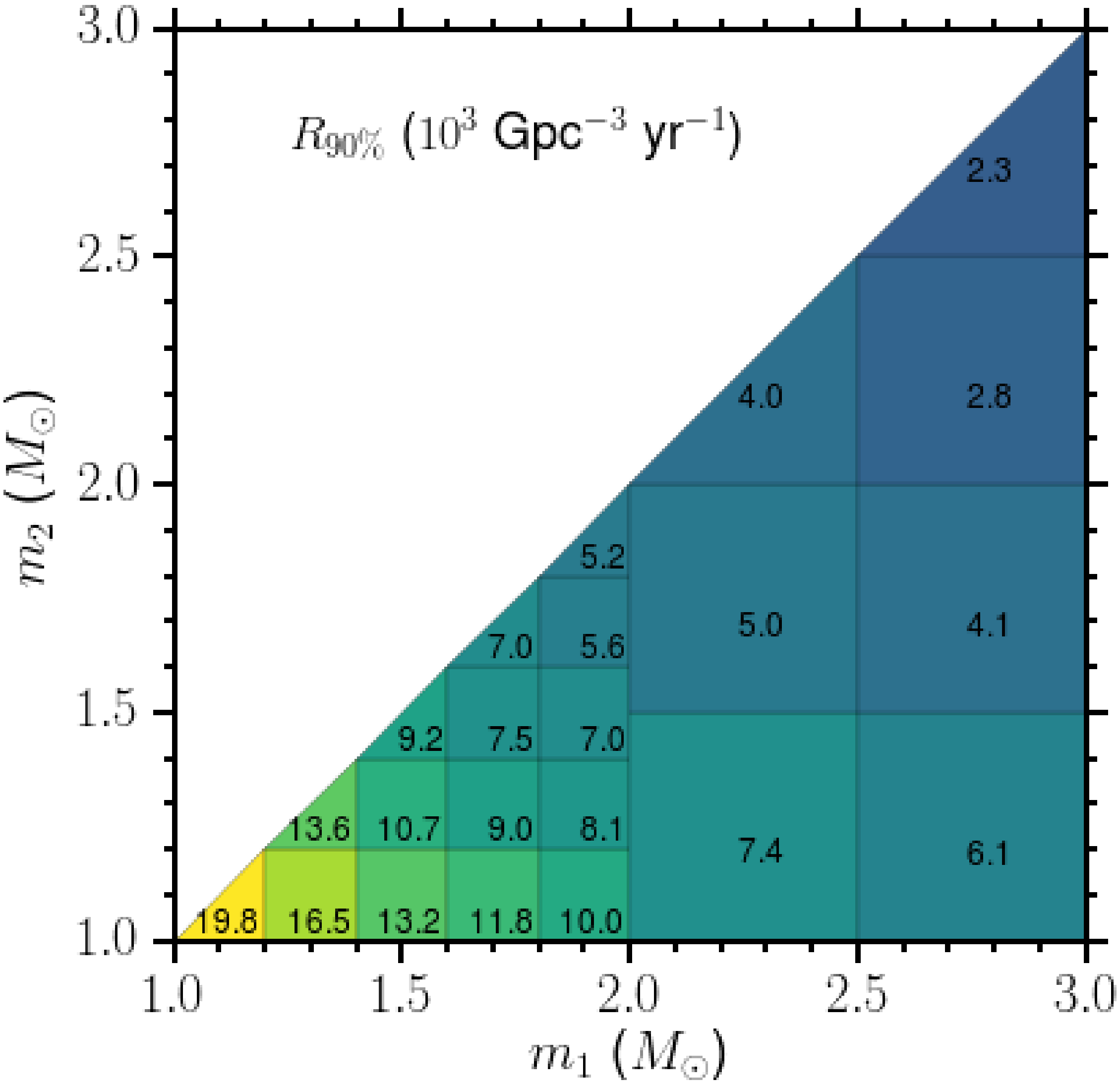


Figure 8: 90% confidence upper limit on the BNS merger rate as a function of the two component masses using the `PyCBC` analysis. Here the upper limit for each bin is obtained assuming a BNS population with masses distributed uniformly within the limits of each bin, considering isotropic spin direction and dimensionless spin magnitudes uniformly distributed in $[0, 0.05]$.

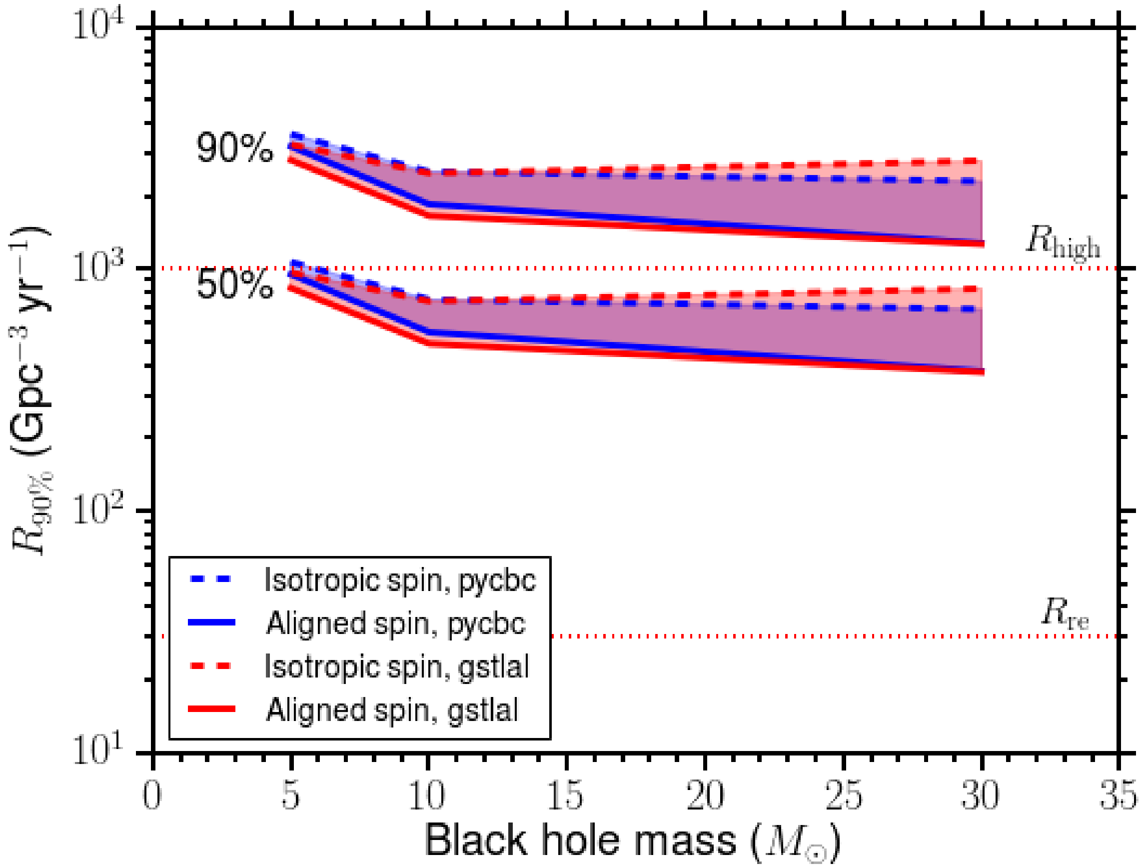


Figure 9: 50% and 90% upper limits on the NSBH merger rate as a function of the BH mass using the more conservative uniform prior for the counts Λ. Blue curves represent the `PyCBC` analysis and red curves represent the `GstLAL` analysis. The NS mass is assumed to be $1.4 M_\odot$. The spin magnitudes were sampled uniformly in the range [0, 0.04] for NS and [0, 1] for BH. For the aligned spin injection set, the spins of both the NS and BH are aligned (or anti-aligned) with the orbital angular momentum. For the isotropic spin injection set, the orientation for the spins of both the NS and BH are sampled isotropically. The isotropic spin distribution results in a larger upper limit. Also shown are the realistic $R_{\rm re}$ and high end $R_{\rm high}$ of the expected NSBH merger rates identified in [6].

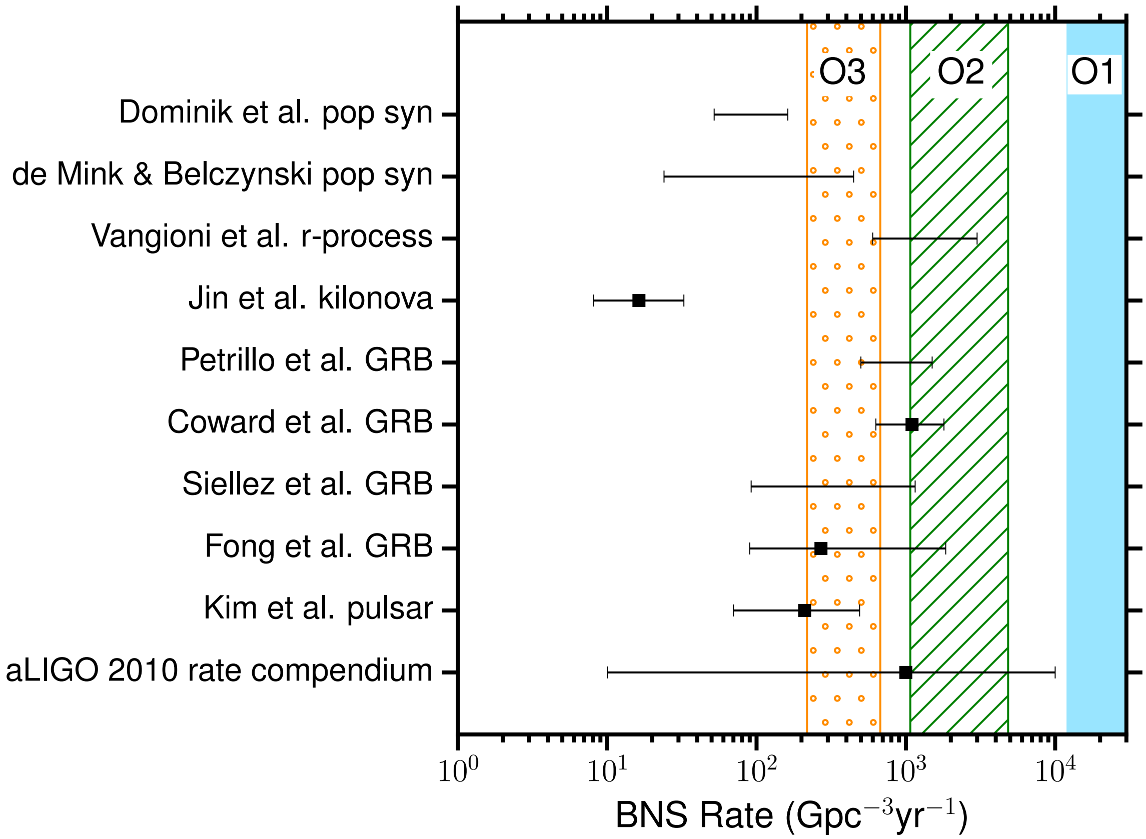


Figure 10: A comparison of the O1 90% upper limit on the BNS merger rate to other rates discussed in the text [6, 7, 8, 9, 10, 11, 12, 13, 14, 15]. The region excluded by the low-spin BNS rate limit is shaded in blue. Continued non-detection in O2 (slash) and O3 (dot) with higher sensitivities and longer operation time would imply stronger upper limits. The O2 and O3 BNS ranges are assumed to be 1-1.9 and 1.9-2.7 times larger than O1. The operation times are assumed to be 6 and 9 months [16] with a duty cycle equal to that of O1 ($\sim$ 40%).

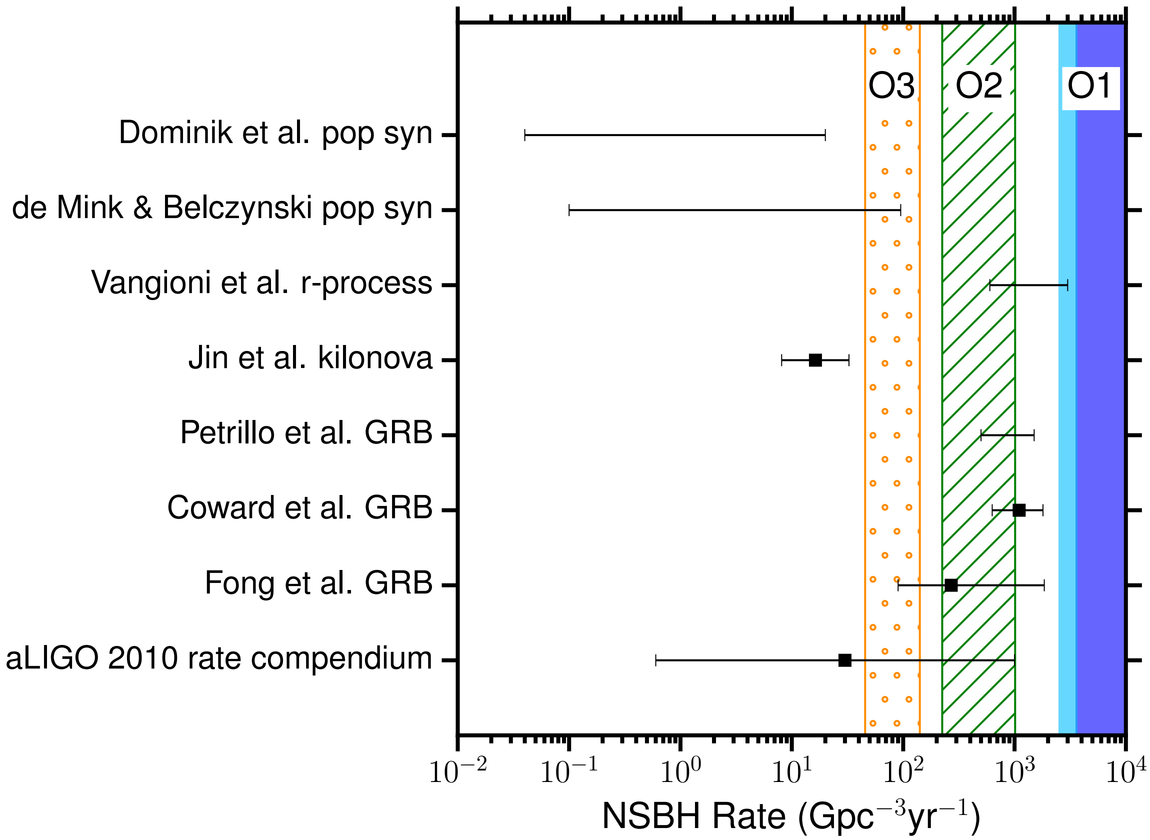


Figure 11: A comparison of the O1 90% upper limit on the NSBH merger rate to other rates discussed in the text [6, 8, 10, 11, 12, 13, 14, 15]. The dark blue region assumes a NSBH population with masses 5–1.4 $M_\odot$ and the light blue region assumes a NSBH population with masses 10–1.4 $M_\odot$. Both assume an isotropic spin distribution. Continued non-detection in O2 (slash) and O3 (dot) with higher sensitivities and longer operation time would imply stronger upper limits (shown for 10–1.4 $M_\odot$ NSBH systems). The O2 and O3 ranges are assumed to be 1-1.9 and 1.9-2.7 times larger than O1. The operation times are assumed to be 6 and 9 months [16] with a duty cycle equal to that of O1 ($\sim$ 40%).

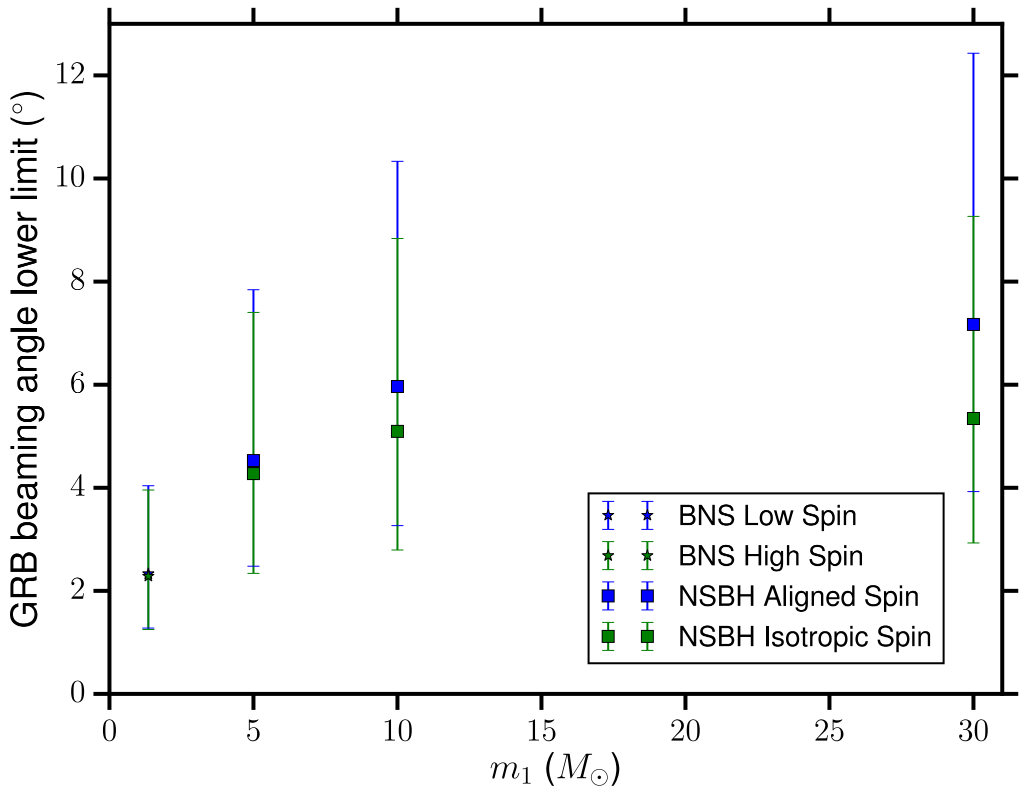


Figure 12: Lower limit on the beaming angle of short GRB, as a function of the mass of the primary BH or NS, m_1. We take the appropriate 90% rate upper limit from this paper, assume all short GRB are produced by each case in turn, and assume all have the same beaming angle θ_j. The limit is calculated using an observed short GRB rate of $10^{+20}_{-7}\mathrm{Gpc}^{-3}\ \mathrm{yr}^{-1}$ and the ranges shown on the plot reflect the uncertainty in this observed rate. For BNS, m_2 comes from a Gaussian distribution centered on $1.35M_\odot$, and for NSBH it is fixed to $1.4M_\odot$.

Injection	Range of spin	$\langle VT \rangle$ (Gpc^3 yr)	Range (Mpc)	$R_{90\%}$ (Gpc^{-3} yr^{-1})
Isotropic low spin	[0, 0.05]	2.09×10^{-4}	73.2	12,100
Isotropic high spin	[0, 0.4]	2.00×10^{-4}	72.1	12,600

Table 2: Sensitive space-time volume $\langle VT \rangle$ and 90% confidence upper limit $R_{90\%}$ for BNS systems. Component masses are sampled from a normal distribution $\mathcal{N}(1.35M_\odot, (0.13M_\odot)^2)$ with samples outside the range $[1, 3]M_\odot$ removed. Values are shown for the `PyCBC` pipeline. $\langle VT \rangle$ is calculated using a FAR threshold of 0.01 yr^{-1}. The rate upper limit is calculated using a uniform prior on $\Lambda = R\langle VT \rangle$ and an 18% uncertainty in $\langle VT \rangle$ from calibration errors.

NS mass ($M_\odot$)	BH mass ($M_\odot$)	Spin distribution	$\langle VT \rangle$ (Gpc^3 yr)	Range (Mpc)	$R_{90\%}$ (Gpc^{-3} yr^{-1})
1.4	5	Isotropic	7.01×10^{-4}	110	3,600
1.4	5	Aligned	7.87×10^{-4}	114	3,210
1.4	10	Isotropic	1.00×10^{-3}	123	2,530
1.4	10	Aligned	1.36×10^{-3}	137	1,850
1.4	30	Isotropic	1.10×10^{-3}	127	2,300
1.4	30	Aligned	1.98×10^{-3}	155	1,280

Table 3: Sensitive space-time volume $\langle VT \rangle$ and 90% confidence upper limit $R_{90\%}$ for NSBH systems with isotropic and aligned spin distributions. The NS spin magnitudes are in the range $[0, 0.04]$ and the BH spin magnitudes are in the range $[0, 1]$. Values are shown for the `PyCBC` pipeline. $\langle VT \rangle$ is calculated using a FAR threshold of 0.01 yr^{-1}. The rate upper limit is calculated using a uniform prior on $\Lambda = R\langle VT \rangle$ and an 18% uncertainty in $\langle VT \rangle$ from calibration errors.

Chapter 4

First Open Gravitational wave Catalog : 1-OGC

4.1 Introduction

Since the publication of the results by [17, 246], improvements to the data-analysis methods used by [214] have been implemented [45, 46, 47]. Using these improvements, we re-analyze the O1 data and provide a full catalog of candidate events from a matched filter search for compact binary coalescences using the O1 data. We call this full, open catalog 1-OGC. This catalog provides estimates of the statistical significance of previously known events and a ranked list of sub-threshold candidates. Although not statistically significant by themselves, these sub-threshold candidates can be correlated with archival data or transient events found by other astronomical observatories to provide constraints on the population of compact-object mergers [247, 248].

Our catalog is based entirely on public, open data and software. We use the LIGO data available from the Gravitational Wave Open Science Center [48], and analyze the data using the open source `PyCBC` toolkit [66, 65, 67]. This toolkit was also used by one of the two analyses described in [214] and in chapter 2 of this dissertation. The lowest mass sources targeted in our search are neutron star binaries with total mass $m_1 + m_2 = 2\,M_{\odot}$. The search space extends to binary black hole systems that produce gravitational waveforms longer than 0.15 s from 20 Hz. This corresponds to a total mass up to $500 M_{\odot}$ for sources with high mass ratios and spins where the component aligned with the orbital angular momentum is positive and large. For binaries with negligible spin, this corresponds to total mass $\lesssim 200 M_{\odot}$. The search space also includes neutron star–black hole binaries. After applying cuts for data quality [206, 249], a total of 48.1 days of coincident data are searched for signals.

The three most significant signals in the catalog correspond to GW150914 [135], LVT151012 [135, 17], and GW151226 [136]. No other statistically significant signals are observed. In the analysis of [17], LVT151012 was the third-most statistically significant event, but it was not detected with a statistical significance to be labeled as an unambiguous detection. With the improved methods employed here, the false alarm rate of this candidate improves by an order of magnitude and it should be considered a true astrophysical event. The analyses of [17, 246] restricted the astrophysical search space to binaries with a total mass less than $100\, M_{\odot}$. Our analysis extends this target space to higher mass signals. No additional signals are detected with statistical significance in this region of parameter space, consistent with the results of [250].

A second observing run (O2) of the Advanced LIGO detectors took place from November 30, 2016 to August 25, 2017 [16]. The Virgo gravitational wave detector also collected data for part of this period, starting from August 1, 2017. The detections reported in this second observing run thus far include nary black hole coalescence even three additional binary black hole coalescence events [251, 252, 253], and a binary neutron star merger [54]. The publication of this catalog precedes the publication of the full O2 data set and therefore the catalog presented here is restricted to the first observing run, O1.

This chapter is organized as follows: In Sec. 4.2 and Sec. 4.3, we summarize our analysis methods, including the parameter space searched, the detection statistic used for ranking candidate events, and our method for calculating the statistical significance of events. The search results are summarized in Sec. 6.5. Our full catalog is available online (`www.github.com/gwastro/1-ogc`). In this chapter, we focus on the detection of compact binary coalescences. Since no new astrophysical events have been observed, we do not consider measurement of the signals' parameters and refer to [17, 254] for discussion of the detected events' source-frame properties. Consequently, we quote binary mass parameters in the detector frame in this work.

4.2 Search Methodology

To search for gravitational waves from compact-object mergers, we use matched filtering [74] implemented in the open-source PyCBC library [66, 65, 67]. Our methods improve on the analyses of [17, 246, 214] by imposing a phase, amplitude and time delay consistency on candidate signals, an improved background model, and a larger search parameter space [45, 46, 47].

4.2.1 Target Search Space

A discrete bank of gravitational-wave template waveforms [91, 92, 97] is used to target binary neutron star, neutron star–black hole, and binary black hole mergers with total mass from $2 - 500 M_\odot$ [47]. The templates are parameterized by their component masses $m_{1,2}$ and their dimensionless spins $\vec{\chi}_{1,2} = c\vec{S}_{1,2}/Gm_{1,2}^2$, where $\vec{S}_{1,2}$ are the spin vectors of each compact object. For compact objects with component masses greater than $2M_\odot$, the template bank covers a wide range of spins, with $\chi_{(1,2)z} \in [\pm 0.998]$, where $\chi_{(1,2)z}$ are the components aligned with the orbital angular momentum. For compact objects with masses less than $2M_\odot$, the spin is restricted to $\chi_{(1,2)z} \in [\pm 0.05]$ [97]. Templates that correspond to sources with a signal duration less than 0.15 s (starting from 20 Hz) are excluded due to the difficulty in separating candidates arising from these templates from populations of instrumental glitches [47]. Consequently, the total mass boundary of the search depends strongly on the "effective spin" [255, 42],

$$\chi_{\text{eff}} = \frac{\chi_{1z} m_1 + \chi_{2z} m_2}{m_1 + m_2}. \tag{4.1}$$

This dependence is visible in the distribution of the approximately 400,000 templates required to cover the space shown in Fig. 13. A dotted line in Fig. 13 denotes the upper boundary of the O1 analysis performed in [17]. For binaries with total mass greater than $4\,M_\odot$, we use the spinning effective-one-body model (SEOBNRv4) [256, 257] as template gravitational waveforms. For sources with total masses less than $4M_\odot$ we use TaylorF2 post-Newtonian waveforms with phasing accurate to 3.5 post-Newtonian order and the dominant amplitude evolution [27, 102, 258, 259]. Our choice of template bank discretization causes less than a 10% loss in detection rate for any source within the boundaries of the template bank. Our search assumes that the source can be adequately described by only the dominant gravitational-wave mode, two component masses, non-precessing spins, and negligible eccentricity.

4.2.2 Creation and Ranking of Candidate Events

For each template and each detector, we calculate the matched filter signal-to-noise ratio (SNR) as a function of time $\rho(t)$ [74]. The template bank is divided into 15 equal sized sub-banks based on the chirp mass $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$ of each template. A single-detector "trigger" is a peak in the SNR time series that is greater than 4 and larger than any other peaks within 1s. See Fig. 5 in Chapter 2 of this dissertation for an example of this single-detector trigger with GW150914. For each sub-bank, the loudest 100 triggers (by ρ) are recorded in ~ 1 s fixed time windows. This method has been shown to improve search sensitivity, while making the rate of single-detector triggers manageable [260]. We have found this choice of sub-banks to be an effective method to ensure the analysis can concurrently record triggers from separate regions of parameter space that respond differently to instrumental noise. Other choices are also possible.

We use the data-quality segments provided by the Gravitational-Wave Open Science Center to exclude triggers that occur in times when there are problems with the detectors' data quality [206, 249]. In addition, very loud transient glitches, corresponding to $> 100\sigma$ deviations from Gaussian noise, are excised from the strain data according to the procedure of [66] before calculation of the SNR time series. However, there remain many types of transient non-Gaussian noise in the LIGO data which produce triggers with large values of SNR [261, 206, 249].

For every trigger with $\rho > 5.5$ we calculate the signal consistency test, χ_r^2 (see Eq. 2.5 in Chapter 2 of this dissertation), introduced in [77]. The statistic χ_r^2 divides the matched filter into frequency bands and checks that the contribution from each band is consistent with the expected signal. The statistic takes values close to unity when the data contains either Gaussian noise or the expected signal and larger values for many types of transient glitches. See Fig. 5 in Chapter 2 of this dissertation for an example with GW150914. We impose the SNR limit as the χ_r^2 test is generally non-informative when $\rho < 5.5$. The χ_r^2 value is used to re-weight the SNR ρ as [123, 108].

We reproduce Eq. 2.6 from Chapter 2 of this dissertation for ease of access

$$\tilde{\rho} = \begin{cases} \rho & \text{for } \chi_r^2 \leq 1 \\ \rho \left[\frac{1}{2} \left(1 + (\chi_r^2)^3 \right) \right]^{-1/6}, & \text{for } \chi_r^2 > 1. \end{cases} \tag{4.2}$$

For single-detector triggers from templates with total mass greater than $40 M_\odot$ we apply an additional test, $\chi^2_{r,\,sg}$, that determines if the detector output contains power at higher frequencies than the maximum expected frequency content of the gravitational-wave signal [46]. This test is only applied for higher mass systems, since these templates are shorter in duration and more difficult to separate from instrumental noise. For other systems, we set $\chi^2_{r,sg} = 1$. Using this statistic, we apply a further re-weighting as

$$\hat{\rho} = \begin{cases} \tilde{\rho} & \text{for } \chi^2_{r,sg} \leq 4 \\ \tilde{\rho} (\chi^2_{r,sg}/4)^{-1/2}, & \text{for } \chi^2_{r,sg} > 4. \end{cases} \tag{4.3}$$

Candidate events are generated when single-detector triggers occur in both the LIGO Hanford and Livingston data within 12 ms (the light-travel time between the observatories extended by 2 ms for signal time-measurement error) and if the triggers are recorded in the same template in each detector [66]. Following the procedure of [45], we model the distribution of single detector triggers from each template as an exponentially decaying function, $\lambda(\hat{\rho}, \theta^{\mathbf{N}})$, where $\theta^{\mathbf{N}}$ allows the parameters of the exponential to vary as a function of total mass, symmetric mass ratio $\eta = m_1 m_2 / M^2$, and χ_{eff}. This fitted model allows us to rescale $\hat{\rho}$ to better equalize the rate of triggers from each template. We produce a λ for each detector as λ_H and λ_L.

We improve upon the ranking of candidates in [246, 17] by also taking into account $p^S(\theta^{\mathbf{S}})$, which is the expected distribution of SNR ρ_H and ρ_L, phase difference $\phi_{c,H} - \phi_{c,L}$, and arrival time delay $t_{c,H} - t_{c,L}$ between the two LIGO instruments for an astrophysical population [45]. No assumption is made about the distribution of intrinsic source parameters in this term. The primary benefit arises from assuming that a population of sources is isotropically distributed in orientation and sky location. We combine the ranking statistic $p^S(\theta^{\mathbf{S}})$ with that of $\lambda(\hat{\rho}, \theta^{\mathbf{N}})_{\mathbf{H}}$ and $\lambda(\hat{\rho}, \theta^{\mathbf{N}})_{\mathbf{L}}$ to get the final ranking statistic $\tilde{\rho}_c$ as

$$\tilde{\rho}_c \propto \left[\log p^S(\theta^{\mathbf{S}}) - \log \left(\lambda_{\mathbf{H}}(\hat{\rho}_{\mathbf{H}}, \theta^{\mathbf{N}}) \lambda_{\mathbf{L}}(\hat{\rho}_{\mathbf{L}}, \theta^{\mathbf{N}}) \right) \right] + \text{const.} \tag{4.4}$$

This expression is normalized so that $\tilde{\rho}_c$ approximates the standard network SNR $\rho_c = (\rho_L^2 + \rho_H^2)^{1/2}$ for candidates from regions of parameter space that are not affected by elevated rates of instrumental noise. Candidates from regions affected by elevated rates of noise triggers are down-weighted and assigned a smaller statistic value by this method. As multiple candidates, which arise from different template waveforms, may occur in response to the same signal, we select only the highest ranked candidate within ten seconds. This is selection of the highest ranked candidate in a 10 s window is similar to the clustering of coincident triggers presented in Chapter 2 of this dissertation and the `PyCBC` analysis in the first observing run. A simpler version of this statistic where the single-detector exponential noise model is only a function of the template duration has also been employed in the analysis of data from LIGO's second observing run [262, 263, 252].

4.2.3 Statistical Significance

The statistical significance of candidate events is estimated by measuring empirically the rate of false alarms (FAR). To measure the noise background rate, we generate additional analyses by shifting the time-stamps of data from one instrument with respect to the other by multiples of 0.1s. Since this time shift is greater than the astrophysical time of flight between observatories, any candidates produced in these analyses are false alarms. This time shift is much greater than the auto-correlation length of our template waveforms of $\mathcal{O}(1\text{ms})$. The time-slid analyses are produced following the same procedure as the search; this is a key requirement for our analysis to produce valid statistical results [214]. The equivalent of more than 50,000 years of observing time can be generated from 5 days of data.

To provide an unbiased measure of the rate of false alarms at least as statistically significant as a potential candidate, the single-detector triggers that compose the candidate event should be included in the background estimation [264]. However, when a real signal with a large $\tilde{\rho}_c$ is present in the data, the rate of false alarms for candidate events with smaller $\tilde{\rho}_c$ tends to be overestimated. This is due to the fact that the loud single-detector triggers from the real event in one detector form coincidences with noise fluctuations in the other detector, producing loud coincident background events. As in [17], an unbiased rate of false alarms can be achieved by a hierarchical procedure whereby a candidate with large $\tilde{\rho}_c$ is removed from the estimation of background for candidates with smaller $\tilde{\rho}_c$; we use this procedure here.

4.3 Evaluating Candidates based on the Astrophysical Population

We find two candidate events with FAR $<$ 1 per 50 000 years, corresponding to GW150914 and GW151226. Although FAR does not give the probability that an event is an astrophysical signal, we can be confident that these events were not caused by chance coincidence between the detectors. It is possible that these events were caused by a correlated source between the detectors. However, detailed followup studies of GW150914 and GW151226 found no correlated noise sources between the detectors that could be mistaken for a gravitational wave [206, 136].

We conclude that GW150914 and GW151226 are astrophysical in origin and use them to constrain the rate of real signals. A "true discovery rate" (TDR) can be constructed for less significant events. The TDR is defined as:

$$\mathrm{TDR}(\tilde{\rho}_c) = \frac{\mathcal{T}(\tilde{\rho}_c)}{\mathcal{T}(\tilde{\rho}_c) + \mathcal{F}(\tilde{\rho}_c)}, \tag{4.5}$$

where $\mathcal{T}(\tilde{\rho}_c)$ is the rate that signals of astrophysical origin are observed with a ranking statistic $\geq \tilde{\rho}_c$ (the "true alarm rate") and $\mathcal{F}(\tilde{\rho}_c)$ is the false alarm rate.

The true discovery rate is the complement of the false discovery rate [265], and can be used to estimate the fraction of real signals in a population. For example, if $\mathrm{TDR}(\tilde{\rho}_c) = 0.9$, it means that 90% of events with a ranking statistic $\geq \tilde{\rho}_c$ are expected to be real signals. The TDR is also independent of the observation time.

Note that TDR is not the probability that a particular event is a signal of astrophysical origin P_{astro}. For that, one needs to model the distribution of signals and noise at a given $\tilde{\rho}_c$. In this work, we use a simple model of these distributions as functions of the ranking statistic $\tilde{\rho}_c$. Models incorporating additional parameters are also possible, but we do not consider them here. As a function of $\tilde{\rho}_c$, P_{astro} can be computed as

$$P_{\mathrm{astro}}(\tilde{\rho}_c) = \frac{\Lambda_S P_S(\tilde{\rho}_c)}{\Lambda_S P_S(\tilde{\rho}_c) + \Lambda_N P_N(\tilde{\rho}_c)}, \tag{4.6}$$

where $P_S(\tilde{\rho}_c)$ and $P_N(\tilde{\rho}_c)$ are the probabilities of an event having ranking statistic $\tilde{\rho}_c$ given the signal and noise hypotheses respectively [266, 267, 139]. Λ_S and Λ_N are the rates of signal and noise events.

Since no binary neutron star or neutron star–black hole candidates are obtained from a search of the O1 data, here we restrict the calculation of both the TDR and P_{astro} to binary black hole (BBH) observations. We include signals with total mass $M \geq 10\,\text{M}_{\odot}$, mass ratio $m_1/m_2 < 5$ (where $m_1 \geq m_2$), and dimensionless spins $|\chi_{(1,2)z}| < 0.5$. These choices are based on a combination of what has been observed [17, 262, 263, 252] and what is expected from models of isolated binary-star evolution ("field" binaries). The mass distribution of field binaries is dependent on a number of unknown parameters, such as the metallicity of the environment [87]. Generally, it is expected that most binaries are close to equal mass, as typically less than 1 in $\mathcal{O}(10^3)$ simulated binaries have mass ratio > 5 in models of field-binary evolution [15]. The majority of observations of nearby X-ray binaries have yielded black holes with masses greater than $5\,\text{M}_{\odot}$, which has led to speculation of a "mass gap" between 3–5 $\text{M}_{\odot}$ [166, 165, 268]. The signals detected so far by LIGO and Virgo are consistent with this: the smaller component mass in the lowest-mass system known to date, GW170608, has an estimated mass of $7^{+2}_{-2}\,\text{M}_{\odot}$ [252].

The spin distribution of black holes is not well constrained [269]. The component spins of the most significant binary black holes detected by LIGO and Virgo are only weakly constrained [17]. The best measured quantity related to spin is χ_{eff}. All of the BBH gravitational-wave signals detected so far have $|\chi_{\text{eff}}| \lesssim 0.2$. A binary with low χ_{eff} may still have component masses with large spin magnitudes, if the spins are anti-parallel or are purely in the plane of the binary. However, it seems unlikely that this would be the case for all of the detections made so far. Hence we include signals that have component spins with $|\chi_{(1,2)z}| < 0.5$. This is consistent with recent population synthesis models, which indicate that black holes must have low natal spin in order to obtain a distribution of χ_{eff} that satisfies gravitational-wave observations [270, 271].

To estimate the rate and distribution of false alarms that arise only from the region consistent with this selected population of binary black hole mergers, we must determine which templates are sensitive to these sources. It is necessary to analyze a simulated set of signals since the template associated with a particular event is not guaranteed to share the true source parameters. We find that the region of the template bank defined by $M > 8.5\,\text{M}_{\odot}$, $m_{1,2} > 2.7\,\text{M}_{\odot}$, and $\chi_{\text{eff}} < 0.9$ is effective at recovering this population of sources. This region is shown in Fig. 13 in red.

To estimate the true rate $\mathcal{T}$, we use the two significant events observed during O1, GW150914 and GW151226. We do not use any of the O2 events because at the time of this catalog the full data is not yet available for analysis, making it difficult to obtain a consistent rate estimate. The total analysis time in O1 was ~ 48 days, giving $\mathcal{T} \approx 15\mathrm{yr}^{-1}$. Given the uncertainty in this estimate based on only two events, we take the rate of observations as a Poisson process, and choose the lower 95% bound on $\mathcal{T}$. This yields a $\mathcal{T} \approx 2.7\mathrm{yr}^{-1}$. For the calculation of the TDR we use this value for all events, independent of their ranking statistic. This means we likely underestimate the TDR for events with detection statistic lower than GW151226 and GW150914, but this is a conservative bias.

To estimate the probability that a given event is astrophysical in origin P_{astro}, we model the distribution of signals and noise as a function of $\tilde{\rho}_c$. It is reasonable to approximate the signal probability distribution $P_S(\tilde{\rho}_c)$ as $\propto \tilde{\rho}_c^{-4}$ [272, 273]. We normalize the signal number density $\Lambda_S P_S(\tilde{\rho}_c)$ so that the number of signals with $\tilde{\rho}_c$ greater than or equal to some threshold $\tilde{\rho}_c^{\dagger}$ is $\approx 2.7\mathrm{yr}^{-1}$. We make the conservative choice to place $\tilde{\rho}_c^{\dagger}$ at the value of the next largest $\tilde{\rho}_c$ value after GW150914 and GW151226.

To approximate the noise number density $\Lambda_N P_N(\tilde{\rho}_c)$, we make a histogram of the $\tilde{\rho}_c$ values of false alarms arising from our selected BBH region. We use only the false alarms which are uncorrelated with possible candidate events to ensure an unbiased estimate of the mean false alarm rate [264]. We fit an exponential decay to this histogram from $8 < \tilde{\rho}_c < 9.2$. For $\tilde{\rho}_c$ much less than 8, $\Lambda_N P_N$ is not well modeled by an exponential due to the effects of applying a threshold to single-detector triggers. We note, however, there is only a 50% chance that an event is astrophysical at $\tilde{\rho}_c \sim 8.6$, and this chance quickly becomes negligible with decreasing $\tilde{\rho}_c$. The result of this procedure is shown in Fig. 14. We caution that P_{astro} for candidates with $\tilde{\rho}_c > 9.2$ will be sensitive to the form of the model chosen since it is not constrained by empirically measured false alarms.

While we do not assess the astrophysical probabilities of sources outside our selected BBH region, we are not precluding that such sources exist. Our P_{astro} is compatible with any model of the true BBH source distribution that allows for a signal rate to be at least as high as our estimate within the chosen region. This holds irrespective of whatever other kinds of sources may also be permitted.

4.4 Results and Binary Black Hole Candidates

The results presented here are generated using the data from the first observing run of Advanced LIGO which ran from September 12, 2015 to January 19, 2016. We divide the 16 kHz LIGO open data into 9 consecutive periods of time and search each time period independently so that each analysis contains roughly five days of observing time. This time interval is set by the disk and memory requirements of the search pipeline, but it is sufficient to estimate the FAR of candidate events to better than 1 in 50,000 years. It is possible to combine these time intervals during the analysis to improve this limit, but we have not done so here. Our analysis is restricted to times marked as observable by the metadata provided by the Gravitational-Wave Open Science Center. After accounting for times which are marked as not analyzable, there remain ~ 48.1 days of data when both the Hanford and Livingston LIGO instruments were operating.

The top candidate events by FAR from the full search are given in Table 4. There are three candidates which are statistically significant. These are the binary black hole mergers GW150914, LVT151012, and GW151226, which were previously reported in [17, 135, 136]. The false alarm rates for GW150914 and GW151226 of 1 per 66,000 and 1 per 59,000 years, respectively, are limits based on the amount of background time available in their respective analysis. These limits are less stringent than those reported in [17] as we have created less background time. There are no other individually convincing candidates. Fig. 15 shows candidate events with $\tilde{\rho}_c > 7.5$. The three binary black hole mergers stand out from the other candidate events and are clustered in a portion of the parameter space that is analyzed with relatively few template waveforms.

Given that there are two binary black hole mergers (GW150914 and GW151226) that are well established from their statistical significance, we can estimate the rate of detecting binary black hole mergers by this analysis. Candidate events that are consistent with our selected binary black hole population are listed in Table 5. We estimate the false alarm rate of events for just this region of the analysis, and using our estimate of the true rate of detections, calculate the true discovery rate as a function of ranking statistic. The TDR at the ranking statistic of the fourth most significant candidate is 0.52. This means that only 52% of candidates with $\tilde{\rho}_c$ at least as large are expected to be of astrophysical origin. For each candidate we estimate its individual probability of being astrophysical in origin, P_{astro}. The fourth event has only a 6% chance of being astrophysical. We do not report P_{astro} and TDR values for the top two events since these events are assumed to be signals in the construction of these statistics.

4.5 Revisiting LVT151012

LVT151012 was first announced in [214], with a FAR of 1 per 2.3 years. Our improved methods yield a false alarm rate for LVT151012 of 1 per 24 years. Restricting attention to our selected BBH region, which is consistent with the other observed binary black hole mergers, gives a FAR for LVT151012 in this region alone of 1 per 446 years. We combine this FAR with our conservative estimate of the rate of detections to estimate that 99.92% of binary black hole merger candidates at least as significant as LVT151012 are astrophysical in origin. We also estimate the probability that specifically LVT151012 is astrophysical in origin to be 97.59%.

These measures both depend on our selected region of binary black hole sources and our estimate of the rate of true detections, but we believe our choices for both of these to be conservative. The FAR of 1 per 446 years is not a statistical statement about the search as a whole and is used only in comparison against the rate of real signals within this same region. Selecting different boundaries for this region would yield a different FAR. However, assuming that the false alarm rate and true alarm rate are both approximately uniform in this region, then P_{astro} and TDR will not change.

As data from future observing runs becomes available, it will be possible to more precisely estimate this rate in a consistent way, and improve our estimate of this event's significance. We have modeled our signal distribution and population of false alarms as being characterized by the ranking statistic $\tilde{\rho}_c$ alone. An improved model could take into account the variation over the parameter space and in time. Fig. 14 shows the probability distribution of our noise and signal models for the analysis which contains LVT151012. Compared to the P_{astro} reported in [17] of 87%, our analysis has improved the ranking of candidate events, the boundaries of our selected BBH distribution differ from what was used there, and we use a more conservative estimate of the signal rate. Given a P_{astro} value of 97.6% we conclude that LVT151012 is astrophysical in origin. For comparison, if we had chosen the rate of observed mergers to be $\approx 15\mathrm{yr}^{-1}$, which is the linear extrapolation of two detections in 48 days, we would find that LVT151012 had a 99.6% probability of astrophysical origin.

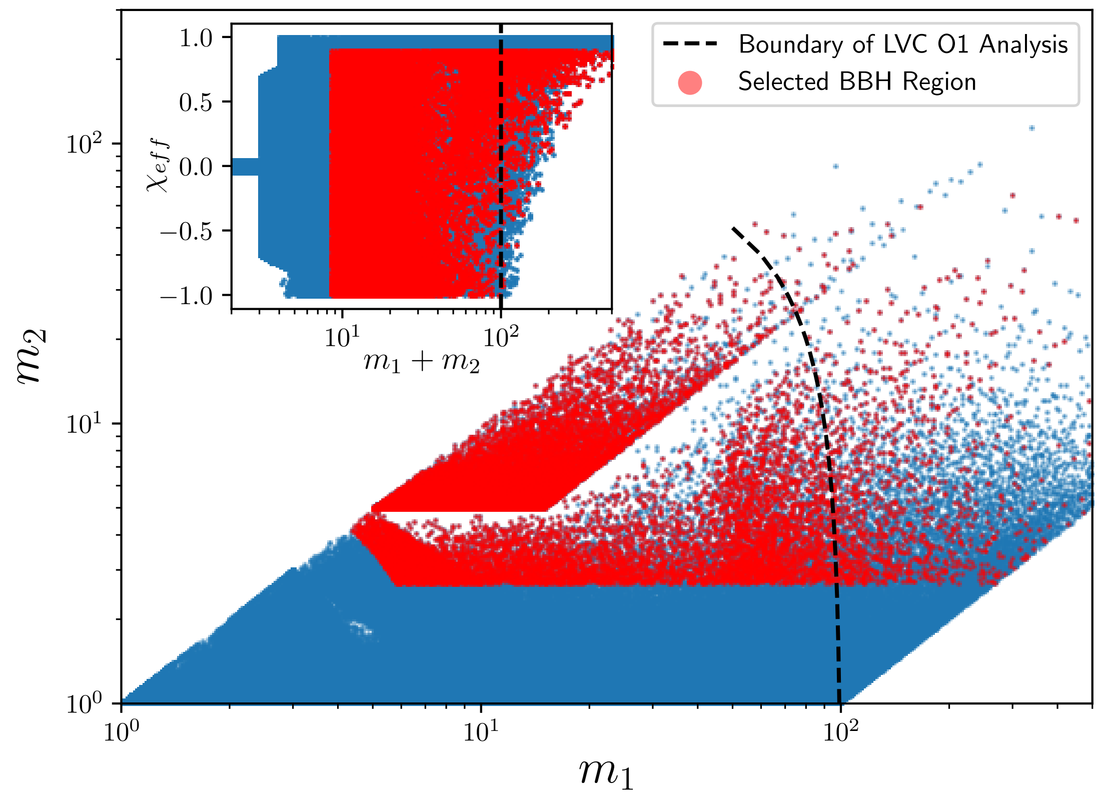


Figure 13: The component masses and spins of the templates used to search for compact binary mergers. Due to the exclusion of short duration templates, there is a dependency on the total mass searched and its effective spin. For binary black holes with negligible spin, this implies that this study only probes sources with total mass less than $200\,\mathrm{M}_\odot$. Visible artifacts due to the procedure for constructing the template bank do not impact performance. Templates which we conservatively consider to produce binary black hole (BBH) candidates consistent with known observations are shown in red as discussed in Sec. 4.3. The upper mass boundary of the analysis performed by the LVC in [17] is shown as a black dotted line.

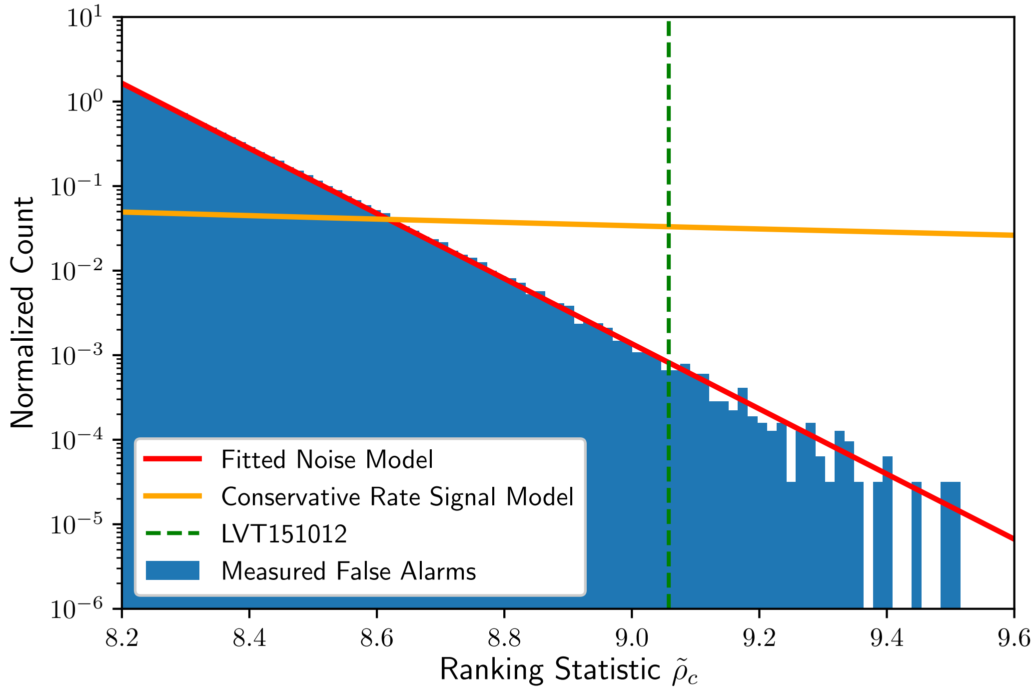


Figure 14: The scaled probability distributions of assumed signals and noise as a function of the ranking statistic $\tilde{\rho}_c$ for the analysis containing LVT151012. Blue shows the normalized histogram of empirically measured false alarms that are within our selected BBH region of the template bank, P_N. Red is the exponential decay model that has been fitted to this set of false alarms, $P_S \Lambda_S / \Lambda_N$, normalized so that the counts can be directly compared to the noise distribution. Orange shows the signal model based on our conservative rate of detections. The value of $\tilde{\rho}_c$ for LVT151012 is shown as a dotted green vertical line. The ratio of signal to noise at this value of $\tilde{\rho}_c$ strongly favors the signal model.

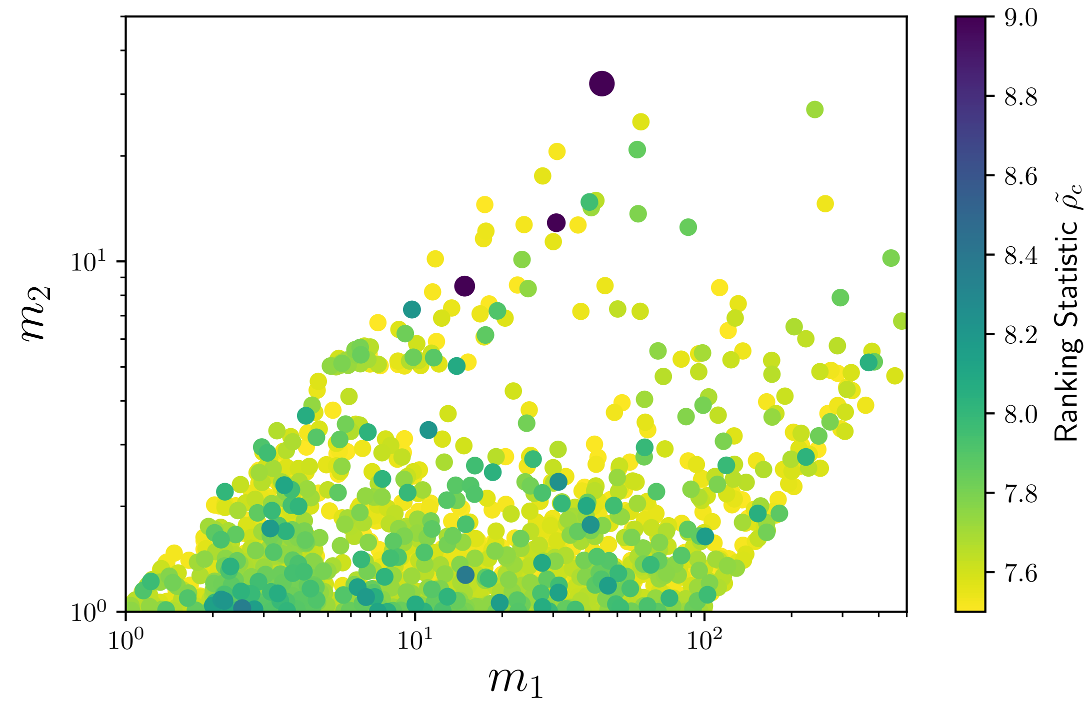


Figure 15: Candidate events with a ranking statistic $\tilde{\rho}_c > 7.5$ from the full search for compact binary mergers in O1 data. The colorbar is capped at 9. The three BBH mergers are clearly visible in the plots, while the remaining events are largely distributed according to the density of the template bank.

Table 4: Candidate events from the full search for compact binary mergers in O1 data. Candidates are sorted by FAR evaluated for the entire bank of templates. The FAR of the top two candidates is limited only by the amount of background time estimated, and only differ due to the variation in time available in their respective analyses to create background. The parameters of the template associated with each candidate are listed. Note that these are not intended as a rigorous estimation of the source parameters. Masses are given in the detector frame.

Designation	Julian Date	$FAR^{-1}(yr)$	$\tilde{\rho}_c$	ρ_H	ρ_L	m_1	m_2	$\chi_{\rm eff}$
150914+09:50:45UTC	2457279.910665	>66000	18.45	19.67	13.38	44.21	32.16	0.09
151226+03:38:53UTC	2457382.652426	>59000	11.62	10.73	7.43	14.83	8.50	0.24
151012+09:54:43UTC	2457307.913420	24	9.06	6.96	6.71	30.75	12.89	-0.05
151019+00:23:16UTC	2457314.516585	0.060	8.39	6.81	5.47	14.93	1.27	0.11
150928+10:49:00UTC	2457293.951122	0.042	8.37	6.05	6.34	2.53	1.02	-0.70
151218+18:30:58UTC	2457375.271929	0.029	8.24	7.11	5.38	31.29	2.35	-0.00
160103+05:48:36UTC	2457390.742504	0.026	8.22	6.01	6.60	9.75	7.29	0.49
151202+01:18:13UTC	2457358.554740	0.025	8.23	6.54	5.73	40.42	1.77	-0.26
160104+03:51:51UTC	2457391.661424	0.021	8.19	5.80	6.39	6.76	1.10	-0.51
151213+00:12:20UTC	2457369.508985	0.019	8.22	5.70	7.24	11.12	3.30	-0.79
150923+07:10:59UTC	2457288.799711	0.014	8.20	6.78	5.84	2.14	1.08	0.65
151029+13:34:39UTC	2457325.066149	0.014	8.21	6.83	5.23	2.19	1.07	-0.27
151206+14:19:29UTC	2457363.097291	0.013	8.17	5.80	6.37	100.60	1.64	0.98
151202+15:32:09UTC	2457359.147751	0.012	8.14	5.93	6.41	6.33	1.18	-0.59
151012+06:30:45UTC	2457307.771774	0.011	8.19	6.74	5.70	3.16	1.73	-0.15
151116+22:41:48UTC	2457343.446120	0.010	8.14	5.79	6.64	2.00	1.04	-0.45
151121+03:34:09UTC	2457347.649138	0.010	8.12	6.48	5.78	7.43	1.00	-0.86
150922+05:41:08UTC	2457287.737317	0.010	8.16	6.05	6.34	2.78	1.02	0.17
151008+14:09:17UTC	2457304.090202	0.008	8.16	5.84	6.10	46.38	1.19	0.38
151127+02:00:30UTC	2457353.584101	0.008	8.10	6.28	5.44	39.12	2.01	0.99

Table 5: Candidate events consistent with the selected population of binary black holes. There are three binary black hole mergers above a threshold corresponding to a true discovery rate of 99.92%. The third most significant event, LVT151012, has a 97.6% probability of being astrophysical in origin. Note that the FARs indicated do not reflect the false alarm rate for the full search, but instead for the limited region of the template bank indicated in red in Fig. 13. The FARs listed for the top two events are limited by the background time generated and so are identical to those in Table 4.

Designation	Julian Date	P_{astro}	TDR	$FAR^{-1}(yr)$	$\tilde{\rho}_c$	ρ_H	ρ_L	m_1	m_2	χ_{eff}
150914+09:50:45UTC	2457279.910665	-	-	>66000	18.45	19.67	13.38	44.21	32.16	0.09
151226+03:38:53UTC	2457382.652426	-	-	>59000	11.62	10.73	7.43	14.83	8.50	0.24
151012+09:54:43UTC	2457307.913420	0.976	0.999	446	9.06	6.96	6.71	30.75	12.89	-0.05
160103+05:48:36UTC	2457390.742504	0.061	0.517	0.396	8.22	6.01	6.60	9.75	7.29	0.49
151213+00:12:20UTC	2457369.508985	0.047	0.455	0.309	8.22	5.70	7.24	11.12	3.30	-0.79
151216+18:49:30UTC	2457373.284799	0.017	0.223	0.106	8.09	6.10	6.01	13.92	5.03	-0.41
151222+05:28:26UTC	2457378.728506	0.012	0.169	0.075	8.03	5.67	6.46	6.86	3.26	-0.74
151217+03:47:49UTC	2457373.658627	0.006	0.088	0.036	7.96	6.69	5.57	40.02	14.77	0.84
151009+05:06:12UTC	2457304.713060	0.005	0.087	0.035	7.99	5.66	5.90	25.55	2.73	-0.05
151220+07:45:36UTC	2457376.823761	0.003	0.053	0.021	7.87	6.55	5.39	17.50	6.17	0.82
151104+04:12:55UTC	2457330.676062	0.003	0.053	0.021	7.91	5.94	6.33	19.25	7.22	0.71
151120+16:20:06UTC	2457347.181049	0.003	0.047	0.018	7.86	6.11	5.44	5.49	3.10	0.79
151216+09:24:16UTC	2457372.892271	0.003	0.045	0.017	7.86	5.76	5.66	58.56	20.84	0.66
151128+14:37:02UTC	2457355.109478	0.003	0.040	0.016	7.83	6.79	5.02	9.25	6.22	-0.87
160109+08:08:42UTC	2457396.839798	0.003	0.035	0.014	7.82	5.24	6.23	24.29	3.45	-0.98
160111+22:49:34UTC	2457399.451507	0.003	0.035	0.013	7.82	5.10	6.55	5.75	3.43	0.23
151124+11:25:19UTC	2457350.976339	0.002	0.033	0.013	7.81	5.65	6.27	98.89	3.89	0.45
150912+15:39:02UTC	2457278.152523	0.002	0.032	0.012	7.84	6.23	5.23	9.86	5.33	-0.01
151006+06:06:50UTC	2457301.755168	0.002	0.031	0.012	7.89	6.77	5.47	11.59	5.31	-0.05
151015+01:40:52UTC	2457310.570466	0.002	0.029	0.011	7.85	5.37	5.92	87.87	12.52	0.75

Chapter 5

Bayesian Hypothesis Testing

5.1 Probability

Here we consider a few rules of probability theory before we begin discussing Bayesian inference and hypothesis testing. We follow the guide of [274] for the rules of probability. Firstly, for a given set of N possible outcomes where each outcome has a probability p_i of occurring then the sum of all possible outcomes must be unity. This can be expressed as

$$\sum_{i=1}^{N} p_i = 1. \tag{5.1}$$

This is also true for probability distribution functions $p(x)$ described by a continuous parameter x. We express this rule of probability as

$$\int p(x)\, dx = 1. \tag{5.2}$$

Next, we consider the probability of independent events occurring and the concept of conditional probability. Two events A and B are said to be independent of one another if the probability of A is unaffected by what we may know about B. This can be stated as:

$$p(A \,\text{and}\, B) \equiv p(A,\, B) = p(A)\, p(B). \tag{5.3}$$

In cases that independence does not hold we can consider the conditional probability of A given the information that we know of B. This conditional probability is stated as

$$p(A \,|\, B) = \frac{p(A,\, B)}{p(B)}. \tag{5.4}$$

Here $p(A \,|\, B)$ is the probability of A given that B has occurred. If A and B are independent events this reduces back to Eq. (5.3). If there are many possibilities for event B, which we label as B_i, then we can also attain the probability of A through the following summation

$$p(A) = \sum_i p(A \,|\, B_i)\, p(B_i). \tag{5.5}$$

This is called marginalization and pertains to summing out nuisance parameters. This technique of marginalization also generalizes to continuous probability distributions and can be expressed as

$$p(A) = \int p(A \,|\, B)\, p(B)\, dB. \tag{5.6}$$

In order to conduct Bayesian inference and hypothesis testing we will make extensive use of conditional probabilities and probability marginalization.

5.2 Bayesian Inference

In Bayesian statistical inference we are interested in using our data to update our beliefs regarding hypotheses and the parameters that come with these hypotheses. We can make use of Bayes' theorem to take prior beliefs about hypotheses, take observations of the data, and use these observations to construct posterior beliefs about hypotheses. We begin this approach by introducing Bayes' theorem

$$p(\mathrm{H})\, p(\mathbf{d} \,|\, \mathrm{H}) = p(\mathrm{H} \,|\, \mathbf{d})\, p(\mathbf{d}). \tag{5.7}$$

Here we have $p(\mathrm{H})$ which represents our prior belief about the hypothesis H. The likelihood $p(\mathbf{d} \,|\, \mathrm{H})$ represents the probability of observing our data under the assumptions implicity stated in our hypothesis H. We use these to update our inference on the probability of hypothesis H as expressed in the posterior probability $p(\mathrm{H} \,|\, \mathbf{d})$. Finally, we have $p(\mathbf{d})$, the probability of obtaining this, data set. For the moment we will consider this as a normalization factor. For ease of reading we will change notation so that the prior $p(H)$ is $\pi(H)$, the likelihood $p(\mathbf{d} \,|\, \mathrm{H})$ is $\mathcal{L}(\mathbf{d} \,|\, \mathrm{H})$, the posterior distribution $p(\mathrm{H} \,|\, \mathbf{d})$ is $\mathcal{P}(\mathrm{H} \,|\, \mathbf{d})$, and the normalization factor $p(\mathbf{d})$ is $\mathcal{Z}(\mathbf{d})$ following the notation of [275].

Within the context of parameter estimation we often use a hypothesis H that is composed from prior distributions on parameter values $\vec{\theta}$ that fully describe the hypothesis. In this context it is more helpful to rewrite Eq. (5.7) as

$$\mathcal{P}(\vec{\theta}\,|\,\mathbf{d},\mathrm{H}) = \frac{\pi(\vec{\theta}\,|\,\mathrm{H})\,\mathcal{L}(\mathbf{d}\,|\vec{\theta})}{\mathcal{Z}(\mathbf{d}\,|\,\mathrm{H})}. \tag{5.8}$$

In Eq. (5.8) the normalization constant $\mathcal{Z}$ is often called the marginal likelihood since it marginalizes all parameters of the model H out of the likelihood [275]:

$$\mathcal{Z}(\mathbf{d}\,|\,\mathrm{H}) = \int \pi(\vec{\theta}\,|\,\mathrm{H})\,\mathcal{L}(\mathbf{d}\,|\,\vec{\theta})d\vec{\theta}. \tag{5.9}$$

This marginal likelihood is often called the evidence since it summarizes the likelihood of obtaining the data given the hypothesis over the entire prior distribution. That is, we have marginalized the likelihood across the entire prior distribution. We can compare the evidences between different hypotheses as a way to see which hypothesis better *predicts* the data. The larger the evidence, the better the hypothesis predicts the data. In most practical cases this integral is analytically intractable due to the dimensionality of the prior distribution. We could consider estimation of this evidence via Monte Carlo sampling techniques. To clarify, we can consider Eq. (5.9) as the prior-weighted average likelihood. We can consider a Monte Carlo simulation where we draw random samples from the prior distribution, measure the likelihood, and then take the arithmetic mean of this likelihood. This arithmetic mean will approximate the marginal likelihood up to some Monte Carlo uncertainty. In practice, this is impractical in realistic astrophysical applications [274]. We will consider methods that are more effective for estimating the Bayesian evidence in Sections 5.4 and 5.5.

We can compare two competing hypotheses by calculating the likelihood ratio of the marginal likelihoods for each hypothesis. If the likelihood ratio is greater than unity, then the hypothesis in the numerator predicts the data with higher likelihood than the hypothesis in the denominator. If this likelihood ratio is smaller than unity, then the hypothesis in the numerator predicts the data with lower likelihood than the hypothesis in the denominator. This likelihood ratio test is known as the Bayes factor. For two hypotheses, H_A, and H_B, with marginal likelihoods $\mathcal{Z}(\mathbf{d}\,|\,\mathrm{H}_A)$ and $\mathcal{Z}(\mathbf{d}\,|\,\mathrm{H}_B)$, respectively, the Bayes factor is defined as

$$\mathcal{B}^{\mathrm{H}_A}_{\mathrm{H}_B} \equiv \frac{\mathcal{Z}(\mathbf{d}\,|\,\mathrm{H}_A)}{\mathcal{Z}(\mathbf{d}\,|\,\mathrm{H}_B)}. \tag{5.10}$$

Therefore if the Bayes factor is unity for these two hypotheses, then each hypothesis is equally likely to have predicted the data. We can convert a Bayes factor into a posterior odds ratio to give the odds of one hypothesis over another via:

$$\mathcal{O}^{\mathrm{H}_A}_{\mathrm{H}_B} = \mathcal{B}^{\mathrm{H}_A}_{\mathrm{H}_B} \times \frac{\pi(\mathrm{H}_A)}{\pi(\mathrm{H}_B)}. \tag{5.11}$$

Here $\mathcal{O}^{\mathrm{H}_A}_{\mathrm{H}_B}$ represents the odds that hypothesis H_A is preferred over hypothesis H_B after the observation of the data. This is called the posterior odds ratio. The ratio $\pi(\mathrm{H}_A)/\pi(\mathrm{H}_B)$ represents our prior odds ratio, that is, how much more did we believe that hypothesis H_A was preferred over hypothesis H_B before observing the data. The prior odds ratio gives us a statement of what level of Bayes factor we would require before we begin to change our minds about the odds of hypothesis H_B being better supported in the data than hypothesis H_A. It is considered good practice to state prior probabilities and prior odds at the outset of an experiment [275]. Choice of prior probabilities on hypotheses are subjective but should not be considered arbitrary since they represent explicit decisions in experimental design which requires scientific expertise. An uninformative prior on each hypothesis, indicating no prior preference for either hypothesis, would set the prior odds ratio to unity.

If there are only two hypotheses being considered then an odds ratio can be converted into a probability of one hypothesis over another hypothesis through the following expression [276]:

$$\mathcal{P}\left(\mathrm{H}_A \,|\, \mathbf{d}\right) = \frac{\mathcal{O}^{\mathrm{H}_A}_{\mathrm{H}_B}}{1 + \mathcal{O}^{\mathrm{H}_A}_{\mathrm{H}_B}}. \tag{5.12}$$

Here the probability of hypothesis H_A after observation of the data is given as $\mathcal{P}(\mathrm{H}_A \,|\, \mathbf{d})$. This is the posterior probability of the hypothesis H_A. Since there are only two possible hypotheses the posterior probability of H_B is $\mathcal{P}(\mathrm{H}_B)) \equiv 1 - \mathcal{P}(\mathrm{H}_A))$. In Fig. 16 we compare the logarithm of the odds ratio with the posterior probability for a hypothesis. When the log odds ratio is 0 the probability of one hypothesis relative to the other hypothesis is 0.5. Furthermore, we can also make a mapping of this probability to a single-tailed z-score of a Gaussian distribution. This is the familiar test statistic σ used in physics (see Chapter 2 for an example of its use for the estimation of the statistical significance of GW150914). Specifically, the conversion from probability to this statistic is given by z-score $= -\sqrt{2}\mathrm{erfc}^{-1}(2p)$, where erfc^{-1} is the inverse complementary error function and p is the probability value[1]. A z-score of 0 (0σ) indicates a 50% probability, while a z-score of 5 (5σ) is $\sim 10^{-5}$% probability. The relationship between this z-score and the log odds ratio is shown in Fig 17.

In cases where there are more than two hypotheses available, we set one model as the fiducial model such that all Bayes factors are calculated relative to this fiducial hypothesis ($\mathrm{H}_{\text{fiducial}}$) [276]. Then posterior probabilities for individual hypotheses can be calculated as

$$\mathcal{P}\left(\mathrm{H}_i \,|\, \mathbf{d}\right) = \frac{\mathcal{O}^{\mathrm{H}_i}_{\mathrm{H}_{\text{fiducial}}}}{\sum_{j=1}^{\mathrm{N}} \mathcal{O}^{\mathrm{H}_j}_{\mathrm{H}_{\text{fiducial}}}}. \tag{5.13}$$

Here the summation is over all N available hypotheses. This posterior probability for an individual hypotheses is often useful to calculate to compare how informative any individual hypothesis is relative to all available models. In practice can be computationally difficult to test a large set of hypotheses, but [277] is an example of testing a large set of hypotheses in the field of gravitational wave astronomy.

As [276] notes we are often not just concerned with the Bayes factors and posterior probabilities on hypotheses but we also want to learn from the inference on parameters conditional on each of our tested hypotheses. We can inspect each of these posterior probabilities on parameters for each hypotheses in isolation or we can take a Bayesian model averaging approach of [278]. Bayesian model averaging involves coherently combining the parameter inference for common parameters from many different hypotheses. To do so we calculate the posterior probability for each hypothesis based on the posterior odds ratio (or Bayes factor if the prior odds ratios for hypotheses are unity) as in Eq. (5.13). Next, we reweight the marginal posterior probabilities for common parameters for each hypothesis by the posterior probability of the hypothesis. Finally, we sum these reweighted posterior probabilities on common parameters to create a multi-model parameter inference on these common parameters. This technique is often used in Bayesian cosmological modeling [275]. The mathematical expression for Bayesian model averaging for parameter inference on a parameter Δ can be stated as

$$\mathcal{P}\left(\Delta \,|\, \mathbf{d}\right) = \sum_{i=1}^{\mathrm{N}} \mathcal{P}\left(\mathrm{H}_i \,|\, \mathbf{d}\right) \, \mathcal{P}\left(\Delta \,|\, \mathbf{d}, \mathrm{H}_i\right). \tag{5.14}$$

[1]In Chapter 2 we converted p-values to single-sided Gaussian standard deviation scores via the inverse error function. The formulation used here is equivalent.

Here the summation is over all N available hypotheses. The term $\mathcal{P}\,(\Delta\,|\,\mathbf{d}, \mathrm{H}_i)$ represents the parameter inference on Δ conditional on hypothesis H_i. If we have a continuous "hyper-parameter" α that connects all the hypotheses we are testing then the summation in Eq. (5.14) becomes an integration over α.

In addition to testing hypotheses on one particular observation $\mathbf{d}$ it is possible to combine the inference from multiple observations, $\{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_\mathrm{N}\}$. If we are testing the same prior distribution and each observation is statistically independent, then the Bayes factor for multiple observations can be combined through multiplication [276, 279]. We temporarily suppress notation on hypotheses, and adopt the notation of [280] for Bayes factors from multiple observations to give the combined Bayes factor as

$$\mathcal{B}(\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_\mathrm{N}) = \mathcal{B}(\mathbf{d}_1) \times \mathcal{B}(\mathbf{d}_2) \times \ldots \times \mathcal{B}(\mathbf{d}_N). \tag{5.15}$$

Here $\mathcal{B}(\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_\mathrm{N})$ represents the Bayes factor after many observations. Even if the Bayes factor for a particular hypothesis is not statistically significant for any individual observation $\mathbf{d}_i$, we can accumulate evidence over many observations to reach a clearer conclusion about how well supported a hypothesis is by the many observations.

We can also update our posterior distributions on parameters over many observations. Here we follow [281] and consider the updated marginal posterior probability $\mathcal{P}\;(\Delta\,|\,\mathrm{H}, \mathbf{d}_1, \ldots, \mathbf{d}_N,)$ on a parameter Δ after N observations

$$\mathcal{P}\;(\Delta\,|\,\mathrm{H}, \mathbf{d}_1, \ldots, \mathbf{d}_N,) = \frac{\pi(\Delta\,|\,\mathrm{H})}{c}\mathcal{L}(\mathbf{d}_1, \ldots, \mathbf{d}_N\,|\,\Delta, \mathrm{H}). \tag{5.16}$$

Here we have a prior $\pi(\Delta\,|\,\mathrm{H})$ representing our belief on the parameter Δ over the entire collection of our observations. Our likelihood function $\mathcal{L}$ is separable if all events are statistically independent. In this case we can then write

$$\mathcal{P}\;(\Delta\,|\,\mathrm{H}, \mathbf{d}_1, \ldots, \mathbf{d}_N,) = \frac{\pi(\Delta\,|\,\mathrm{H})}{c}\mathcal{L}(\mathbf{d}_1\,|\,\Delta, \mathrm{H}) \times \mathcal{L}(\mathbf{d}_2\,|\,\Delta, \mathrm{H}) \times \ldots \mathcal{L}(\mathbf{d}_N\,|\,\Delta, \mathrm{H}). \tag{5.17}$$

Here we can substitute the likelihood $\mathcal{L}(\mathbf{d}_i\,|\,\Delta, \mathrm{H})$, where i marks the ith observation in the list of observations from 1 to N, using Bayes' theorem. The likelihood can be substituted with the expression $\mathcal{P}'(\Delta\,|\,\mathbf{d}_i, \mathrm{H})\,/\,\pi'(\Delta\,|\,\mathrm{H})$. We use a prime notation here on the marginal posterior distribution and marginal prior distribution on the parameter Δ to denote the distributions from observation of $\mathbf{d}_i$. Marginal posterior distributions on parameters can be approximated from a Monte Carlo simulation or a Markov-chain Monte Carlo simulation. This procedure is generic and we can combine this result with the multi-model inference of Eq. (5.14) so that our parameter inference after N observations is not completely dependent on any one particular parameter hypothesis.

We have so far treated the Bayesian evidence and Bayes factor as exact quantities that can be estimated exactly. We now consider error propagation and uncertainty estimation for Bayesian evidences and Bayes factors.

5.3 Bayes Factor Uncertainty Estimation

When comparing hypotheses practically we must confront the fact that it is often too difficult to calculate the evidence analytically and so we often turn to Markov Chain Monte Carlo (MCMC) techniques to approximate the evidence [278, 276, 275, 274]. Since these methods are approximate it is useful to have some sense of the statistical or systematic uncertainties that arise from these MCMC methods[275]. In our treatment we consider model the logarithm of the evidence as a Gaussian distribution in units of log likelihood. This distribution has mean $\mu_{\widehat{\ln \mathcal{Z}}}$ representing a point estimate from the MCMC method, and standard deviation $\sigma_{\widehat{\ln \mathcal{Z}}}$ representing systematic or statistical uncertainties from the MCMC method. The logarithm of the evidence can be represented as

$$p(\widehat{\ln \mathcal{Z}}) = \left(\frac{1}{\sqrt{2\pi\sigma_{\widehat{\ln \mathcal{Z}}}}}\right) \exp\left\{-\frac{\left(\ln \mathcal{Z} - \mu_{\widehat{\ln \mathcal{Z}}}\right)^2}{2\sigma_{\widehat{\ln \mathcal{Z}}}}\right\}. \tag{5.18}$$

The Bayes factor is the ratio of two evidences and hence the logarithm of the Bayes factor is the difference of the logarithms of two evidences. Here we suppress notation on hypothesis H_A and hypothesis H_B, instead calling them A and B, such that the logarithm of the Bayes factor can be expressed as

$$\ln \mathcal{B}^A_B = \ln \mathcal{Z}_\mathrm{A} - \ln \mathcal{Z}_\mathrm{B}. \tag{5.19}$$

However, since we treat $\ln \mathcal{Z}_\mathrm{A}$ as a random variable we must deal with the uncertainty in $\widehat{\ln \mathcal{Z}_\mathrm{A}}$. The logarithm of the Bayes factor then becomes the difference between two probability distribution functions. This can be solved via convolution and has been solved for the Gaussian case [282]. From [282], we can express $\widehat{\ln \mathcal{B}_B^A}$ as a Gaussian distribution function with mean $\mu_{\widehat{\ln \mathcal{B}_B^A}} = \mu_{\widehat{\ln \mathcal{Z}_A}} - \mu_{\widehat{\ln \mathcal{Z}_B}}$ and standard deviation $\sigma_{\widehat{\ln \mathcal{B}_B^A}} = \sqrt{\sigma^2_{\widehat{\ln \mathcal{Z}_A}} + \sigma^2_{\widehat{\ln \mathcal{Z}_B}}}$. This gives us the following expression for the distribution function on the logarithm of the Bayes factor:

$$p(\widehat{\ln \mathcal{B}_B^A}) = \left(\frac{1}{\sqrt{2\pi\sigma_{\widehat{\ln \mathcal{B}_B^A}}}}\right) \exp\left\{-\frac{\left(\widehat{\ln \mathcal{B}_B^A} - \mu_{\widehat{\ln \mathcal{B}_B^A})}\right)^2}{2\sigma^2_{\widehat{\ln \mathcal{B}_B^A}}}\right\}. \tag{5.20}$$

The expression in Eq. (5.20) is a Gaussian distribution function in $\widehat{\ln \mathcal{B}_B^A}$, but we often prefer to know the estimate on $\mathcal{B}_B^A$ and so we must use a transformation of coordinates. Rather than do this explictly we could recognize that Eq. (5.20) describes a log-normal distribution, which can be expressed as

$$p(\widehat{\mathcal{B}_B^A}) = \frac{1}{\widehat{\mathcal{B}_B^A}\,\sigma_{\widehat{\ln \mathcal{B}_B^A}}}\frac{1}{2\pi}\exp\left\{-\frac{\left(\ln \widehat{\mathcal{B}_B^A} - \mu_{\widehat{\ln \mathcal{B}_B^A})}\right)^2}{2\sigma^2_{\widehat{\ln \mathcal{B}_B^A}}}\right\}. \tag{5.21}$$

It is worth noting that for a sufficiently small standard deviation on the logarithm of the Bayes factor, the probability distribution function of the Bayes factor will look approximately Gaussian in shape. This log-normal Bayes factor distribution has a median value that is identical to the point-estimate Bayes factor, $\mathcal{B}_B^A = \exp\left[\ln \mathcal{Z}_{\mathrm{A}} - \ln \mathcal{Z}_{\mathrm{B}}\right]$. The expectation value (mean) of this log-normal distribution is always right of the median, while the mode of the distribution is always left of the median. Large standard deviations on the logarithm of the evidence will create very long tails for the distribution of the Bayes factor, which makes decision-making based on Bayes factors more risky. Studies that use this estimation of the Bayes factor should consider trying to limit the error on the logarithm of the evidence.

In Sec. 5.2 we gave an equation Eq. (5.15) for combining the Bayes factors from a series of observations. If we consider the logarithm of the Bayes factor from Eq. (5.15) then combining the combined logarithm of the Bayes factor is the sum of the logarithm of the Bayes factor from each observation. In this section we have modeled the logarithm of the Bayes factor as a Gaussian distribution, and so the combined logarithm of the Bayes factor across observations is the sum of a series of Gaussian distributions, which itself is a Gaussian distribution [282]. The combined logarithm of the Bayes factor across many observations is the Gaussian distribution described by

$$p(\widehat{\ln \mathcal{B}_B^A})\,(\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_N) = \mathcal{N}\left(\mu = \sum_{i=1}^{N} \mu_i,\ \sigma = \sqrt{\sum_{i=1}^{N} \sigma_i^2}\right). \tag{5.22}$$

Here $p(\widehat{\ln \mathcal{B}_B^A})\,(\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_N)$ is the combined logarithm of the Bayes factor of hypothesis A vs hypothesis B. Here μ_i is the point-estimate log-Bayes factor from observation $\mathbf{d}_i$, and σ_i is the standard deviation of the log-Bayes factor from observation $\mathbf{d}_i$. If the μ_i and σ_i are all comparable across observations, then we note that the mean value of the combined logarithm of the Bayes factor estimate tends to grow linearly for N observations, while the standard deviation of the combined logarithm of the Bayes factor grows as $\sim \sqrt{N}$. And so, even if we cannot reduce σ_i very well for individual observations $\mathbf{d}_i$, in the long-run of observations we may expect the combined logarithmic Bayes factor to tend in a direction of statistical significance where a decision on the hypothesis can be confidently made. If the combined logarithm of the Bayes factor constantly oscillates around 0 over many observations then the error term σ will overcome the mean-value μ and the log Bayes factor will have no utility in decision making.

To illustrate this growth in Bayes factor uncertainty we consider a toy-model where three estimators of the logarithm of the Bayes factor for hypothesis H_A and hypothesis H_B. We denote these three log Bayes factor estimators as E1, E2, and E3. We consider 400 observations, where the true value of the log Bayes factor is 0.05 for every observation. This is weak or anecdotal evidence for hypothesis H_A for each observation, but the evidence accumulates over several observations. After 400 observations the combined log Bayes factor is $400 \times 0.05 = 20$, which is a combined Bayes factor of 4.8×10^8. This would provide very large evidential support for hypothesis H_A (see Table 6 for rule-of-thumb statistical significance interpretation). The first Bayes factor estimator E1 measures an unbiased estimate the log Bayes factor for each observation such that $\mu_i = 0.05$ and $\sigma_i = 0.05$. The second Bayes factor estimator E2 also has an unbiased estimate the log Bayes factor for all observations such that $\mu_i = 0.05$, but has trouble getting a good error estimate, measuring with uncertainty $\sigma_i = 0.3$ in the log Bayes factor for each observation. Finally, the third Bayes factor estimator E3 uses a method of Bayes factor estimation that is unintentionally biased such that they measure $\mu_i = 0$ for all estimates, but their estimator gives a statistical uncertainty of $\sigma_i = 0.05$. Figure 18 shows the behavior of these three Bayes factor estimators and their uncertainty at the 90% confidence interval. After 200 observations, the true combined log Bayes factor is 10, and so E1 has estimated a combined log Bayes factor of (8.8, 10, 11.2), while E2 has estimated a combined log Bayes factor of (3, 10, 17), and finally E3 has estimated a combined log Bayes factor of (-7, 0, 7) at the (5th, 50th, 95th) percentiles respectively. Here, E1 is the best estimator of the log Bayes factor, however E2 also correctly follows positive support for hypothesis H_A. We see that after 200 observations E3 has an unresolved Bayes factor with no evidential support for either hypothesis. The uncertainty confidence intervals for E2 and E3 still overlap however. After 400 observations we find E1 has measured (18.3, 20, 21.6), E2 has measured (29.9, 20, 10), and E3 has measured (-9.9, 0, 9.9). After 400 observations, the 90% confidence intervals of E2 and E3 no longer overlap. We also find that E1 and E2 have decisive levels of statistical significance to give support to hypothesis A. Meanwhile, E3 has accumulated no relative evidence between hypotheses and the uncertainty in the logarithm of the Bayes factor has become extremely large.

If we want to use Bayes factors to make decisions on the credibility of hypotheses in gravitational wave astronomy we should use the most accurate and unbiased methods for their estimation. Reducing the bias and variance of our Bayes factor estimators will play an important role in our ability to discriminate theories in physics that are supported by the data from those that are not. We now move on to practical methods for estimating the Bayesian evidence, Bayes factors, and their uncertainties via Markov-Chain Monte Carlo methods.

5.4 The Thermodynamic Integration Method for Estimating the Bayesian Evidence

The first Markov-Chain Monte Carlo (MCMC) method that we consider for Bayesian hypothesis testing is the thermodynamic integration method. Many MCMC samplers use a chain of multiple temperatures to simulate annealing as a means for gradually guiding the MCMC sampler from the prior distribution to the posterior distribution [283, 284, 285, 286]. These multiple temperatures are helpful for finding the global maxima and modes of the posterior distribution. In addition to finding the modes of the posterior distribution, this method is also useful for estimating the logarithm of the evidence. We focus on the importance of thermodynamic integration for the estimation of the logarithm of the evidence. In particular we follow the of [49, 50] which is called the method of power-posteriors. In this method, each temperature describes a tempered posterior distribution. These tempered posterior distributions are called power-posterior because they can be expressed as

$$\mathcal{P}\left(\vec{\theta}|\mathbf{d}, H\right)_{\beta} \propto \pi\left(\vec{\theta}|H\right)\mathcal{L}\left(\mathbf{d}|\vec{\theta}, H\right)^{\beta}. \tag{5.23}$$

Here $\mathcal{P}\left(\vec{\theta}|\mathbf{d}, H\right)_{\beta}$ is the power-posterior. The likelihood is raised to the power of the inverse-temperature β. The prior distribution is left identical to the standard, untempered prior distribution used in hypothesis testing and parameter inference. For a value of $\beta = 0$ the power-posterior is equivalent to the prior distribution, while for a value of $\beta = 1$ the power-posterior is equivalent to the posterior distribution. The normalization constant for the power-posterior in Eq. (6.4) is the normalizing constant for that power-posterior, given as $\mathcal{Z}(\mathbf{d}\,|\;, H)_{\beta} \equiv \int \pi\left(\vec{\theta}\,|\,\mathrm{H}\right)\mathcal{L}\left(\mathbf{d}|\vec{\theta}, \mathrm{H}\right)^{\beta} d\vec{\theta}$.

From these power-posterior distributions we can use a thermodynamic integration method [49, 50] to estimate the logarithm of the evidence. For the derivation of this thermodynamic integration method we follow [287]. We begin by considering the following expression implied by the 2nd Fundamental theorem of Calculus:

$$\ln \mathcal{Z}_{\beta=1}(\mathbf{d}) - \ln \mathcal{Z}_{\beta=0}(\mathbf{d}) = \int_0^1 \left(\frac{d\left(\ln \mathcal{Z}_{\beta}(\mathbf{d})\right)}{d\beta}\right) d\beta = \int_0^1 \frac{1}{\mathcal{Z}_{\beta}(\mathbf{d})}\frac{d\,\mathcal{Z}_{\beta}(\mathbf{d})}{d\beta} d\beta. \tag{5.24}$$

For a properly normalized prior, $\pi(\vec{\theta},\ \ln \mathcal{Z}_{\beta=0}(\mathbf{d}) = 0$. This leaves the marginal likelihood at $\beta = 1$ that we are interested in which is the untempered $\ln \mathcal{Z}(\mathbf{d})$. Now we can expand Eq. (5.24) as:

$$\ln \mathcal{Z}(\mathbf{d}) = \int_0^1 \frac{\int \frac{d}{d\beta}\left[\pi\left(\vec{\theta}\right)\mathcal{L}\left(\mathbf{d}|\vec{\theta}\right)^{\beta}\right] d\vec{\theta}}{\int \pi\left(\vec{\theta}\right)\mathcal{L}\left(\mathbf{d}|\vec{\theta}\right)^{\beta} d\vec{\theta}} d\beta. \tag{5.25}$$

Suppressing notation on θ and $\mathbf{d}$, for clarity, we find the derivative in the numerator of Eq. (5.4) as:

$$\ln \mathcal{Z} = \int_0^1 \frac{\int \pi \quad (\ln \mathcal{L}) \quad \mathcal{L}^{\beta} d\theta}{\int \pi \mathcal{L}^{\beta} d\theta} d\beta. \tag{5.26}$$

Using Bayes' theorem we can replace the numerator and denominator with $\mathcal{P}_\beta = \pi\mathcal{L}^\beta \,/\, \mathcal{Z}_\beta$ to get:

$$\ln \mathcal{Z} = \int_0^1 \frac{\int \mathcal{P}_\beta \, (\ln \mathcal{L}) \, d\theta}{\int \mathcal{P}_\beta d\theta} d\beta. \tag{5.27}$$

This simplifies to:

$$\ln \mathcal{Z} = \int_0^1 \langle \ln \mathcal{L} \rangle_{\mathcal{P}_\beta} \, d\beta. \tag{5.28}$$

Therefore, the logarithm of the evidence is given by the one dimensional integral in Eq. (5.4). Here $\langle \ln \mathcal{L} \rangle_{\mathcal{P}_\beta}$ represents the average untempered log-likelihood under the measure described by the power-posterior distribution at β. This is the average untempered log likelihood when drawing random samples from the power-posterior distribution at β. We suppress this notation to write $\langle \ln \mathcal{L} \rangle_{\mathcal{P}_\beta} \equiv \langle \ln \mathcal{L} \rangle_\beta$. With the thermodynamic integration method we have reduced a potentially large N dimensional integral into a one dimensional integral. This method is an unbiased estimator of the logarithm of the evidence provided that samples of $\langle \ln \mathcal{L} \rangle_\beta$ can be drawn in an unbiased manner from power-posteriors [288].

It is also convenient to describe additional derivatives of the thermodynamic integrand. In general, n^{th} derivatives of the form $\ln \mathcal{Z}$ can be solved by referring to Eq. 0.435 of [289][2]:

$$\frac{d^n}{d\beta^n} (\ln \mathcal{Z}) = \sum_{k=1}^{n} \frac{(-1)^{(k+1)} \binom{n}{k}}{k \mathcal{Z}^k} \frac{d^n}{d\beta^n} \left(\mathcal{Z}^k \right). \tag{5.29}$$

The first derivative, $n = 1$, we have already solved as $\langle \ln \mathcal{L} \rangle_\beta$. The next derivative, $n = 2$, was found in [290] as $\mathrm{Var}(\ln \mathcal{L})_\beta = \langle (\ln \mathcal{L})^2 \rangle_\beta - \langle \ln \mathcal{L} \rangle_\beta^2$. This is the variance of the untempered log likelihood samples when drawn from the power-posterior at β. We solve the next derivative, $n = 3$, as:

$$\frac{d^3}{d\beta^3} (\ln \mathcal{Z}) = \langle (\ln \mathcal{L})^3 \rangle_\beta + 2\langle \ln \mathcal{L} \rangle_\beta^3 - 3\langle (\ln \mathcal{L})^2 \rangle_\beta \langle \ln \mathcal{L} \rangle_\beta. \tag{5.30}$$

In practice, we have found this third derivative in Eq. (5.30) is not computationally stable in our applications in gravitational wave astronomy where the log likelihood can be $\sim \mathcal{O}(10^{-7})$. However, we observe that the pattern of the nth derivative of $\ln \mathcal{Z}$ with respect to β follows the pattern of the nth cumulant [4] of the power-posterior distribution at β [288]. See Table 7 for examples up to the 4th derivative.

[2] Note that the solution in [289] has a minor typo, which we correct here.

In fact, $\mathcal{Z}$ is analagous to a partition function in Bayesian statistical inference [288, 291]. This cumulant property is helpful because it can make computation of values of higher order derivatives more numerically stable since cumulants of order ≥ 2 are shift-invariant [4]. We can make the transformation of variables, $\widetilde{\ln \mathcal{L}} \equiv \ln \mathcal{L} - \ln \mathcal{L}_{\mathrm{max}}$ for every power-posterior before computing the numerical value of higher order derivatives of $\ln \mathcal{Z}$. We have tested this transformation rule on higher order derivatives and found it to be both accurate and numerically stable, confirming the cumulant properties of the derivatives. However, we have also found that the samples drawn from power-posterior simulation using the parallel-tempered *emcee* sampler [283, 284] may not be accurate enough to permit accurate calculation of derivatives higher than order 3 in all cases.

In Fig. 19 we show the thermodynamic integrand and the next two derivatives for a gravitational wave analysis that uses 51 temperatures. We also show the thermodynamic integrand and the next two derivatives on a logarithmic scale in β in Fig. 20 so that the curvature of the thermodynamic integrand is easier to see. In practice, plots like Figs. 19, 20 are helpful to inspect for places where the integrand may not be well sampled in β and hence require additional inverse-temperatures [292, 293, 294] for an accurate estimate of the logarithm of the evidence. Of particular note is the instability in the middle (bottom) subplot of Fig. 20 where the second (third) derivative is not perfectly smooth in β. We expect the thermodynamic integrand to be smooth and monotonically increasing as β goes from 0 to 1 [287]. In Fig. 19 there is some numerical instability at $\beta \sim 10^{-9}$. The other derivatives of the thermodynamic integrand should also be smooth. This instability implies the need for a better tempering sampler or bias-corrective terms in the sampling such as those found in the multi-tempering samplers of [295, 296]. The instability in Figs. 19, 20 is very slight however and we would not expect an effect like this to significantly impact the Bayes factor estimation.

After inspection that the MCMC sampler has produced a smooth and well-behaved thermodynamic integrand we can use numerical quadrature routines to integrate the thermodynamic integrand with respect to β. In Section 5.4.1 we describe different polynomial-based quadrature integration techniques for computing the thermodynamic integral from a finite set of inverse-temperatures β.

5.4.1 Numerical Quadrature

The thermodynamic integral in Eq. (5.4) can be estimated through numerical quadrature rules such as the trapezoidal rule or Simpson's rule. Since optimal placements of inverse-temperatures β are not typically uniformly distributed between 0 and 1 [287], it is helpful to consider integration rules that do not depend on equally spaced abscissa (β in the context of thermodynamic integration). A polynomial interpolant that does not make of equally spaced abscissa is the Newton's divided difference polynomial, see [297, 298, 5] for how to construct these polynomials. We can then integrate these interpolants to create numerical quadrature rules.

First we consider the trapezoidal rule which in the context of thermodynamic integration is

$$\widehat{\ln \mathcal{Z}}_{\text{Trapz}} = \sum_{i=0}^{N_\beta - 1} \frac{1}{2} \left(\beta_{i+1} - \beta_i\right) \left(\langle \ln \mathcal{L} \rangle_{\beta_{i+1}} + \langle \ln \mathcal{L} \rangle_{\beta_i}\right) \tag{5.31}$$

Here N_β represents the number of β being summed over in the integration estimation. The error correction term to the trapezoidal rule can be found by integrating the next-to-leading order Taylor polynomial [5], yielding:

$$\widehat{\ln \mathcal{Z}}_{\text{Trapz}+} \approx \widehat{\ln \mathcal{Z}}_{\text{Trapz}} + \sum_{i=0}^{N_\beta - 1} -\frac{1}{12} \left(\beta_{i+1} - \beta_i\right)^2 \left[f'(\beta_{i+1}) - f'(\beta_i)\right]. \tag{5.32}$$

Here $f'(\beta_i)$ represents the second derivative of $\ln \mathcal{Z}$ with respect to β. It was found in [290] that this corresponds to the variance of the untempered log likelihood as drawn from the power-posterior at β_i.

Simpson's rule for unequally spaced abscissa under Newton's divided difference interpolation [299] is:

$$\widehat{\ln \mathcal{Z}}_{\text{Simps}} = \sum_{\text{i is even, } i=0}^{N_\beta - 2} \frac{h_i + h_{i+1}}{6} \left[A \langle \ln \mathcal{L} \rangle_{\beta_i} + B \langle \ln \mathcal{L} \rangle_{\beta_{i+1}} + C \langle \ln \mathcal{L} \rangle_{\beta_{i+2}}\right], \tag{5.33}$$

for the expressions:

$$\begin{aligned} A &= \left[(2h_i - h_{i+1})\right] / h_i, \\ B &= \left[(h_i + h_{i+1})^2\right] / \left[h_i h_{i+1}\right], \\ C &= \left[(2h_{i+1} - h_i)\right] / h_{i+1}. \end{aligned} \tag{5.34}$$

Here $h_i \equiv \beta_{i+1} - \beta_i$, and $h_{i+1} \equiv \beta_{i+2} - \beta_{i+1}$. The error corrective term for Simpson's rule can thus be solved in the same manner as for the trapezoidal rule and we find:

$$\widehat{\ln \mathcal{Z}}_{\text{Simps}+} \approx \widehat{\ln \mathcal{Z}}_{\text{Simps}} + \sum_{\text{i is even, } i=0}^{N_\beta - 2} \frac{1}{72} \left(\beta_{i+2} - \beta_i\right)^2 \left(\beta_i - 2\beta_{i+1} + \beta_{i+2}\right) \frac{f''(\beta_{i+2}) - f''(\beta_i)}{\beta_{i+2} - \beta_i}. \tag{5.35}$$

Here $f''(\beta_i)$ represents the third derivative of $\ln \mathcal{Z}$ with respect to β, which is in Eq. (5.30).

The cubic integration rule for unequally spaced abscissa under Newton's divided difference interpolation can be found in [297, 298, 300] or can be derived through the tools in [5]. We present the form given in [300]:

$$\widehat{\ln \mathcal{Z}}_{\text{cubic}} = \sum_{\text{i is a multiple of 3, } i=0}^{N_\beta - 3} \frac{h_i + h_{i+1} + h_{i+2}}{12} \left[A \left\langle \ln \mathcal{L} \right\rangle_{\beta_i} + B \left\langle \ln \mathcal{L} \right\rangle_{\beta_{i+1}} + C \left\langle \ln \mathcal{L} \right\rangle_{\beta_{i+2}} + D \left\langle \ln \mathcal{L} \right\rangle_{\beta_{i+3}}\right], \tag{5.36}$$

for expressions:

$$\begin{aligned}
A &= \left[3h_i^2 - h_{i+1}^2 + h_{i+2}^2 + 2h_i h_{i+1} - 2h_i h_{i+2}\right] \,/\, \left[h_i(h_i + h_{i+1})\right], \\
B &= \left[(h_i + h_{i+1} + h_{i+2})^2 (h_i + h_{i+1} - h_{i+2})\right] \,/\, \left[h_i h_{i+1}(h_{i+1} h_{i+2})\right], \\
C &= \left[(h_i + h_{i+1} + h_{i+2})^2 (h_{i+1} + h_{i+2} - h_i)\right] \,/\, \left[h_{i+1} h_{i+2}(h_i + h_{i+1})\right], \\
D &= \left[h_i^2 - h_{i+1}^2 + 3h_{i+2}^2 - 2h_i h_{i+2} + 2h_{i+1} h_{i+2}\right] \,/\, \left[h_{i+2}(h_{i+1} + h_{i+2})\right].
\end{aligned} \tag{5.37}$$

Here we have defined $h_i \equiv \beta_{i+1} - \beta_i$, $h_{i+1} \equiv \beta_{i+2} - \beta_{i+1}$, and $h_{i+2} \equiv \beta_{i+3} - \beta_{i+2}$.

We recommend caution in using the thermodynamic integral through a higher order polynomial quadrature rule as the integrand may not be well interpolated by higher order polynomials. Additionally, higher order cumulants are difficult to estimate from sampled data [4]. Thus there may be very little incentive for going to higher order polynomial rules as improved accuracy is not always guaranteed by going to higher order polynomial integration rules [301]. Since the true value of the logarithm of the evidence is not usually known we recommend comparing the estimates from all available quadrature rules to ensure consistency.

Future studies may make use of quadrature rules from Taylor series polynomial interpolants for unequally spaced abscissa, from ratios of Taylor series polynomials through the Padé approximant [302], or other interpolant functions. Improvement in numerical integration for thermodynamic integration may also be improved by focusing on increasing the number of inverse-temperatures β and by improved placement of β.

5.4.2 Monte Carlo Error Estimation

Here we follow the discussion from [287] for estimating the Monte Carlo error in the thermodynamic integral under a generic quadrature rule. The Monte Carlo error for thermodynamic integration is the uncertainty in the estimate of the integral due to only having a finite set of samples from a Markov-Chain Monte Carlo simulation. This uncertainty enters into the integration as an uncertainty in the $\langle \ln \mathcal{L} \rangle$. The variance of the thermodynamic integral estimator, $\widehat{\ln \mathcal{Z}}$, from Monte Carlo error can be found in two steps. First, calculate the thermodynamic integral for each sample of untempered log likelihoods drawn from the power-posterior at β. For N samples drawn from each power-posterior this generates N thermodynamic integral values. The quadrature rule for the integration is generic; we can use the trapezoidal rule, Simpson's rule, etc. From these N integral values we take the sample variance and then divide by the number of samples N to generate an estimate of the population variance of the thermodynamic integration. This population variance of the logarithm of the evidence represents a long-run estimate of the variance of the estimator. This variance can be represented as:

$$\sigma^2_{\text{MC}} = \frac{1}{N}\sigma^2_{\text{sample}}. \tag{5.38}$$

Here, σ^2_{MC} represents the Monte-Carlo variance for the thermodynamic integration estimator while σ^2_{sample} is the sample variance and N represents the number of available samples. This σ^2_{MC} can be seen as the standard error of the mean value of the logarithm of the evidence due to Monte Carlo uncertainty. See Fig. 21 for a visualization of this procedure.

Repeated runs where the random seed for the MCMC analysis was changed has shown that the variance estimate from presented in [287] is a plausible confidence interval estimate for Monte Carlo error.

5.4.3 Convergence Error Estimation

The procedure for estimating the marginal likelihood from power-posterior simulation requires that the power-posteriors all converge to a final stationary distribution [49]. To do this we inspect the stability of the thermodynamic integral and integrand over the course of the MCMC analysis. To investigate the convergence of the evidence over the course of the MCMC analysis we must draw independent and identically distributed samples of the chains of the MCMC analysis at different intervals [287].

Gathering independent and identically distributed samples from a power-posterior can be done in `PyCBC` Inference by calculating the autocorrelation length of the MCMC chains of that power-posterior. In practice, `PyCBC` Inference calculates the autocorrelation length of all of the temperature chains and uses the largest posisble autocorrelation length as the autocorrelation length for all temperatures [303]. This is a safe and conservative practice for ensuring that samples drawn from the Markov Chain Monte Carlo simulation are not correlated. Thus, to track the thermodynamic integrand at various iterations in the MCMC simulation we divide the MCMC analysis into 12 equally spaced segments. In practice any number of segments will do, but it is computationally intensive to sample more segments. The segments do not need to be equally spaced in MCMC iterations but we find equally spaced segments to be useful for inspecting the progression of the thermodynamic integrand. Using this number of segments, each segment is partitioned in half, where the first half is discarded as burn-in samples, and the autocorrelation length is calculated from the remaining half of the samples. Independent samples are drawn from this half of the segment by drawing a sample for every autocorrelation length. This is the generic procedure of the $\mathrm{n_{acl}}$ algorithm implemented in `PyCBC` Inference for drawing independent samples from the Markov chains [303]. The segmenting and partitioning procedure is shown in Fig. 22.

Having drawn independent samples from 12 segments of the MCMC analysis we can visually inspect the stability of the thermodynamic integrand at 12 iterations in the MCMC analysis. We can also inspect the convergence of the thermodynamic integral. When the logarithm of the evidence has converged to $\mathcal{O}(10^{-2}$ accuracy, we usually consider the power-posteriors to have converged to their final stationary distribution. Figure 23 shows the progression of the convergence of the thermodynamic integrand as a function of the MCMC iteration. The MCMC iteration denotes how far along the MCMC has progressed, where the final iteration value indicates where the MCMC analysis was terminated. Figure 24 shows the convergence rate of the thermodynamic integral as a function of the MCMC iteration for a variety of integration techniques.

Finally, the absolute value of the difference between the last two thermodynamic integration estimates from this segmenting procedure are then used for for the standard deviation of the error for the log evidence due to convergence error, $\sigma_{\text{convergence}}$:

$$\sigma_{\text{convergence}} \sim | \ln \mathcal{Z}_{\text{partition}\, N} - \ln \mathcal{Z}_{\text{partition}\, N-1} | \tag{5.39}$$

This provides a rough estimate for estimating the consequences of potentially terminating the MCMC analysis too early.

During the development of this technique a similar technique based on a moving-block bootstrap method was developed in [304] within the context of gravitational wave analysis for error estimation of the logarithm of the evidence from the thermodynamic integration method. We have not investigated this technique thoroughly enough to compare its performance with our own method.

5.4.4 Temperature Placement Bias

The placement of inverse-temperatures β also affects the results of the numerical integration for the evidence [49, 51]. Research into the proper placement of β is ongoing in the field of Statistics [305, 287]. The studies of [50, 51, 304] have used geometric placements of β or drawn β from a power-law distribution as default temperature placement estimates. This is often a good place to start when choosing temperatures before conducting a multi-tempered MCMC analysis. The suggestion presented in [292, 294, 287] is to conduct a pilot MCMC analysis where inverse temperatures are placed according to one of these default temperature placement rules. Then a followup re-analysis can be conducted where additional inverse-temperatures are placed where the thermodynamic integrand changes rapidly or the behavior of the first derivative of the thermodynamic integrand is not smooth or well-behaved. Additional re-analyses can be conducted if the evidence integration method does not seem stable or is not measured at a precision adequate for the analysis. It is recommended in [287] that more than 40 temperatures be used, and we have used > 50 inverse-temperatures within the context of gravitational wave data analysis. In our studies we have relied primarily on visual inspection and the suggestions in this paragraph for temperature placement.

A potential improvement for temperature placement that improves upon visual inspection of the thermodynamic integrand is presented in [290]. The method of [290] calculates the intersection of the linear slopes taken from the derivatives of the thermodynamic integrand from two adjacent β. We defer to [290] for additional details on stopping rules for β placement. Finally, an additional possible improvement is to consider the placement of β from as a Bayesian inference problem. This Bayesian inference approach to numerical integration is known as Bayesian quadrature [306], and it has been specifically applied to the problem of thermodynamic integration in [307].

5.5 The Steppingstone Method for Estimating the Bayesian Evidence

We can also use another Bayesian evidence estimation technique that makes use of multi-tempered MCMC analyses. The steppingstone method is very similar to thermodynamic integration in that it requires multiple inverse-temperatures between 0 and 1 in the evidence calculation. The steppingstone method uses importance sampling between adjacent temperatures to estimate the contribution to the marginal likelihood $\mathcal{Z}$ at each interval β_{i-1}-β_i. Before we present the derivation of the steppingstone method we provide a brief, but useful derivation of another identity, called the harmonic mean estimator for the evidence [308]. For the following section we suppress use of $\vec{\theta}$ and $\mathbf{d}$ in our notation for probability functions. We also choose to use $d\theta$ in place of $d\vec{\theta}$.

For the derivation of the harmonic mean estimator we follow a simplified version of the derivation presented in [308]. We can re-express the definition of the evidence as

$$\frac{1}{\mathcal{Z}} = \frac{1}{\int \pi \, \mathcal{L} \, d\theta}. \tag{5.40}$$

Here we can substitute the numerator with $\int \pi d\theta = 1$, which gives:

$$\frac{1}{\mathcal{Z}} = \frac{\int \pi \, d\theta}{\int \pi \, \mathcal{L} \, d\theta}. \tag{5.41}$$

Now we multiply both the numerator and denominator by $\mathcal{P}/\mathcal{P}$ to get:

$$\frac{1}{\mathcal{Z}} = \left[\int \frac{\pi}{\mathcal{P}} \mathcal{P} \, d\theta \right] \Big/ \left[\int \frac{\pi \, \mathcal{L}}{\mathcal{P}} \mathcal{P} \, d\theta \right]. \tag{5.42}$$

Which we simplify using Bayes' theorem to substitute out for $1/\mathcal{P}$ to give:

$$\frac{1}{\mathcal{Z}} = \left[\int \frac{\pi \mathcal{Z}}{\pi \mathcal{L}} \mathcal{P} \, d\theta \right] \Big/ \left[\int \frac{\pi \, \mathcal{L} \mathcal{Z}}{\pi \mathcal{L}} \mathcal{P} \, d\theta \right]. \tag{5.43}$$

Cancelling out terms of π and moving terms of $\mathcal{Z}$ out of the integral to cancel, this gives:

$$\frac{1}{\mathcal{Z}} = \left[\int \frac{1}{\mathcal{L}} \mathcal{P} \, d\theta \right] \Big/ \left[\int \mathcal{P} \, d\theta \right] = \int \frac{1}{\mathcal{L}} \mathcal{P} \, d\theta. \tag{5.44}$$

Therefore we can express the inverse of the evidence as:

$$\frac{1}{\mathcal{Z}} = \langle \mathcal{L}^{-1} \rangle_{\mathcal{P}}. \tag{5.45}$$

In Eq. (5.45) we have the inverse of the evidence is equal to the average value of the inverse of the likelihood when sampled from the posterior distribution. The harmonic mean estimator of the evidence is typically poorly behaved in the context of MCMC analysis but it is a useful identity [51].

We follow [287] in the derivation of the steppingstone estimator. Recall from Eq. (5.4) that the marginal likelihood can be expressed as:

$$\ln \mathcal{Z} = \ln \mathcal{Z}_{\beta=1} - \ln \mathcal{Z}_{\beta=0}, \tag{5.46}$$

which is equivalent to:

$$\mathcal{Z} = \frac{\mathcal{Z}_{\beta=1}}{\mathcal{Z}_{\beta=0}}. \tag{5.47}$$

Without loss of generality we can consider a set of 100 inverse-temperatures β uniformly distributed between 0 and 1 such that Eq. (5.47) can be expressed as

$$\mathcal{Z} = \frac{\mathcal{Z}_{\beta=0.01}}{\mathcal{Z}_{\beta=0}} \times \frac{\mathcal{Z}_{\beta=0.02}}{\mathcal{Z}_{\beta=0.01}} \times \ldots \times \frac{\mathcal{Z}_{\beta=0.99}}{\mathcal{Z}_{\beta=0.98}} \times \frac{\mathcal{Z}_{\beta=1}}{\mathcal{Z}_{\beta=0.99}}. \tag{5.48}$$

This generalizes to

$$\mathcal{Z} = \prod_{i=1}^{N_\beta} \frac{\mathcal{Z}_{\beta_i}}{\mathcal{Z}_{\beta_{i-1}}}. \tag{5.49}$$

Here we use the ordering on β, as $\beta_0 = 0 < \beta_1 < \ldots < \beta_{N_\beta - 1} < \beta_{N_\beta} = 1$. Finally, then, consider the evidence for the power-posterior at inverse-temperature β_i given as:

$$\mathcal{Z}_{\beta_i} = \int \pi \mathcal{L}^{\beta_i} \, d\theta. \tag{5.50}$$

We now divide by 1 using $\int \pi \, d\theta$ and multiply by 1 using $\mathcal{P}_{\beta_{i-1}} \,/\, \mathcal{P}_{\beta_{i-1}}$ in the numerator and denominator to get:

$$\mathcal{Z}_{\beta_i} = \left(\int \frac{\pi \mathcal{L}^{\beta_i}}{\mathcal{P}_{\beta_{i-1}}} \mathcal{P}_{\beta_{i-1}} \, d\theta \right) \Big/ \left(\int \frac{\pi}{\mathcal{P}_{\beta_{i-1}}} \mathcal{P}_{\beta_{i-1}} \, d\theta \right). \tag{5.51}$$

Using Bayes' theorem we substitute $\mathcal{P}_{\beta_{i-1}} = (1/\mathcal{Z}_{\beta_{i-1}})\, \pi \, \mathcal{L}^{\beta_{i-1}}$ to get:

$$\mathcal{Z}_{\beta_i} = \left(\int \frac{\pi \mathcal{L}^{\beta_i} \mathcal{Z}_{\beta_{i-1}}}{\pi \mathcal{L}^{\beta_{i-1}}} \mathcal{P}_{\beta_{i-1}} \, d\theta \right) \Big/ \left(\int \frac{\pi \mathcal{Z}_{\beta_{i-1}}}{\pi \mathcal{L}^{\beta_{i-1}}} \mathcal{P}_{\beta_{i-1}} \, d\theta \right). \tag{5.52}$$

Terms of $\mathcal{Z}_{\beta_{i-1}}$ are independent of θ and so can be moved out of the integral where they cancel. We can also cancel terms of π to get

$$\mathcal{Z}_{\beta_i} = \left(\int \frac{\mathcal{L}^{\beta_i}}{\mathcal{L}^{\beta_{i-1}}} \mathcal{P}_{\beta_{i-1}} \, d\theta \right) \Big/ \left(\int \frac{1}{\mathcal{L}^{\beta_{i-1}}} \mathcal{P}_{\beta_{i-1}} \, d\theta \right). \tag{5.53}$$

Finally we recognize that in the denominator we have Eq. (5.45) for the inverse of the evidence at the inverse-temperature β_{i-1}. With some additional simplifications in the numerator we have

$$\mathcal{Z}_{\beta_i} = \mathcal{Z}_{\beta_{i-1}} \int \mathcal{L}^{\beta_i - \beta_{i-1}} \mathcal{P}_{\beta_{i-1}} \, d\theta. \tag{5.54}$$

Thus we arrive at the key ingredient for the steppingstone estimator:

$$\frac{\mathcal{Z}_{\beta_i}}{\mathcal{Z}_{\beta_{i-1}}} = \int \mathcal{L}^{\beta_i - \beta_{i-1}} \mathcal{P}_{\beta_{i-1}} \, d\theta = \langle \mathcal{L}^{\beta_i - \beta_{i-1}} \rangle_{\mathcal{P}_{\beta_{i-1}}}. \tag{5.55}$$

We suppress some of the notation in Eq. (5.55) such that $\langle \mathcal{L}^{\beta_i - \beta_{i-1}} \rangle_{\mathcal{P}_{\beta_{i-1}}} \equiv \langle \mathcal{L}^{\beta_i - \beta_{i-1}} \rangle_{\beta_{i-1}}$ and combine Eq. (5.55) into Eq. (5.49) to give the steppingstone estimator for the evidence:

$$\mathcal{Z} = \prod_{i=1}^{N_\beta} \langle \mathcal{L}^{\beta_i - \beta_{i-1}} \rangle_{\beta_{i-1}}. \tag{5.56}$$

Some care needs to be taken in the implementation of Eq. (5.56) as the form presented is not numerically stable and we often must use the log likelihood and log evidence in place of the likelihood and the evidence. A numerically stable form of the logarithm of Eq. (5.56) is presented in [51]. It is noted by [51] that the logarithm of the evidence in the steppingstone estimator exhibits some level of bias as an estimator of the marginal likelihood. This bias is small when many inverse-temperatures are used, and it was shown in [51] that the steppingstone estimator in many cases outperforms the trapezoidal rule for thermodynamic integration with the same inverse-temperatures.

5.5.1 Monte Carlo Error

In [51] there is also an expression for the estimated variance of the logarithmic steppingstone estimator using an approximation method called the δ method [309]. The δ method makes use of the asymptotic normal behavior of an estimator due to the central limit theorem to approximate the variance of the estimator. The expression in [51] for the variance of the logarithm of the evidence is however not presented in a form that makes use of the log likelihood and thus can suffer from numerical instability. We use a numerically stabilized version of the variance estimator in our gravitational wave analysis as implemented in `PyCBC` Inference. We have found the variance estimate from the δ method is typically comparable to the Monte Carlo error estimate in the thermodynamic integration method. Repeated runs where the random seed for the MCMC analysis was changed has indicated that the variance estimate from presented in [51] is a plausible confidence interval estimate for the Monte Carlo error.

5.5.2 Convergence Error

The method for calculating the error on the steppingstone estimator due to convergence error is algorithmically identical to the thermodynamic integration method presented in Section 5.4.3. The convergence of the evidence estimation for the steppingstone estimator for a gravitational wave analysis can be seen in Fig. 24. We rely on the stability of the thermodynamic integrand in addition to inspecting the convergence of the steppingstone log evidence estimate to decide if the method has converged to a stationary estimate of the log evidence. This convergence can be influenced by number of inverse-temperatures used and the placement of those temperatures.

5.5.3 Temperature Placement Bias

The problem of optimal placement of inverse-temperatures β remains an active area of research [51, 287]. For a large number of inverse-temperatures β the study of [51] showed that the steppingstone estimator converges to the correct evidence estimate faster than the thermodynamic integration method and [304] showed that this is also true in applications to gravitational wave astronomy. Both [51] and [304] indicate that the evidence estimation of the steppingstone estimator may be less sensitive to temperature placement than the thermodynamic integration method implemented via the trapezoidal rule. We therefore do not inspect temperature placement for the steppingstone estimator but instead rely on inspection of the thermodynamic integrand for insight on where to place β.

5.6 The Savage-Dickey Density Ratio Method for Estimating Bayes Factors

We also consider another MCMC method for estimating Bayes factors to verify the results of the multi-tempered methods. The Savage-Dickey density ratio method requires two models, wherein one hypothesis is nested in the parameter space of the other hypothesis. This requirement that the two models be nested is very restrictive on the types of hypotheses that can be tested, but it has been used considerably in the field of Cosmology [275]. We derive the method following [310]. We can consider two hypotheses, $\mathrm{H}_{\mathrm{simple}}$ and $\mathrm{H}_{\mathrm{complex}}$. The hypothesis $\mathrm{H}_{\mathrm{simple}}$ is nested within $\mathrm{H}_{\mathrm{complex}}$ via a parameter A. When A in $\mathrm{H}_{\mathrm{complex}}$ tends towards a critical value, $\mathrm{H}_{\mathrm{complex}}$ becomes identical to $\mathrm{H}_{\mathrm{simple}}$. In our use of the Savage-Dickey density ratio this critical value is $A = 0$. We can formalize this by stating the prior distributions from these hypotheses in the following way:

$$\pi(\vec{\theta}_{\mathrm{simple}} \,|\, \mathrm{H}_{\mathrm{simple}}) \;\equiv\; \pi\left(\alpha, \beta, \gamma, \ldots\right\} \,|\, \mathrm{H}_{\mathrm{simple}}) \tag{5.57}$$

$$\pi(\vec{\theta}_{\mathrm{complex}} \,|\, \mathrm{H}_{\mathrm{complex}}) \;\equiv\; \pi\left(\alpha, \beta, \gamma, \ldots, A\right\} \,|\, \mathrm{H}_{\mathrm{complex}}) \,. \tag{5.58}$$

Here $\pi(\vec{\theta}_{\mathrm{simple}} \,|\, \mathrm{H}_{\mathrm{simple}})$ represents the prior distribtuion on parameters in the simple hypothesis, while $\pi(\vec{\theta}_{\mathrm{complex}} \,|\, \mathrm{H}_{\mathrm{complex}})$ represents the prior distribution on parameters in the complex hypothesis. The parameters α, β, γ are nuisance parameters in the context of this Savage-Dickey density ratio method, and are shared parameters between the simple and complex hypotheses. We abbreviate the notation by writing the prior under the simple hypothesis as $\pi_{\text{complex}}(\psi)$ and the prior under the more complex hypothesis as $\pi_{\text{complex}}(\psi, A)$. Here the dependence on hypotheses is denoted by the subscript simple or complex, and ψ denotes all parameters that are not A. In order for the Savage-Dickey Density Ratio method to hold true the following expression must be satisfied:

$$\lim_{A\to 0} \pi_{\text{complex}}(\psi|A) = \pi_{\text{simple}}(\psi). \tag{5.59}$$

This states that in the limit that A tends to 0, the prior parameter space in the complex hypothesis becomes identical to the prior parameter space in the simple hypothesis. Under these conditions we then consider the definition of the Bayes factor:

$$\mathcal{B}^{\text{complex}}_{\text{simple}} \equiv \frac{\mathcal{Z}_{\text{complex}}(\mathbf{d})}{\mathcal{Z}_{\text{simple}}(\mathbf{d})}. \tag{5.60}$$

The denominator can be expressed as

$$\mathcal{Z}_{\text{simple}}(\mathbf{d}) = \int \pi_{\text{simple}}(\psi)\, \mathcal{L}_{\text{simple}}(\mathbf{d} \,|\, \psi)\, d\psi. \tag{5.61}$$

Since the models are nested, the prior and the likelihood under the complex hypothesis at $A = 0$ is equivalent to the prior and likelihood under the simple hypothesis. This gives

$$\pi_{\text{complex}}(\psi, A = 0) = \pi_{\text{simple}}(\psi) \tag{5.62}$$

and

$$\mathcal{L}_{\text{complex}}(\mathbf{d} \,|\, \psi, A = 0) = \mathcal{L}_{\text{simple}}(\mathbf{d} \,|\, \psi). \tag{5.63}$$

If we substitute Eqss. (5.62) and (5.63) into Eq. (5.61) we get:

$$\mathcal{Z}_{\text{simple}}(\mathbf{d}) = \int \pi_{\text{complex}}(\psi, A = 0)\, \mathcal{L}_{\text{complex}}(\mathbf{d} \,|\, \psi, A = 0)\, d\psi. \tag{5.64}$$

Integrating this over all ψ, leaves the $A = 0$ unintegrated over leaving us with $\mathcal{Z}_{\text{simple}} = \mathcal{L}_{\text{complex}}(\mathbf{d} \,|\, A = 0)$. Using Bayes theorem, we can rewrite $\mathcal{L}_{\text{complex}}(\mathbf{d} \,|\, A = 0) = [\mathcal{P}_{\text{complex}}(A = 0 \,|\, \mathbf{d})\, \mathcal{Z}_{\text{complex}}(\mathbf{d})] \,/\, \pi_{\text{complex}}(A = 0)$. This leaves us with:

$$\mathcal{Z}_{\text{simple}}(\mathbf{d}) = \frac{\mathcal{P}_{\text{complex}}(A = 0 \,|\, \mathbf{d})\, \mathcal{Z}_{\text{complex}}(\mathbf{d})}{\pi_{\text{complex}}(A = 0)}. \tag{5.65}$$

Finally, then we arrive at the Savage Dickey density ratio Bayes factor

$$\mathcal{B}_{\text{simple}}^{\text{complex}} = \frac{\pi_{\text{complex}}\,(A=0)}{\mathcal{P}_{\text{complex}}\,(A=0\,|\,\mathbf{d})}. \tag{5.66}$$

Here the Bayes factor for the complex hypothesis relative to the simple hypothesis is the ratio of the prior density at the critical value A=0 to the posterior density at the critical value A=0 when sampling from the complex hypothesis' parameter space. This all assumes that the simple hypothesis is nested within the complex hypothesis. A more rigorous derivation that avoids potential division by 0 is presented in [311] where an additional term is multiplied to Eq. (6.31). For our purposes we will not need to use this correction term.

The Savage-Dickey density ratio method requires the estimation of probability densities at a point in the marginal prior and posterior distributions for the nesting parameter A. We explore four methods for accurate probability density estimation. The first two methods are simple histogram approaches, the third estimate is a Gaussian kernel density estimator, and the last method is a cubic spline density estimator.

5.6.1 Histogram Methods

A simple method for estimating the probability density function is to histogram the samples by counts and then normalize the histogram to integrate to unity. Under this methodology the only relevant parameter to fitting the histogram to the data is in the choice of bin-width, sometimes called bandwidth.

We use two bin-width algorithms to fit a histogram to the data. The first method is Scott's rule [312] . Scott's rule is optimal if the underlying density of the function is normally distributed. The bin-width h for this rule is defined as:

$$h = \frac{3.5\,\sigma}{N^{1/3}}. \tag{5.67}$$

Here σ represents the sample standard deviation of the data, and N represents the number of samples in the data.

The second binning method is the Freedman-Diaconis binning method [313]. The binning method makes use of the interquartile range (IQR) of the data rather than the standard deviation σ. The IQR is defined as the difference between the 75th and 25th percentiles of the data. The bin-width h in the Freedman-Diaconis is

$$h = \frac{2\,\text{IQR}}{N^{1/3}.} \tag{5.68}$$

Where N represents the number of samples in the data.

There are other bin-width algorithms available but we have found that these two methods are very robust towards estimating the density functions of some common and expected probability distribution functions.

5.6.2 A Gaussian Kernel Density Estimator

We also use a Gaussian kernel density estimator available in the Python package GetDist [20] for probability density estimation. GetDist is intended to accurately estimate one-dimensional and two-dimensional posterior probability distribution functions from sample data produced from Bayesian MCMC analyses. A Gaussian kernel density estimator uses small truncated-Gaussian distributions centered at the samples of the data rather than individual bin-counts such as in a histogram. The Gaussian kernel density estimator then combines the sum of the these truncated Gaussian distributions to create a smooth probability distribution function.

The advantage that GetDist offers over other Gaussian kernel density estimators is that it comes with a robust linear-boundary bias correction to the standard Gaussian kernel density estimator. Sharp boundaries on the distribution function can cause bias to the probability distribution function estimation for Gaussian kernel density estimators [20]. The Savage-Dickey density ratio method often requires us to estimate the density of the posterior distribution at the boundary of the distribution and so an estimator that can give an unbiased estimate of this density is highly desirable. There are additional features and bandwidth optimization algorithms present in GetDist for accurate density estimation but we defer to [20] for the full details.

5.6.3 The Logspline Estimator

The logspline estimator of [314] is written as a software package in R. The logspline estimator estimates the probability density function of a set of data through a univariate cubic spline fit to the logarithm of the probability density. The package places knots of a cubic spline along the axis of the data through a likelihood function evaluated through a data censoring procedure. The package uses Bayesian model-selection based on the Akaike Information Criterion (AIC) [315] and the Bayesian Information Criterion (BIC) [316] to ensure goodness of fit and to avoid overfitting to the data.

The AIC is a model selection routine founded in information theory, while the BIC is based on an approximation of the Bayes factor. The full details of the logspline density estimator are beyond the scope of this dissertation. We utilize the maximum likelihood (best) fit to the probability distribution function from the packages' model selection routine. The logspline density estimator of [314] is considered one of the most accurate one-dimensional probability distribution function estimators in the R language [317] and the logspline package comes specifically recommended in [310] for the Savage-Dickey density ratio test.

5.6.4 Error Analysis for the Savage-Dickey Density Ratio

There are many different approaches to estimate the uncertainty in the Savage-Dickey density ratio method. The most straightforward method is to conduct multiple MCMC analyses using the same prior distribution to get multiple statistically independent estimates of the posterior distribution. The histogram methods using Scott's binning rule, the Freedman-Diaconis binning rule, the GetDist Gaussian kernel density estimator, and the logspline density estimator can then be checked for the posterior distribution for each MCMC analysis. A distribution of the Savage-Dickey density ratio estimates can then be constructed and a 90% confidence interval can be esitmated. Bayesian MCMC analyses, especially in the context of gravitational wave data analysis, can be computationally expensive and so we consider a different approach.

A simple method for a constructing a confidence interval on a test statistic is through a resampling technique known as the bootstrap method [318]. The bootstrap method resamples the data N times with replacement providing a set of N datasets to estimate a test statistic on. In our case the test statistic is the Savage-Dickey density ratio from our four density estimators. A confidence interval can be estimated for these point-estimates of the Savage-Dickey density ratio by resampling the marginal posterior distribution with replacement and calculating the Savage-Dickey density ratio. The bootstrap method is resampling technique known as cross-validation. There are many different kinds of cross-validation techniques available with potential to improve our confidence interval estimation [319, 320], but we do not explore them here.

Now that we have established all of the tools that we need for Bayesian hypothesis testing we discuss their application in gravitational wave analysis.

5.7 Prior Distributions for Gravitational Wave Analysis

Choosing a set of prior distributions for our Bayesian inference requires us to make choices on acceptable distributions of plausible values for parameters. Choice of these parameters and the probability distributions we want to ascribe to them describe the physics of the gravitational waves that we wish to model. Here we briefly describe some considerations when choosing prior distributions for Bayesian parameter estimation for gravitational waves from compact binary coalescences. The choices of parameters that describe gravitational waves have been described in Chapters 2, 3, and 4, although not within the context of Bayesian parameter estimation. In this section we follow the discussion of the choice of prior distributions for Bayesian parameter estimation from [303].

There are a number of parameters that describe the gravitational waves radiated from a compact binary coalescence. These parameters include, the component masses $m_{1,2}$ of the binary as well as the three-dimensional spin vectors $\vec{S}_{1,2}$ of each of the compact objects [321]. This gives us 8 intrinsic parameters to describe gravitational waves from compact binary mergers such as binary black hole mergers. There are additional possible intrinsic parameters that we can consider when modeling the physics of binary black hole mergers, and there are yet more parameters that we will need to model the physics of compact objects with matter such as binary neutron stars and neutron star-black hole binaries. We do not consider them in this section.

Additional parameters are needed to describe gravitational waves from compact binary mergers. The gravitational waveform observed by an Earth-based detector network depends on six additional parameters: the signal's time of arrival $t_{\mathrm{coalescence}}$, the binary's luminosity distance d_{L}, and four Euler angles that describe the transformation from the binary's radiation frame to the detector network frame [322]. These angles are the binary's right ascension α, declination δ, a polarization angle Ψ, and the inclination angle ι. The inclination angle is the angle between the binary's angular momentum axis $\hat{L}$ and the line of sight. These parameters are typicaly considered as extrinsic parameters since they are not properties of the source of the gravitational waves. Bearing all of these parameters in mind this brings the dimensionality of the problem of Bayesian inference up to 14.

Modeling gravitational waves from compact binaries is currently a challenging problem for Bayesian inference due to the dimensionality of the signal parameter space. This is further complicated by degeneracies between many of the signal's parameters. For example, to leading order the gravitational waveform depends on the chirp mass $\mathcal{M} \equiv (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$ [35]. The mass ratio q enters the waveform at higher orders and is more difficult to measure. This results in an amplitude-dependent degeneracy between the component masses [323]. Similarly, the binary's mass ratio can be degenerate with its spin [324]. In Chapter 6 of this dissertation we will explore difficulties of Bayesian inference due to parameter degeneracy.

Given a set of parameters $\vec{\theta}$, one can obtain a model of the gravitational-wave signal from a binary merger using a variety of different waveform modeling methods, including, but not limited to: post-Newtonian theory (see e.g. Ref. [36] and references therein) and analytic models calibrated against numerical simulations [37, 38, 39, 40, 41, 42, 43]. The specific choice of waveform model and marginal prior distribution probabilities on parameters specify a $\pi(\vec{\theta}\,|\,\mathrm{H})$ for a Bayesian analysis. The decision of what parameters $\vec{\theta}$ and prior $\pi(\vec{\theta}\,|\,\mathrm{H})$ to use depends on the physics that we wish to explore in the model. A variety of waveform models are available for use in `PyCBC` Inference, either directly implemented in `PyCBC` or via the LIGO Algorithm Library (LAL) [325].

5.8 The Likelihood function for Gravitational Wave Analysis

In this section we follow the description of the likelihood function given in [303]. The gravitational wave strain data observed by gravitational-wave detector networks enters Bayes' theorem through the likelihood $\mathcal{L}(\mathbf{d}\,|\,\vec{\theta})$ in Eq. (5.8). Currently, `PyCBC` Inference assumes that each detector produces stationary, Gaussian noise $n_i(t)$ that is uncorrelated between the detectors in the network. The observed data is then $d_i(t) = n_i(t) + s_i(t)$, where $s_i(t)$ is the gravitational waveform observed in the ith detector. For detectors that are not identical and co-located (as in the case of the Advanced LIGO-Virgo network), each detector observes a slightly different waveform due to their different antennae patterns which are functions of the sky position (right ascension and declination) and polarization [322].

Under these assumptions, the appropriate form of $\mathcal{L}(\mathbf{d}\,|\,\vec{\theta})$ is the likelihood for a signal of known morphology in Gaussian noise (see e.g. Ref. [34] for its derivation), which is given by

$$\begin{aligned}\mathcal{L}(\mathbf{d}\,|\,\vec{\theta}) &\propto \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\langle\tilde{n}_i(f)|\tilde{n}_i(f)\rangle\right] \\ &= \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\left\langle\tilde{d}_i(f)-\tilde{s}_i(f,\vec{\theta})|\tilde{d}_i(f)-\tilde{s}_i(f,\vec{\theta})\right\rangle\right],\end{aligned} \tag{5.69}$$

where N is the number of detectors in the network. The constant of proportionality for this Gaussian likelihood is not important for Bayesian inference under MCMC simulation since the constant drops out when proposals are made by the MCMC for how to traverse the parameter space [326]. This proportionality constant also cancels out when calculating Bayes factors. The inner product $\langle\tilde{a}\,|\,\tilde{b}\rangle$ is

$$\left\langle\tilde{a}_i(f)\,|\,\tilde{b}_i(f)\right\rangle = 4\mathcal{R}\int_0^\infty \frac{\tilde{a}_i(f)\tilde{b}_i(f)}{S_n^{(i)}(f)}\mathrm{d}f\,, \tag{5.70}$$

where $\mathcal{R}$ denotes the real part of the integral, $S_n^{(i)}(f)$ is the power spectral density of the of the ith detector's noise. Here, $\tilde{d}_i(f)$ and $\tilde{n}_i(f)$ are the frequency-domain representations of the data and noise, obtained by a Fourier transformation of $d_i(t)$ and $n_i(t)$, respectively. The model waveform $\tilde{s}_i(f,\vec{\theta})$ is the waveform in the frequency domain. For a specified prior distribution on parameters, `PyCBC` Inference calculates Eq. (5.69) and completes Bayesian inference for estimating parameter posterior distributions and estimating Bayesian evidences.

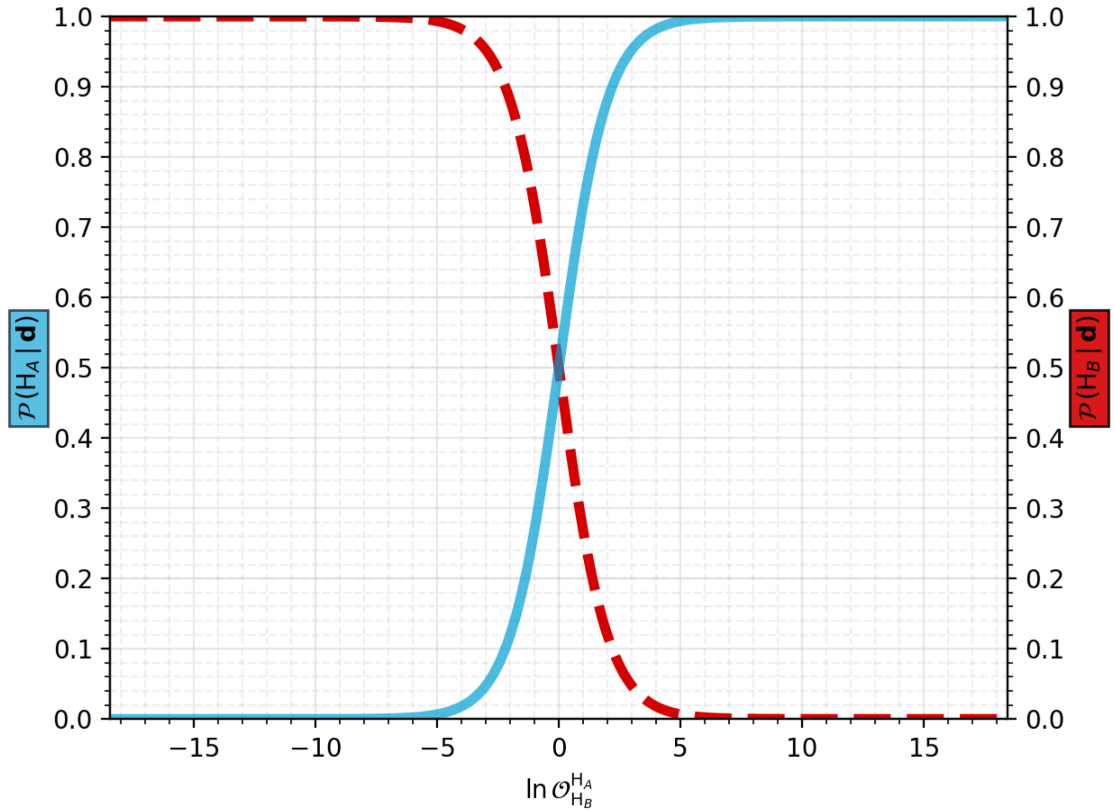


Figure 16: (*Light blue, solid*) The probability of hypothesis H_A being favored over hypothesis H_B after observation of the data $\mathbf{d}$ when considering calculating the natural log of the odds ratio for each hypothesis. (*Red, dashed*) The posterior probability of H_B is the complement of the posterior probability of H_A if there are only two hypotheses available to test. When $\log_{10} \mathcal{O} = 0$, the probability for each hypothesis is 50%.

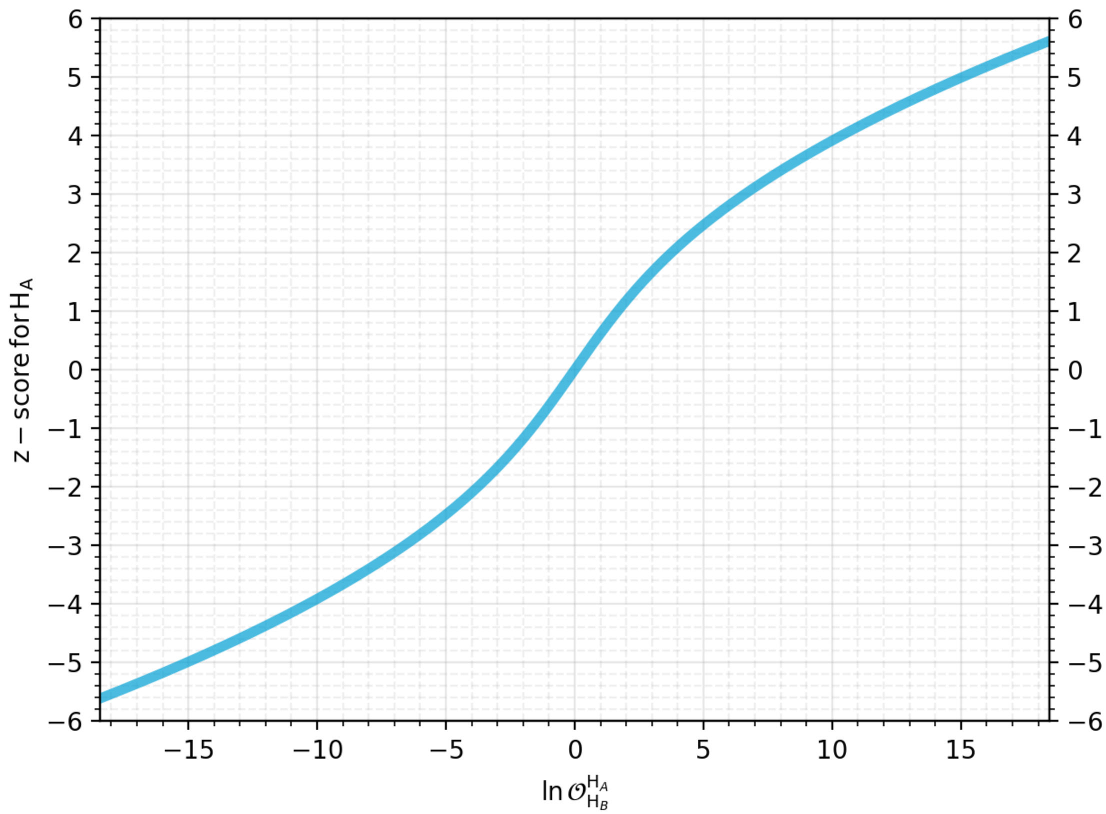


Figure 17: The z-score pertaining to the same level of probability for hypothesis 1 being favored over hypothesis 2 when considering the $\ln \mathcal{O}^{\mathrm{H}_A}_{\mathrm{H}_B}$. When $\ln \mathcal{O}^{\mathrm{H}_A}_{\mathrm{H}_B} = 0$, the z-score is 0σ and the probability for each hypothesis is 50%. A z-score of $> 5\sigma$ has the same probability value as an odds ratio of $> 10^7$.

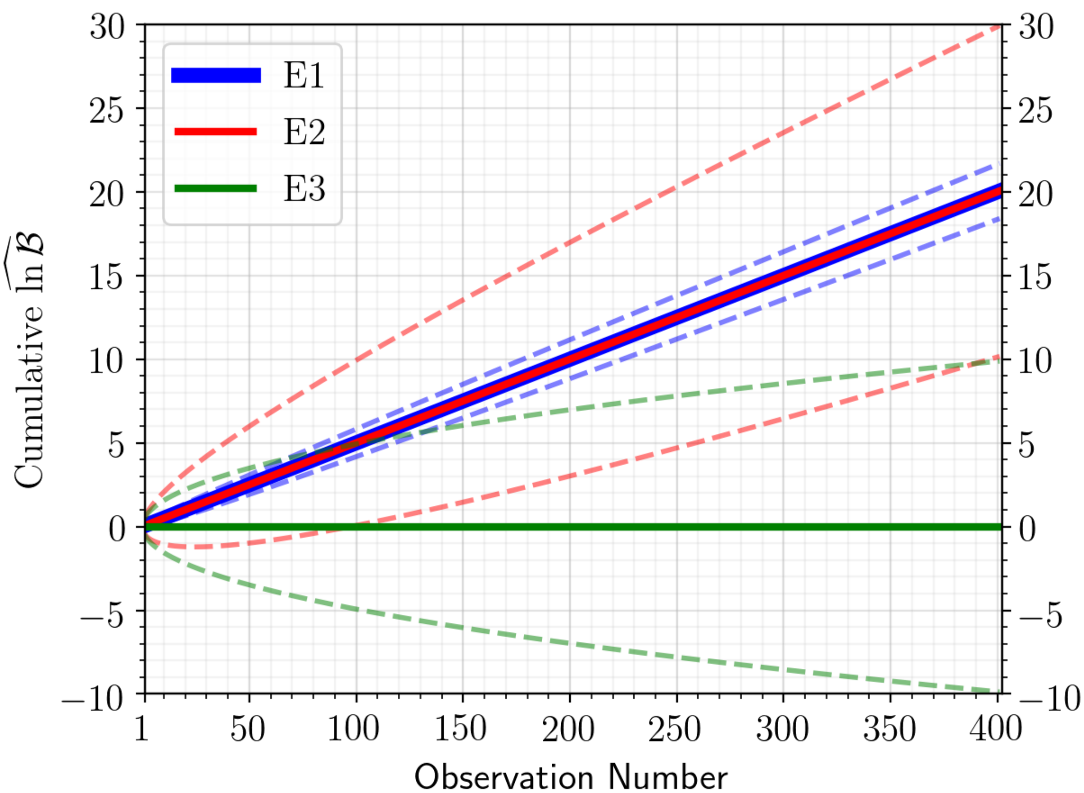


Figure 18: The divergence of statistical inference for three hypothetical log Bayes factor estimators E1, E2, E3 who are estimating the combined logarithm of the Bayes factor over many observations. Each observation has a true ln $\mathcal{B}$ of hypothesis H_A relative to H_B of 0.05, indicating no statistical significance at the level of a single observation. Over 400 events however the combined logarithm of the Bayes factor is 20. The estimator E1 (light blue) estimates the logarithm of the Bayes factor for each observation through an unbiased method, measuring a mean value of $\mu_i = 0.05$ for each observation with standard deviation $\sigma_i = 0.05$. Here i denotes the observation number. The estimator E2 (light red) estimates the logarithm of the Bayes factor for each observation through an unbiased method, measuring a mean value of $\mu_i = 0.05$ for each observation with a very large standard deviation $\sigma_i = 0.3$. The estimator E3 (light green) estimates the logarithm of the Bayes factor for each observation through a slightly biased method, instead measuring a mean value of $\mu_i = 0$ for each observation but has a small measuring uncertainty of $\sigma_i = 0.05$. The inferences of these estimators diverge after many observations due to systematic and statistical uncertainty.

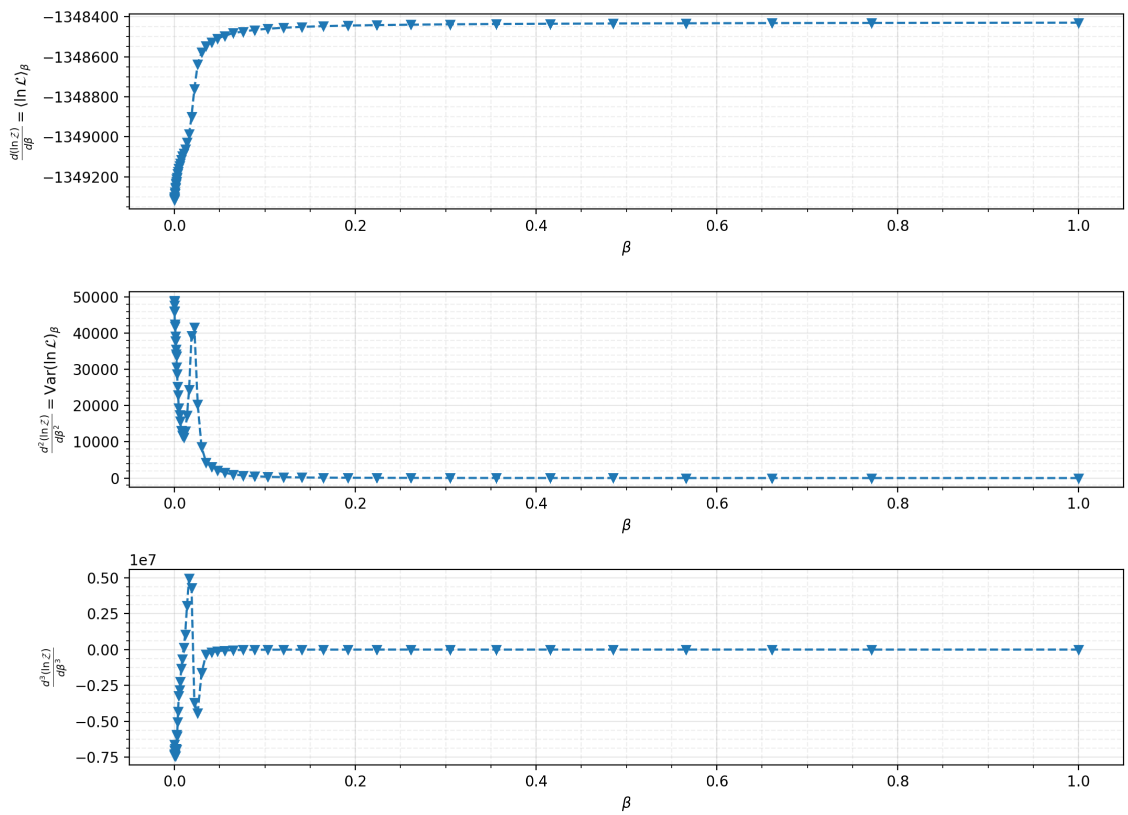


Figure 19: The subplots of the thermodynamic integrand and subsequent derivatives of the thermodynamic integral. (*Top*) The thermodynamic integrand when compared to the inverse-temperature β. The curve should be smooth and montonic, however it is very difficult to inspect the integrand on a linear β scale. (*Middle*) The second derivative of the logarithm of the evidence is the variance of the power-posterior at an inverse temperature β. There is some indication that an inflection point happens in the curvature of the integrand at high temperature. (*Bottom*) The third derivative of the logarithm of the evidence is also the third-order cumulant of the power-posterior distributions at an inverse-temperature β. It is difficult to inspect the behavior of this derivative on the linear β scale.

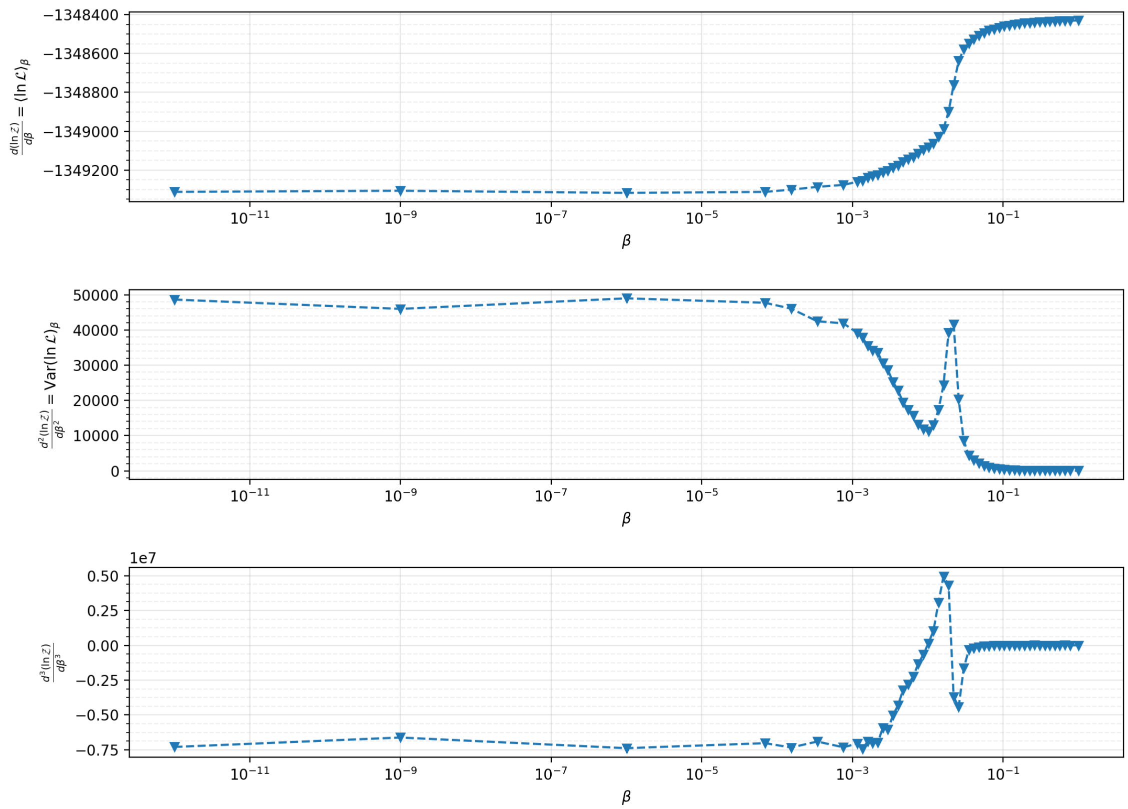


Figure 20: The subplots of the thermodynamic integrand and subsequent derivatives of the thermodynamic integral. (*Top*) The thermodynamic integrand when compared to the inverse-temperature β. The curve should be smooth and montonic, however there is some indication at $\beta = 10^{-9}$ that this condition is not strictly met in the Markov Chain Monte Carlo simulation. (*Middle*) The second derivative of the logarithm of the evidence is the variance of the power-posterior at an inverse temperature β. This function should also be smooth however there is some indication that at high temperature that the derivatives are not stable. (*Bottom*) The third derivative of the logarithm of the evidence is also the third-order cumulant of the power-posterior distributions at an inverse-temperature β. Here we can see that the derivatives are not very sable or smooth. This may motivate moving our analysis to new multi-tempered samplers that are optimized for thermodynamic integration.

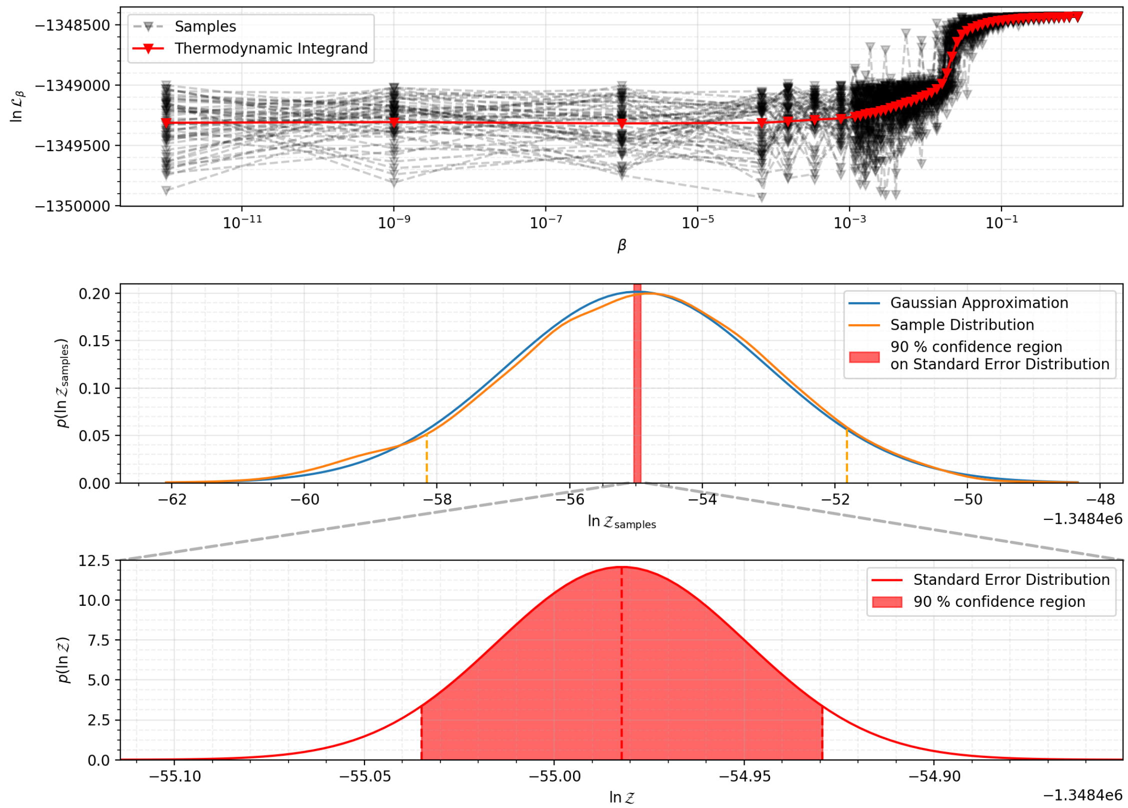


Figure 21: The first subplot denotes the untempered log-likelihood samples when drawn from the power-posteriors at β. The expectation value of the untempered log-likelihood when drawn from these power-posteriors is the thermodynamic integrand and is plotted in red. The thermodynamic integral over all geometric paths given from the samples is drawn in the second subplot. The sample-log-integral distribution is approximately a Gaussian distribution. The standard error of the mean value of the log evidence is given by the sample standard deviation divided by the square root of the number of samples. The 90% confidence interval on the sample distribution in the log-evidence is drawn in dashed orange lines. The 90% confidence region from this standard error is shaded in red. The final subplot is a zoom-in on this 90% confidence region showing the error estimate on the thermodynamic integral due to Monte Carlo sampling.

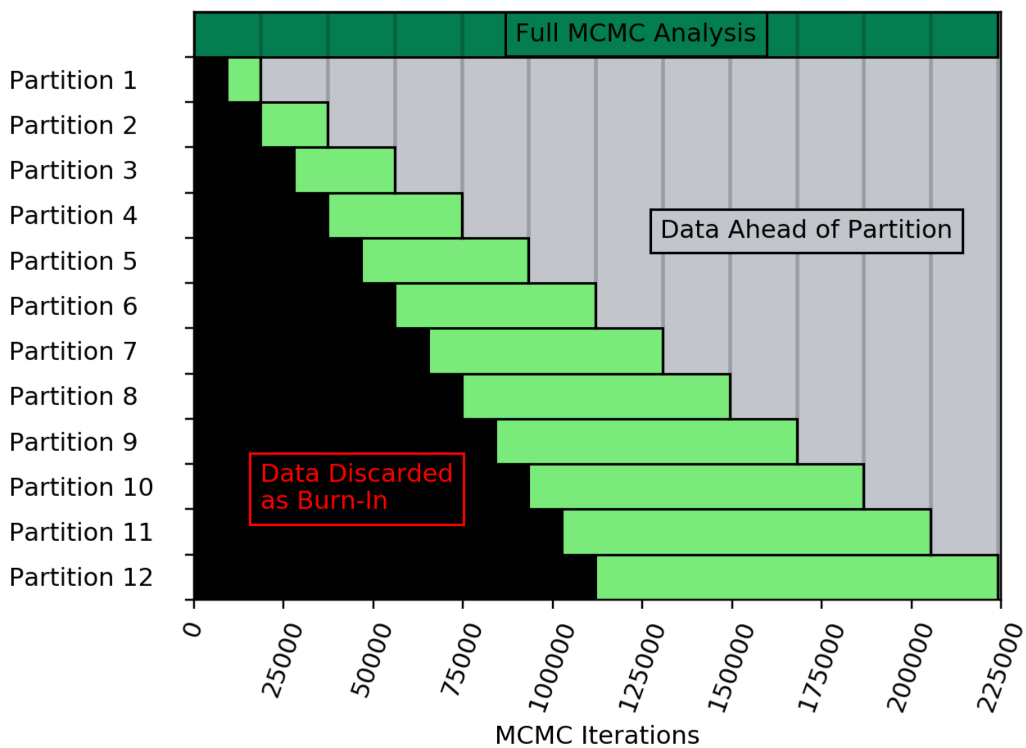


Figure 22: The partitioning of the MCMC analysis to check on the convergence of the thermodynamic integrand and the thermodynamic integration. The dark-green bar at the top represents all of the samples collected by the MCMC analysis. This analysis is divided into 12 partitions represented by the dark gray lines. The light-green segments represent partitions of the analysis that independent samples can be drawn from. The black region represents samples discarded as burn-in samples for the MCMC. The light grey region represents data that is ahead of the partition and thus not used in drawing independent samples for that partition. Partition 12 produces the identical samples as drawing independent samples according to the n_{acl} algorithm from PyCBC at the end of the analysis.

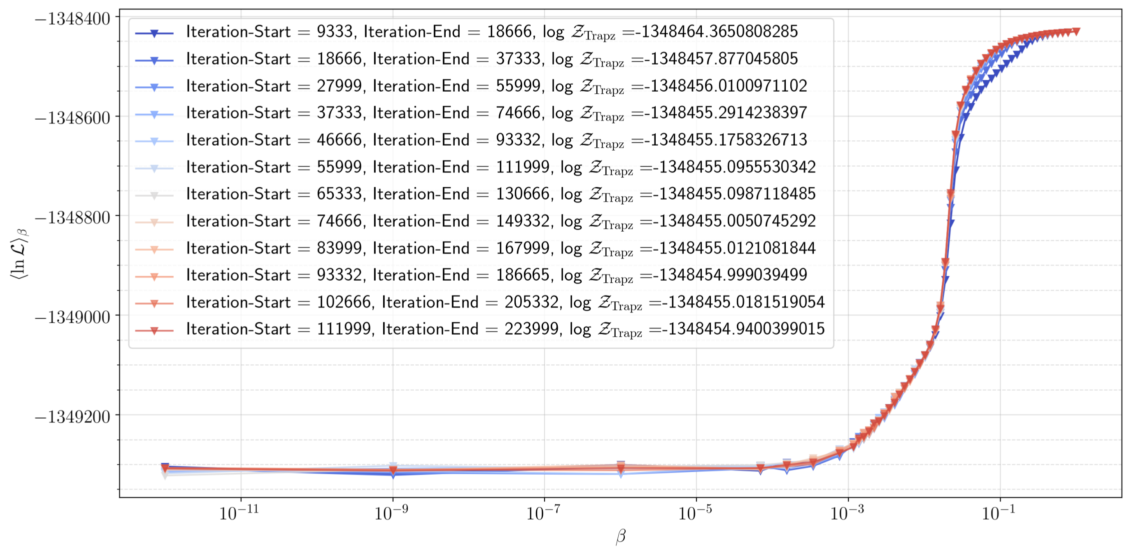


Figure 23: The convergence of the thermodynamic integrand for a gravitational wave analysis using 51 temperatures. This analysis neglected $\beta = 0$, but is otherwise an acceptable representation of the thermodynamic integrand. The Iteration-Start denotes the point is taken from a segment beginning with that MCMC iteration and ending with the MCMC iteration denoted as Iteration-End. These iterations correspond to the segments found in Fig. 22. The logarithm of the evidence is shown also in the figure caption, and as the MCMC analysis progresses the integral converges to a set value. The thermodynamic integrand can be visually seen to converge to a S-like curve but the shape and curvature are unique to hypotheses and choice of data. Early in the MCMC analysis the thermodynamic integrand can be mishaped as the power-posteriors have not all converged. Experience has told us that the power-posteriors that take the longest to converge tend to be in the region where the average log likelihood changes rapidly. Here this is in the region between $\beta \in (10^{-2}, 1)$.

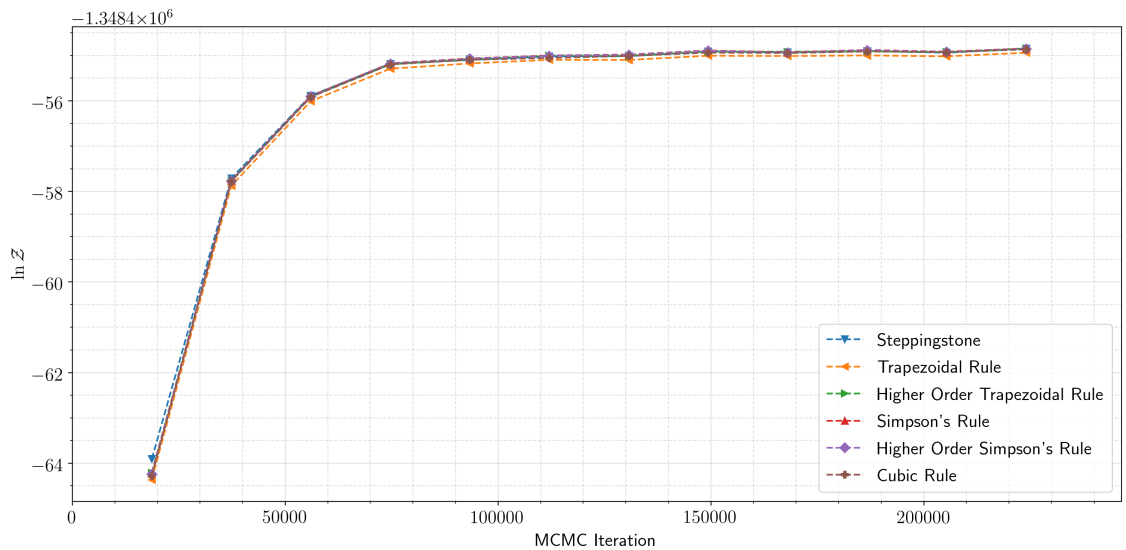


Figure 24: The convergence of the thermodynamic integral for a gravitational wave analysis using 51 temperatures as a function of the MCMC iteration. These choice of points of iterations correspond to the segments found in Fig. 22. As the analysis progresses the logarithm of the evidence from all quadrature methods tend towards a fixed value.

Bayes factor $\mathcal{B}^{\mathrm{H}_A}_{\mathrm{H}_B}$	log Bayes factor $\ln \mathcal{B}^{\mathrm{H}_A}_{\mathrm{H}_B}$	Posterior Probability $\mathcal{P}\left(\mathrm{H}_A \,\vert\, \mathbf{d}\right)$	z-score σ	Interpretation -
2.7×10^{-7} to 2.7×10^{-5}	-15.1 to -10.5	2.7×10^{-7} to 2.7×10^{-5}	-5 to -4	Very strong evidence for H_B
2.7×10^{-5} to 0.0014	-10.5 to -6.6	2.7×10^{-5} to 0.0014	-4 to -3	Very strong evidence for H_B
0.001 to 0.022	-6.6 to -3.8	0.0014 to 0.02	-3 to -2	Strong evidence for H_B
0.022 to 0.18	-3.8 to -1.7	0.02 to 0.15	-2 to -1	Moderate evidence for H_B
0.18 to 1.0	-1.7 to 0	0.15 to 0.5	-1 to 0	Weak evidence for H_B
1.0 to 5.5	0 to 1.7	0.5 to 0.85	0 to 1	Weak evidence for H_A
5.5 to 45	1.7 to 3.8	0.85 to 0.98	1 to 2	Moderate evidence for H_A
45 to 740	3.8 to 6.6	0.98 to 0.9986	2 to 3	Strong evidence for H_A
740 to 3.6×10^{4}	6.6 to 10.5	0.9986 to 0.999972	3 to 4	Very strong evidence for H_A
3.6×10^{4} to 3.6×10^{6}	10.5 to 15.1	0.999972 to 0.99999972	4 to 5	Very strong evidence for H_A

Table 6: An empirical scale for evaluating the relative strength of evidence between hypothesis H_A and an alternative hypothesis H_B loosely adapted from Tables 1 & 2 of [3]. We assume a prior odds ratio of unity between the two hypotheses, indicating no prior preference for either hypothesis. This makes the Bayes factor identical to the odds ratio and so we neglect the odds ratio here. Note that we have rounded to 2 significant digits leaving a minor rounding discrepancy with Table 1. The interpretation column is only a rule of thumb and interpretations vary across scientific fields and experimental contexts.

Derivative	Result	Moment Representation	Cumulant
$\frac{d(\ln \mathcal{Z})}{d\beta}$	$\frac{\mathcal{Z}'}{\mathcal{Z}}$	$\langle \ln \mathcal{L} \rangle_\beta$	μ_1
$\frac{d^2(\ln \mathcal{Z})}{d\beta^2}$	$\frac{\mathcal{Z}''}{\mathcal{Z}} - \frac{\mathcal{Z}'^2}{\mathcal{Z}^2}$	$\langle (\ln \mathcal{L})^2 \rangle_\beta - \langle \ln \mathcal{L} \rangle_\beta^2$	μ_2 - μ_1^2
$\frac{d^3(\ln \mathcal{Z})}{d\beta^3}$	$\frac{\mathcal{Z}^{(3)}}{\mathcal{Z})} + 2\frac{\mathcal{Z}'^3}{\mathcal{Z}^3} - 3\frac{\mathcal{Z}'\mathcal{Z}''}{\mathcal{Z}^2}$	$\langle (\ln \mathcal{L})^3 \rangle_\beta + 2\langle \ln \mathcal{L} \rangle^3 - 3\langle (\ln \mathcal{L})^2 \rangle_\beta \langle \ln \mathcal{L} \rangle_\beta$	$\mu_3 + 2\mu_1^3 - 3\mu_2\mu_1$
$\frac{d^4(\ln \mathcal{Z})}{d\beta^4}$	$\frac{\mathcal{Z}^{(4)}}{\mathcal{Z}} - 3\frac{\mathcal{Z}''^2}{\mathcal{Z}^2} - 6\frac{\mathcal{Z}'^4}{\mathcal{Z}^4} - 4\frac{\mathcal{Z}^{(3)}\mathcal{Z}'}{\mathcal{Z}^2} + 12\frac{\mathcal{Z}'^2\mathcal{Z}''}{\mathcal{Z}^3}$	$\langle (\ln \mathcal{L})^4 \rangle_\beta - 3\langle (\ln \mathcal{L})^2 \rangle_\beta^2 - 6\langle \ln \mathcal{L} \rangle_\beta^4$ - 4 $\langle (\ln \mathcal{L})^3 \rangle_\beta \langle \ln \mathcal{L} \rangle_\beta + 12\langle (\ln \mathcal{L})^2 \rangle_\beta \langle \ln \mathcal{L} \rangle_\beta^2$	$\mu_4 - 3\mu_2^2 - 6\mu_1^4$ $-4\mu_3\mu_1 + 12\mu_2\mu_1^2$

Table 7: The table of derivatives of the thermodynamic integral along with the moment representation of the derivatives of the thermodynamic integral and a set of reference of cumulants [4]. The cumulants are represented in terms of the non-central moments of a distribution so that the representation is made clear with the moment representation of the derivatives of the thermodynamic integral. The non-central moment μ_i represents the expectation value $\langle x^i \rangle$, where x is the one dimensional independent variable of a probability distribution function [5]. We only present four derivatives of the thermodynamic integral for both brevity and because our data are typically not accurate enough to estimate higher order cumulants. We have inspected that this relation holds up to the 7th derivative. Further orders could also be checked, although a formal proof may be more satisfactory.

Chapter 6

Searching for a Measurable Pressure-Gravity Mode Instability in GW170817

6.1 Introduction

The discovery of the binary neutron star merger GW170817 [54] has given us a new way to explore the physics of neutron stars. Recent studies have measured the star's tidal deformability and placed constraints on the equation of state of the neutron stars [54, 55, 56, 57, 19, 58, 59, 60, 61, 62]. [63] have suggested that the star's tidal deformation can induce nonresonant and nonlinear daughter wave excitations in p- and g-modes of the neutron stars via a quasi-static instability. This instability would remove energy from a binary system and possibly affect the phase evolution of the gravitational waves radiated during the inspiral. Although [327] concluded that there is no quasi-static instability and hence no effect on the inspiral, [64] claims that the instability can rapidly drive modes to significant energies well before the binary merges. However, the details of the instability saturation are unknown and so the size of the effect of the p-g mode coupling on the gravitational-waveform is not known [64]. The discovery of the binary neutron star merger GW170817 by Advanced LIGO and Virgo provides an opportunity to determine if there is evidence for nonlinear tides from p-g mode coupling during the binary inspiral.

Since the physics of the p-g mode instability is uncertain, [328] developed a parameterized model of the energy loss due to nonlinear tides. This model is parameterized by the amplitude and frequency dependence of the energy loss, and the gravitational-wave frequency at which the instability saturates and the energy loss turns on. For plausible assumptions about the saturation, [328] concluded that $> 70\%$ of binary merger signals could be missed if only point-particle waveforms are used, and that neglecting nonlinear tidal dynamics may significantly bias the measured parameters of the binary. Bayesian inference can be used to place constraints on nonlinear tides during the inspiral of GW170817. An analysis by [18] computed Bayes factors that investigate whether the GW170817 signal is more likely to have been generated by a model which includes nonlinear tides or one which does not. [18] find a Bayes factor of order unity, and conclude that the GW1701817 signal is consistent with both a model that neglects nonlinear tides and with a model that includes energy loss from a broad range of p-g mode parameters. However, the prior space used in this analysis includes a large region of parameter space where the amplitude of the effect produces a gravitational-wave phase shift that is extremely small. In this case, a waveform that includes p-g mode parameters will have a likelihood that is identical to the likelihood of the waveform without the p-g mode instability. The p-g mode model extends the standard waveform model by adding additional parameters that describe the nonlinear tidal effects. However, when including new parameters in a hypothesis if the likelihood does not vary across large portions of the prior volume for these new parameters relative to the likelihood of the original model, then the Bayes factor will not penalize this additional prior volume, nor will it penalize any extraneous parameters in the model (see e.g. [278, 275]). We examine prior space of p-g model used by [18] and find that although the p-g model model contains regions that are not consistent with the standard model, there are large regions of the prior space where the likelihood is high because the p-g mode model is degenerate with the standard model. These regions of prior space dominate the evidence and hence the Bayes factor neither favors nor disfavors the inclusion of p-g mode parameters.

We investigate a variety of different prior distributions on the p-g mode parameters beginning with a prior distribution that is similar to that tested in [18] and includes large regions of the parameter space that produce a negligible gravitational-wave phase shift. When comparing the evidence for this model with the standard waveform model used by [19] we find a Bayes factor of order unity, as expected. We then investigate a prior distribution in which the p-g mode instability parameters are constrained to induce a phase shift to the waveform that is greater than 0.1 radians. This phase shift is calculated from the time the waveform enters the sensitive band of the detector to the time when the waveform reaches the innermost stable circular orbit. We choose this threshold to exclude trivial regions of the parameter space that produce a non-measurable effect. However, we again find a Bayes factor of order unity when compared to the model hypothesis that does not model the p-g mode instability. Investigation of these results showed that this is due to parameter degeneracies between the p-g mode model and the intrinsic parameters of the standard waveform model.

Finally, we reduce the prior space to contain only the regions where the p-g mode waveform is not degenerate with the standard model by computing the fitting factor [329] of p-g signals against a set of standard waveforms. We do this to restrict the region of parameter space to that where the p-g effect is *measurably* distinct from a model that neglects nonlinear tides. We calculate the Bayes factor as a function of the fitting factor. We find that as the p-g mode parameter space is restricted to exclude regions that have a high fitting factor with standard waveforms, the Bayes factor decreases significantly. Regions of the nonlinear tide parameter space that have a fitting factor of less than 99% (98.5%) are strongly disfavored by a Bayes factor of 15 (25). While certain prior distributions of p-g mode parameters are consistent with the data, we find that these distributions are ones that contain large regions of non-measurable parameter space either because the effect produced is too small to measure, or the effect is degenerate with other parameters of the standard model. We conclude that the consistency of the GW170817 signal with the model of [328] is due to degeneracies and that regions where non-linear tides produce a measurable effect are strongly disfavored.

6.2 A Waveform Model for Nonlinear Tides from a p-g mode instability

As two neutron stars orbit each other, they lose orbital energy E_{orbital} due to gravitational radiation $\dot{E}_{GW}$. The gravitational waveform during the inspiral is well modeled by post-Newtonian theory (see e.g. [210]). The effect of the p-g mode instability is to dissipate orbital energy by removing energy from the tidal bulge of the stars [63, 64, 328]. Once unstable, the coupled p- and g-modes are continuously driven by the tides, giving rise to an extra energy dissipation $\dot{E}_{NL}$ for each star in the standard energy-balance equation [35]

$$\dot{E}_{\mathrm{orbital}} = -\dot{E}_{\mathrm{GW}} - \dot{E}^{1}_{\mathrm{NL}} - \dot{E}^{2}_{\mathrm{NL}}. \tag{6.1}$$

Since the details of how the nonlinear tides extract energy from the orbit is not known, [328] constructed a simple model of the energy loss and calculated plausible values for the model's parameters. In this model, the rate of orbital energy lost during the inspiral is modified by

$$\dot{E}_{\mathrm{NL}} \propto A f^{n+2} \Theta(f - f_0), \tag{6.2}$$

where A is a dimensionless constant that determines the overall amplitude of the energy loss, n determines the frequency dependence of the energy loss, and f_0 is the frequency at which the p-g mode instability saturation occurs and the effect turns on. By solving Eq. (6.1), [328] computed the leading order effect of the nonlinear tides on the gravitational-wave phase as a function of A, n, and f_0. In this analysis, they allowed each star to have independent values of A, f_0, and n, but found that the energy loss due to nonlinear tides depends relatively weakly on the binary's mass ratio. Hence, they consider a model that performs a Taylor expansion in the binary's component mass [330] and include only the leading order terms in the binary's phase evolution. Given this, we parameterize our nonlinear tide waveform with a single set of parameters A, n, and f_0, by setting $\dot{E}^{1}_{\mathrm{NL}} = \dot{E}^{2}_{\mathrm{NL}}$. We keep only the leading order nonlinear tide terms when we obtain the quantities $t(f)$ and $\phi(f)$ used to compute the stationary phase approximation [27, 102, 331]. This approach is reasonable for GW170817, since both neutron stars have similar masses and radii [19].

The dependence of A, n, and f_0 on the star's physical parameters is not known [64]. [328] estimate that plausible parameter ranges are $A \lesssim 10^{-6}$, $0 \lesssim n \lesssim 2$, and $30 \lesssim$

$f_0 \lesssim 80$ Hz. [332] found that the frequency at which the instability begins to grow is equation-of-state dependent and can occur at gravitational-wave frequencies as high as 700 Hz. [333] suggest that the instability may only act during the late stages of inspiral, (above 300 Hz), otherwise the large energy dissipation will cause the temperature of the neutron stars to be very large.

In this chapter, we compare two models for the gravitational waves radiated by GW170817. The first is the standard restricted stationary-phase approximation to the Fourier transform of the gravitational waveform $\tilde{h}(f)$, known as the TaylorF2 waveform [27]. We begin with the same waveform model used by [19], which is accurate to 3.5 PN order in the orbital phase, 2.0 PN order in spin-spin, self-spin and quadrupole-monopole interactions, 3.5 PN order in spin-orbit coupling, and includes the leading and next-to-leading order corrections from the star's tidal deformability [84, 83, 103, 216, 208, 334, 209, 211, 335, 336, 337]. We then construct a second model that adds the leading order effect of nonlinear tides computed using the model of [328]. Below we detail the construction of this second model, by computing the leading order nonlinear tidal Fourier phase term for the TaylorF2 model as well as the leading order nonlinear tidal energy dissipation.

We begin our derivation with the energy balance equation presented in [328],

$$\dot{E}_{orbit} = -\dot{E}_{GW} - 2\dot{E}_{NL}, \tag{6.3}$$

for $\dot{E}_{orbit}$ being the rate of energy loss of a quasi-circular orbit, $\dot{E}_{GW}$ being the energy rate loss due to gravitational waves in the point-particle model, and $\dot{E}_{NL}$ being the rate of energy loss from each star's p-g mode instability. We assume that the energy losses from p-g mode instability will be comparable in each star. The $\dot{E}$ notation refers to the derivative of the energy with respect to time. We now give explicit values to these energy rates with respect to gravitational wave frequency, f.

$$\dot{E}_{\mathrm{orbit}} = -\frac{G^{2/3}\pi^{2/3}\mathcal{M}^{5/3}\dot{f}}{3f^{1/3}} \tag{6.4}$$

is the orbital energy decay. The gravitational wave energy rate as a function of frequency is

$$\dot{E}_{\mathrm{GW}} = \frac{32G^{7/3}\left(\pi\mathcal{M}f\right)^{10/3}}{5c^5}. \tag{6.5}$$

Finally, we take from [328] that each star, indexed by i, should have an energy dissipation rate of

$$\dot{E}_{\mathrm{NL},i} = \frac{(2Gm_i)^{2/3} m_1 m_2}{M} (\pi f_{\mathrm{ref}})^{5/3} A \left(\frac{f}{f_{\mathrm{ref}}}\right)^{n+2} \Theta(f - f_0) \tag{6.6}$$

where m_i is the component mass of the neutron star, M is the total mass ($M = m_1 + m_2$). Assuming that the binaries have equal mass in Eq.. (6.6) and solving for $\dot{f}$, we arrive at the following expression

$$\frac{df}{dt} = \pi \left(\frac{f}{f_{\mathrm{ref}}}\right)^{7/3} f_{\mathrm{ref}}^2 \times \left[\frac{96}{5}\left(\frac{G\mathcal{M}\pi f_{\mathrm{ref}}}{c^3}\right)^{5/3}\left(\frac{f}{f_{\mathrm{ref}}}\right)^{4/3} + 6A\left(\frac{f}{f_{\mathrm{ref}}}\right)^{n} \Theta(f - f_0)\right]. \tag{6.7}$$

Given Eq. 6.7, we can now consider a time domain signal of the form, $h(t) = A(t)$ $e^{\phi(t)}$, where $h(t)$ is the strain of the gravitational wave at some time t before merger, $A(t)$ is the amplitude of the gravitational wave strain at that same time, and $\phi(t)$ is the orbital phase of the binaries[338]. This stationary phase approximation lets us approximate the Fourier transform of this time domain signal as

$$\tilde{h}(f) = \int_{-\infty}^{\infty} h(t)dt = \int_{-\infty}^{\infty} A(t)e^{-2\pi i f t + \phi(t)} df \approx \tilde{B}(f)e^{-i\Psi(f)} \tag{6.8}$$

where $\tilde{B}(f)$ is the Fourier amplitude of the frequency domain waveform, and $\Psi(f)$ is the Fourier phase of the frequency domain waveform. We express this Fourier phase as

$$\Psi(f) = 2\pi f t(f) - \phi(f).. \tag{6.9}$$

One can derive $t(f)$ by solving the differential equation given in Eq.. (6.7). For convenience we redefine and reorganize this differential equation as:

$$\int_{t_c}^{t} dt = \int_{x_c}^{x} \frac{f_{\mathrm{ref}}}{\kappa} \frac{x^{-7/3}dx}{\alpha x^{4/3} + \Theta(x - x_0)\beta x^n} \tag{6.10}$$

where $x = f/f_{\mathrm{ref}}$, $dx = df/f_{\mathrm{ref}}$, $x_0 = f_0/f_{\mathrm{ref}}$, and $\kappa = \pi f_{\mathrm{ref}}^2$. The integration bounds are the time of coalescence ($t_c = 0$) to some time t prior to merger, and from dimensionless frequency at coalescence ($x_c = f_c/f_{\mathrm{ref}} = \infty$) to dimensionless frequency x prior to merger. Here α and β are given by the following expressions:

$$\alpha = \frac{96}{5}\left(\frac{G\pi\mathcal{M}f_{\mathrm{ref}}}{c^3}\right)^{5/3} \tag{6.11}$$

$$\beta = 6A \tag{6.12}$$

We can simplify the differential equation given in Eqn. (6.10) if we assume that the point particle gravitational wave contribution dominates ($\alpha \gg \beta$), we take a power series expansion assuming large α relative to β. This gives to lowest order in β:

$$\int_{t_c}^{t} dt = \int_{x_c}^{x} \frac{f_{\rm ref}}{\kappa} \left(\frac{1}{\alpha x^{11/3}} + \frac{\Theta(x-x_0)\beta x^{n-5}}{\alpha^2 (n-5)} \right) dx \tag{6.13}$$

The first term in Eq.. (6.13) corresponds to the lowest order post-Newtonian result from the point-particle model. Integrating the second term and respecting the $\Theta(x - x_0)$ so as to align the waveform at merger ($t = 0$), we arrive at the leading order contribution of p-g mode instability to $t(f)$:

$$\Delta t(f) = \begin{cases} \frac{-25}{1536} \frac{1}{\pi} \frac{A}{n-4} \left(\frac{G\mathcal{M}\pi f_{\rm ref}}{c^3}\right)^{-10/3} \left(\frac{f_0}{f_{\rm ref}}\right)^{n-4}, & f < f_0 \\ \\ \frac{-25}{1536} \frac{1}{\pi} \frac{A}{n-4} \left(\frac{G\mathcal{M}\pi f_{\rm ref}}{c^3}\right)^{-10/3} \left(\frac{f}{f_{\rm ref}}\right)^{n-4}, & f \geq f_0 \end{cases} \tag{6.14}$$

Following a similar approach we can calculate $\phi(f)$ via $d\phi = 2\pi f dt = 2\pi(f/\dot{f})df$. Taking the same power series expansion, integrating so that the waveform coalesces at $t = 0$, and examining the leading order contribution from p-g mode instability we arrive at

$$\int_{\phi_c}^{\phi} d\phi = \int_{x_c}^{x} \frac{f_{\rm ref}}{\kappa} \left(\frac{1}{\alpha x^{8/3}} + \frac{\Theta(x-x_0)\beta x^{n-4}}{\alpha^2 (n-4)} \right) dx. \tag{6.15}$$

Integrating this through from ϕ_c, the phase at coalescence, to some earlier ϕ prior to coalescence, and integrating the right hand side of Eq.. (6.15) we get the lowest order post-Newtonian correction to the phase for the point particle model in integrating the $x^{-8/3}$ term and the lowest order correction due to p-g mode instability in integrating the x^{n-4} term. Thus the correction to the gravitational wave phase due to p-g mode instability is

$$\Delta \phi(f) = \begin{cases} \frac{-25}{768} \frac{A}{n-3} \left(\frac{G\mathcal{M}\pi f_{\rm ref}}{c^3}\right)^{-10/3} \left(\frac{f_0}{f_{\rm ref}}\right)^{n-3}, & f < f_0 \\ \\ \frac{-25}{768} \frac{A}{n-3} \left(\frac{G\mathcal{M}\pi f_{\rm ref}}{c^3}\right)^{-10/3} \left(\frac{f}{f_{\rm ref}}\right)^{n-3}, & f \geq f_0 \end{cases} \tag{6.16}$$

Finally, we can express the Fourier phase in terms of Eq.s. (6.14) and (6.16) as:

$$\Delta\Psi(f) = \begin{cases} 2\pi f \Delta t(f_0) - \Delta\phi(f_0), & f < f_0 \\ \\ 2\pi f \Delta t(f) - \Delta\phi(f), & f \geq f_0 \end{cases} \tag{6.17}$$

which fully expanded becomes:

$$\Delta\Psi(f) = \begin{cases} -\frac{25}{768} A \left(\frac{f_0}{f_{\rm ref}}\right)^{n-3} \left[\frac{f}{f_0}\frac{1}{n-4} - \frac{1}{n-3}\right], & f < f_0 \\ \\ -\frac{25}{768} A \left(\frac{G\mathcal{M}\pi f_{\rm ref}}{c^3}\right)^{-10/3} \left(\frac{f}{f_{\rm ref}}\right)^{n-3} \left[\frac{1}{n-4} - \frac{1}{n-3}\right], & f \geq f_0 \end{cases} \tag{6.18}$$

Here, $f_{\rm ref}$ is a reference frequency which we set to 100 Hz following [328], G is Newton's gravitational constant, c is the speed of light, and $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$ is the chirp mass of the binary.[1] This waveform model can have a degeneracy in the gravitational wave phasing with chirp mass when $n = 4/3$. For this value of n, the Fourier phase in Eq. (6.18) for nonlinear tides is $\Psi(f) \propto f^{-5/3}$, which is the same power law dependence as the chirp mass phasing. A degeneracy occurs when f_0 is comparable or lower than the frequency at which chirp mass can be accurately measured. In this case, the p-g mode instability is degenerate with changing the chirp mass. In principle, there will be other degeneracies with other intrinsic parameters of the gravitational wave signal for other values of n.

We generate the standard TaylorF2 waveform using the LIGO Algorithm Library [325] and multiply this frequency-domain waveform by the term due to the nonlinear tides,

$$\tilde{h}_{\rm TaylorF2+NL}(f) = \tilde{h}_{\rm TaylorF2}(f) \times \exp[-i\Psi_{\rm NL}(f)]. \tag{6.19}$$

The Fourier phase for the nonlinear tides is implemented as a patch to the version of the PyCBC software [339] used by [19]. Both the standard and nonlinear tide waveform models are terminated when the gravitational-wave frequency reaches that

[1] Appendix A of [328] gives the change to the gravitational-wave phase $\phi(f)$ as a function of frequency and not the change to the Fourier phase $\Psi(f)$ (see e.g. [331] for a discussion of how these differ). The former quantity is useful to compute the change in the number of gravitational-wave cycles, but the latter is required to compute the modification to the TaylorF2 waveform. The study by [18] corrects this mistake.

of a test particle at the innermost stable circular orbit of a Schwarszchild black hole of mass $M = m_1 + m_2$. For the neutron star masses considered here, this frequency is between 1.4 kHz and 1.6 kHz.

We also derive the first-order energy dissipation from p-g modes in the frequency-domain. This can be solved as,

$$E_{\mathrm{NL},i}\left(f'\right) = \int_0^{f'} \left(\frac{dE_{\mathrm{NL},i}}{dt}\right)\left(\frac{dt}{df}\right) df. \tag{6.20}$$

In this derivation we only keep the leading order in A, so we take $\frac{dt}{df}$ from the point-particle term in the approximation and neglect terms in A^2. The derivation in Eq. (6.7) made use of the simplification that $m_1 = m_2$, but we do not take this approach here. The point-particle form of $\frac{dt}{df}$ is [338]:

$$\frac{dt}{df} = \frac{5}{96}\frac{c^5}{G^{5/3}\pi^{8/3}\mathcal{M}^{5/3}f^{11/3}}. \tag{6.21}$$

Placing this equation into Eq. (6.6) and then placing Eq. (6.21) into Eq. (6.20) gives:

$$\frac{dE_{\mathrm{NL},i}}{df} = \frac{5}{96}\frac{\left(2m_i\right)^{2/3}\, m_1\, m_2\, A\, c^5}{G(m_1+m_2)\,\pi\,\mathcal{M}^{5/3}} f_{\mathrm{ref}}^{-n-1/3} f^{(n-5/3)}\Theta(f-f_0) \tag{6.22}$$

Integrating this Eq. (6.22) over all frequencies gives us the energy dissipated by the p-g mode instability for a neutron star of mass m_i:

$$E_{\mathrm{NL},i}(f) = \frac{5}{96}\frac{\left(2m_i\right)^{2/3}\, m_1\, m_2\, A\, c^5}{G\,\pi\,(m_1+m_2)\mathcal{M}^{5/3}} f_{\mathrm{ref}}^{-n-1/3}\left(f^{n-2/3} - f_0^{n-2/3}\right)\frac{1}{n-2/3}. \tag{6.23}$$

Dimensional analysis confirms that Eq. (6.23) is in the form of Joules. In our case however, we are only concerned with the energy dissipated by the p-g mode instability at f_{ISCO} when the stars have finally merged. For neutron stars f_{ISCO} is always greater than f_0, and so the energy dissipation, summing over the contributions from both stars, is:

$$E_{\mathrm{NL}}(f_{\mathrm{ISCO}}) = \frac{5}{96}\frac{\left(2m_1+2m_2\right)^{2/3}\, m_1\, m_2\, A\, c^5}{G\,\pi\,\mathcal{M}^{5/3}} f_{\mathrm{ref}}^{-n-1/3}\left(f_{\mathrm{ISCO}}^{n-2/3} - f_0^{n-2/3}\right)\frac{1}{n-2/3}. \tag{6.24}$$

In the next section we move on to describing appropriate prior distribution for the p-g mode instability parameters as well as the other intrinsic parameters for GW170817.

6.3 Bayesian Model Priors

Bayes' theorem offers a methodology for evaluating the plausibility of models relative to a given data set, and then updating these prior model beliefs with better hypotheses. Bayes' theorem states, in the notation of Chapter 5, that

$$\mathcal{P}\left(\vec{\theta}|H,\mathbf{d}\right)=\frac{\pi\left(\vec{\theta}|H\right)\,\mathcal{L}\left(\mathbf{d}|H,\vec{\theta}\right)}{\mathcal{Z}\left(\mathbf{d}|H\right)}, \tag{6.25}$$

where $\mathcal{Z}\left(\mathbf{d}|H\right)$ is the evidence of the model H, $\pi\left(\vec{\theta}|H\right)$ is the is the prior distribution of the parameters given the signal model, $\mathcal{L}\left(\mathbf{d}|H,\vec{\theta}\right)$ is the likelihood of the data for a particular set of parameters $\vec{\theta}$, and $\mathcal{P}\left(\vec{\theta}|H,\mathbf{d}\right)$ is the posterior distribution of the parameters given the signal model. The likelihood used in this analysis assumes a Gaussian model of detector noise and depends upon the noise-weighted inner product between the gravitational waveform and the data from the gravitational-wave detectors [340, 341]. The choice of prior distributions on the parameters of the signal model represent the hypothesis that we want to test. The posterior distributions reflect how to update ones beliefs with respect to the likelihood and the data. Thus, by examining many different parameter hypotheses we can investigate the extent to which GW170817 is accurately modeled by p-g mode instability waveform models.

In our analysis, we fix the sky location and distance to GW170817 [342, 343] and assume that both neutron stars have the same equation of state by imposing the common radius constraint [19]. In the case of the standard TaylorF2 waveform H_{TaylorF2}, our analysis is identical to that described in [19]. This analysis considered three prior distributions on the binary's component mass. Here, we only consider the uniform prior on each star's mass, with $m_{1,2}\sim U[1,2]\,M_{\odot}$, and the Gaussian prior on the component masses $m_{1,2}\sim N(\mu=1.33,\sigma=0.09)\,M_{\odot}$ [148]. For both mass priors, we restrict the chirp mass to the range $1.1876M_{\odot}<\mathcal{M}<1.2076M_{\odot}$. Since our analysis is identical to that of [19], we refer to that paper for the details of the data analysis configuration.

Given the uncertainty on the range of the nonlinear tide parameters, we follow [18] and let $n\in U[-1.1,2.999]$, draw A from a distribution uniform in $\log_{10}$ between 10^{-10} and $10^{-5.5}$, and $f_0\in U[10,100]$ Hz. We use this along with a uniform prior distribution on the mass from [19].

We also consider two alternative choices of drawing f_0: we draw f_0 from a uniform distribution between 15 and 100 Hz, as used by [328], and from a uniform distribution between 15 and 800 Hz to allow for the larger values of f_0 suggested by [332] and [333]. For these choices we consider A uniform in $\log_{10}$ between 10^{-10} to 10^{-6}. For these alternative prior distributions we also consider applying a further constraint on the parameters. Since some combinations of A, n, and f_0 can produce extremely small gravitational-wave phase shifts [328], we place a cut on the gravitational-wave phase shift due to nonlinear tides

$$\delta\phi(f_{\rm ISCO}) = \frac{-25}{768}\frac{A}{n-3}\left(\frac{G\mathcal{M}\pi f_{\rm ref}}{c^3}\right)^{-10/3}\left[\left(\frac{f_0}{f_{\rm ref}}\right)^{n-3} - \left(\frac{f_{\rm ISCO}}{f_{\rm ref}}\right)^{n-3}\right], \qquad (6.26)$$

where $f_{\rm ISCO}$ is the termination frequency of the waveform (which is always larger than f_0 in our analysis). This gravitational-wave phase shift from the p-g mode instability is strictly negative, but we take the convention of using the absolute value of the phase shift for convenience. We restrict the prior space to values of $\delta\phi > 0.1$ rad. Phase shifts of $\delta\phi \approx 0.1$ rad have an overlap between the two waveform models greater than 99.98%. This cut means that the resulting priors on A, n, and f_0 are not uniform, but are biased in favor of combinations of parameters that may produce a measurable effect on the phasing of the waveform due to nonlinear tides. While $\delta\phi$ is a simple proxy for how similar or dissimilar two waveforms are, formally this is given by the match between two waveforms. A $\delta\phi$ of 1 radian may have a low overlap with a waveform if the radian is accumulated over a large bandwidth but a high overlap if the radian is accumulated near the very end of the signal. Fig. 25 shows a depiction of the prior distributions used when using a permissive prior on $\delta\phi$, similar to [18], and when using a constraint on the p-g mode parameters such that $\delta\phi > 0.1$ rad.

A stricter approach to constructing a prior distribution that considers p-g mode effects that are distinguishable from standard waveforms is to examine the fitting factor between a distribution of p-g mode waveforms and a set of comparable TaylorF2 waveforms. To do so, we examine the fitting factor of our Bayesian inference analysis with respect to a template bank of non-spinning, mass-only TaylorF2 waveforms. We construct a template bank of $\sim 20,000$ non-spinning, mass-only waveforms of comparable masses to the prior distribution on the mass parameters. The template bank is constructed with component masses, $m_{(1,2)} \in (1.0, 2.0)M_{\odot}$, chirp masses, $\mathcal{M}_c \in (1.1826, 1.2126)M_{\odot}$, and a minimal match placement of 99.9%. We then place a threshold on the evidence calculation from the Bayesian analysis based on the maximum overlap with this template bank of standard waveforms. This permits an analysis of the Bayes factor for nonlinear tides where the prior distribution on p-g mode parameters is determined by the fitting factor with a set of standard signals.

6.4 Bayesian Parameter Estimation and Hypothesis Testing Methods

We use the gravitational-wave strain data from the Advanced LIGO and Virgo detectors for the GW170817 event, made available through the GW Open Science Center [48, 344]. We then repeat the analysis of [19] using the waveform model $H_{\mathrm{TaylorF2+NL}}$ to compute the evidence $p(\mathbf{d}\,|\,\mathrm{H}_{\mathrm{TaylorF2+NL}})$.

We use Bayesian model selection to determine which of the two waveform models described in Sec. 6.2 is better supported by the observation of GW170817. Bayes' theorem in Eq. (6.25) permits us a method for model comparison through the ratio of the evidence from each model. This ratio of the model evidences is called the Bayes factor, which we denote as $\mathcal{B}$. A Bayes factor greater than unity indicates support for the model in the numerator, while a Bayes factor less than unity indicates support for the model in the denominator. The Bayes factor can be written as,

$$\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}} = \frac{\mathcal{Z}(\mathbf{d}\,|\,H_{\mathrm{NL}})}{\mathcal{Z}(\mathbf{d}\,|\,H_{!\mathrm{NL}})}. \tag{6.27}$$

The numerator of Eq. (6.4) is the evidence for nonlinear tides $\mathcal{Z}(\mathbf{d}\,|\,\mathrm{NL})$. For the denominator of Eq. (6.4), we use the evidence $\mathcal{Z}(\mathbf{d}\,|\,!\mathrm{NL})$ provided as supplemental materials by [19]. We have used NL (!NL) to denote the nonlinear tidal (standard models).

Posterior distributions for parameters of interest can be also computed by marginalizing the posterior probability distribution over other parameters. Marginalization to obtain the posterior probabilities and the evidence is performed using Markov Chain Monte Carlo (MCMC) techniques. To compute posterior probability distributions and evidences, we use the *PyCBC Inference* software [339, 303] using the parallel-tempered *emcee* sampler [283, 284]. This sampler allows the use of multiple temperatures to sample the parameter space [283, 285, 286]. These multiple temperatures β permit the construction of tempered posterior distributions that form a slow thermodynamic transition from the prior distribution to the posterior distribution in Eq. 6.25. Tempered posteriors are called power-posteriors in [49, 50]. The power-posterior can be according to:

$$\mathcal{P}(\vec{\theta} \,|\, \mathbf{d}, H)_\beta \propto \pi(\vec{\theta} \,|\, H)\mathcal{L}(\mathbf{d} \,|\, \vec{\theta}, H)^\beta. \tag{6.28}$$

The normalization constant for a power-posterior is the evidence for that power-posterior, given as $\mathcal{Z}(\mathbf{d} \,|\, H)_\beta = \int \pi \left(\vec{\theta} \,|\, H\right) \mathcal{L}(\mathbf{d} \,|\, \vec{\theta}, H)^\beta d\vec{\theta}$.

From these power-posterior distributions we use the thermodynamic integration method [49, 50] to estimate the logarithm of the evidence, $\ln \mathcal{Z}$, given as:

$$\ln \mathcal{Z} = \int_0^1 \langle \ln \mathcal{L} \rangle_\beta \, d\beta. \tag{6.29}$$

The estimate of the evidence is determined by the integral over inverse temperatures, β, of the average untempered log likelihood, $\langle \ln \mathcal{L} \rangle_\beta$, drawn from the power-posterior corresponding to the inverse temperature β. An approximation to this integral can be made through use of trapezoid rule integration method. Following [19] we use 51 temperatures where we use a combination of geometric and logarithmic temperature placements to improve the accuracy of the integral [292].

We verify the results of the thermodynamic integration evidence calculation by comparing it with the steppingstone algorithm [51], which utilizes the same likelihoods from multi-tempering sampling as the thermodynamic integration method. Both trapezoidal rule thermodynamic integration and steppingstone methods can have some bias in the estimate of the logarithm of the Bayesian evidence due to a finite number of temperatures being used. This bias is mitigated by an increased number of temperatures [51, 304]. Additionally, this bias can be mitigated in thermodynamic integration by improving the order of the quadrature integration [290]. We also use a higher order trapezoidal rule from [290] and verify that the results are consistent.

We also estimate the error for each method of evidence calculation. The thermodynamic integration method and steppingstone algorithm both contain Monte Carlo error [287]. For the thermodynamic integration method the Monte Carlo error on the thermodynamic integral can be estimated following the methodology of [287]. We use this same uncertainty estimate for the higher order trapezoidal rule as well. In [51] there is a Monte Carlo variance estimate for the logarithm of the evidence from the steppingstone method that we also use here.

The last source of error in the evidence calculation that we consider is whether the MCMC has converged to stable likelihood values across all of the temperatures. This requires examining the stability of the evidence calculations as the MCMC progresses. Independent samples are drawn according to the $\mathrm{n}_{\mathrm{acl}}$ method as described by [303] at various points in the run. This method takes a specific endpoint iteration, takes half the endpoint iteration as the starting point iteration, and calculates the autocorrelation length of the samples between the starting point and the endpoint iteration. Independent samples are drawn in intervals of the maximum autocorrelation length for the samples within this segment. We divide the full run into 12 segments and calculate the evidence from each one of these segments to examine how the evidence progresses along the MCMC iterations. Gradually the evidence begins to settle towards a constant value as the MCMC progresses. We take the difference between the last two evidence estimates as the convergence error.

We estimate the total error on our evidence calculations, $\sigma_{\ln \mathcal{Z}}$, by adding the errors in quadrature according to,

$$\sigma_{\ln \mathcal{Z}} = \sqrt{\sigma^2_{\mathrm{MC}} + \sigma^2_{\mathrm{convergence}}}\,. \tag{6.30}$$

Here, the error σ_{MC} is the Monte Carlo error and $\sigma_{\mathrm{convergence}}$ is the convergence error. Finally, to estimate the Bayes factors we model the log evidence as a normal distribution, with mean given from the log evidence calculation, and standard deviation given by the error propagation formula in Eq. (6.30). The logarithm of the Bayes factor can then be calculated from the difference in the logarithm of the evidences. The standard Bayes factor is then the exponential of the logarithm of the Bayes factor.

As a means of verifying the results from the above Bayes factor calculations we also make use of the Savage-Dickey density ratio method [52, 53, 310] for calculating the Bayes factor of the model where the p-g mode parameters were chosen independently of one another. This is the approach taken in [18].

For certain kinds nested models where prior distributions on parameters are *factorizable*, or independent from one another, there exists a method for deriving the Bayes factor for two models from one parameter estimation Markov-Chain Monte Carlo analysis. If there exists a parameter A for which at a critical value A_{crit} the parameter model is equivalent to a nested model that has no parameterization in A, then the Bayes factor for the model with A compared to the model without A

is taken as the limit of the prior density at A_{crit} relative to the posterior density at A_{crit} when sampled from the model that includes A. This method does not require a multi-dimensional integral or one-dimensional integral to be approximated. In the case of the p-g mode instability the parameter that effectively turns on and turns off the instability is the amplitude factor A. The Bayes factor form the Savage-Dickey density ratio is the ratio of the probability densities between the prior distribution density as $A \to 0$ and the posterior distribution density as $A \to 0$. This expression can be written as:

$$\mathcal{B}^{\text{NL}}_{!\text{NL}} = \lim_{\text{A}\to 0} \frac{\pi\left(\text{A} \,|\, H^{NL}\right)}{\mathcal{P}\left(\text{A} \,|\, \mathbf{d}, H^{NL}\right)}. \tag{6.31}$$

Formally, the parameter model is constructed such that the prior density on A is distributed uniformly in $\log_{10} A$ and so the limit cannot be strictly taken from within the data acquired in these analyses. However, when $Ais10^{-10}$, the matched-filter is not sensitive enough to to distinguish the difference between $A = 0$ and $A = 10^{-10}$ to $> 99.99\%$ overlap. This indicates that substituting $A \to 0$ for $A \to 10^{-10}$ will generate indistinguishable results in this analysis.

This changes the problem of inference from numerical integration to that of probability density estimation. In our analysis, only the unconstrained $\delta\phi$ model has a marginal prior distribution on A that is independent of all of the other parameter priors. This model is similar to [18], where A is uniform in $\log_{10}$ between 10^{-10} and $10^{-5.5}$. Our prior distributional density is analytic and we know the exact prior probability density at 10^{-10} is 0.22222. This reduces the probability density inference to the marginal posterior distribution density on A at 10^{-10}. There are a variety of methods for estimating the density of a probability distributions from samples of data that we introduced in Chapter 5. We consider the histogram method using Scott's binning method [312], the histogram method using the Freedman-Diaconis binning method [313], the Gaussian kernel density estimator GetDist from [20], and the logspline estimator from [314].

6.5 Results

Compared to the standard waveform model, we find that the p-g mode model with priors where $\delta\phi$ is unconstrained gives a Bayes factor of order unity. When we use p-g mode priors where $\delta\phi > 0.1$ radians we also find a Bayes factor of order unity.

Following the Bayes factor interpretation of [278, 345], these Bayes factors cannot be considered to be statistically significant. A Bayes factor of unity indicates that whatever prior beliefs we had about the plausibility of the p-g mode instability prior to the observation of GW170817 is unchanged by the observation of GW170817. For the narrow range of $15 \leq f_0 \leq 100$ Hz where $\delta\phi > 0.1$ rad, we find that the Bayes factor is ~ 0.7. This is also true of the prior range $10 \leq f_0 \leq 100$ Hz with unconstrained $\delta\phi$. The broader range $15 \leq f_0 \leq 800$ Hz, where $\delta\phi > 0.1$ rad, we find that $\mathcal{B} \sim 0.7$ as well. Our estimated statistical error on Bayes factors due to Monte Carlo uncertainty and convergence uncertainty is $\sim \pm 0.1$ at the 90% confidence level. In Section 6.5.1 we discuss the performance of the Bayes factor estimation from the multi-tempered Bayesian evidence estimators. In Section 6.5.2 we verify the results of the multi-tempered Bayes factor estimates with the Savage-Dickey density ratio test for the unconstrained $\delta\phi$ model.

6.5.1 Multi-Tempered Bayes Factors

Our Bayes factor estimation from 6 multi-tempered estimators on the logarithm of the Bayes factor can be seen in Fig. 27 when comparing the hypothesis on p-g mode instability for the unconstrained $\delta\phi$ prior to the hypothesis presented in [19] for the uniform mass prior with a common equation of state constraint. The 6 multi-tempered estimators are fully described in Chapter 5; they are the thermodynamic integration method with the trapezoidal rule, a first-order correction to the trapezoidal rule, Simpson's rule, a first-order correction to Simpson's rule, a cubic integration rule, and the steppingstone method. The different methods give similar probability distributions on the estimate of the Bayes factor. Those Bayes factor uncertainty distributions follow a log-normal distribution and have tails that skew towards a Bayes factor of unity.

The Bayes factors for all hypotheses using all of the quadrature methods in Chapter 5 can be seen in Table 8. The median values of the Bayes factors range between roughly 0.63 and 0.76, with the 5th and 95th percentile interval being $\sim \pm 0.1$. Under a binary choice between the p-g mode instability model and the standard model we can calculate a posterior probability of the nonlinear tidal hypothesis. Without giving preference to either model, we calculate a posterior probability of $\mathcal{P}(\mathrm{H}_{NL} \,|\, \mathbf{d})$ between 34 % and 46% at the 5% and 95% confidence levels. These posterior probabilities correspond to an odds of $0.5 : 1$ and $0.85 : 1$, which are not statistically significant and indicate that the data are uninformative to testing either hypothesis. If we consider all models collectively, the posterior probability on any one particular model reduces significantly due to the increase in number of hypotheses available to consider.

6.5.2 Savage-Dickey Density Ratio Bayes Factors

We report on the Bayes factors from the Savage-Dickey density ratio test on the model for p-g mode instability for an unconstrained $\delta\phi$ prior compared to a standard model. The Savage-Dickey density ratio requires us to know the probability density for the marginal prior and posterior on A at 10^{-10}.

Since the marginal prior and posterior distribution functions on A are distributed logarithmically, it is convenient to do the density estimation in the $\log_{10} A$. Under this change of variables the marginal prior distribution on $\log_{10} A$ is uniform between -10 and -5.5, hence the prior distribution function is:

$$\pi\left(\log_{10} A\right) = \frac{1}{-5.5-(-10)} = 0.22, \qquad -10.0 \leq \log_{10} A \leq -5.5 \tag{6.32}$$

Following this we estimate the marginal posterior probability density of $\log_{10} A$ using the Savage-Dickey density ratio methods in Chapter 5. We use the two histogram density estimator methods with Scott's binning rule and the Freedman-Diaconis binning rule for the posterior probability density estimation. We also use the Gaussian kernel density estimator GetDist, and the logspline density estimating package found in R for this posterior probability density estimation. A comparison of the density estimates for the marginal posterior probability density on $\log_{10} A$ for the different density estimators can be seen in Fig. 28. We then calculate the Savage-Dickey density ratio Bayes factor and use a bootstrap resampling method to resample the posterior distribution $5,000$ times to get a confidence interval on our Bayes factor estimates. The results can be seen in Fig. 29, and are summarized in Table 9.

6.5.3 Parameter Estimation Results

Bayes factor hypothesis testing only provides half of the Bayesian inference method [276]. In this section we examine the results of the parameter estimation if we assume that nonlinear tides are present in the GW170817 signal. Remarkably, when we consider the way that the nonlinear tides enter the Fourier phase in Eq. (6.18), we see that if $n = 4/3$ then the nonlinear tides enter the Fourier phase of the waveform with the same power law dependence on frequency f as the chirp mass, that is $\Psi(f) \propto f^{-5/3}$. We also note that for the effect of nonlinear tides to be degenerate with chirp mass, they must turn on at a frequency f_0 that is close to the low-frequency limit of the detector's sensitive band. If the effect turns on at higher frequencies, then the phasing will change in the detector's sensitive band and it is more difficult to compensate for the nonlinear tide effect with a change in chirp mass.

The marginalized posterior distributions on parameters shown in Fig. 33 show a strong degeneracy between the source-frame chirp mass $\mathcal{M}^{\rm src}$ and nonlinear tides that creates a tail in the chirp mass posterior skewed towards lower values of chirp mass than the value measured using the standard waveform model, $\mathcal{M}^{\rm src} = 1.1867 \pm 0.0001\, M_\odot$ [19]. We see a peak in the posteriors of n and f_0 at $n \lesssim 4/3$ and $f_0 \lesssim 35$ Hz. This parameter degeneracy is also correlated with large A, where $10^{-8} \lesssim A < 10^{-6}$. The samples with large posterior values of $\delta\phi$ seen in Fig. 33 are strongly correlated with source-frame chirp masses $\mathcal{M}^{\rm src} \lesssim 1.1866$. We have examined the change to the posterior distribution when changing the low-frequency cutoff of the likelihood integration from 20 Hz to 25 Hz, and to 30 Hz. In these analyses, the peak in the posterior of f_0 tracks the low-frequency cutoff of the likelihood integration, confirming that this effect is due to the chirp-mass degeneracy with the low-frequency cutoff. The chirp mass degeneracy is also present in the analysis with the broader range of f_0, however it is not as pronounced in the posterior samples due to the larger prior space being explored.

We also examine the leading order estimated energy dissipated through nonlinear tides for the case of a uniform prior on the mass, with $15 \leq f_0 \leq 100$ Hz, with a $\delta\phi > 0.1$ radian constraint. In our analysis, the 95th percentile of the estimated energy dissipated through nonlinear tides from our prior distribution is approximately 2.6×10^{51} ergs at the terminating frequency of the TaylorF2 waveform, $f_{\rm ISCO}$. The estimated energy radiated by gravitational waves by neutron stars of the estimated mass range of GW170817 is greater than $\sim 10^{53}$ ergs. Our analysis finds the energy dissipated through nonlinear tides at the 95% posterior credible percentile is 3×10^{50} ergs. We find our 95% posterior credible percentile to be less than the 90% confidence interval constraint of $\lesssim 2.7 \times 10^{51}$ ergs in [18]. Samples from our posterior distribution that have dissipation energies greater than the 90% credible interval tend to come from two modes in the parameter space. The first mode is from parts of the parameter space with large A, for $n \sim 4/3$, low f_0, and $\delta\phi \sim 100$ rad. The second mode is from parts of the parameter space with $A \gtrsim 10^{-8}$, for $1.6 \lesssim n < 3.0$, and $\delta\phi \sim 1 - 10$ rad. The high end of the nonlinear tidal energy constraints are thus dominated by waveforms that are degenerate with the standard signal.

6.5.4 Improving the Chirp Mass Degeneracy with an Independent Electromagnetic Observation

In this section we consider whether the chirp mass degeneracy could be mitigated by the measurement of the chirp mass by an independent electromagnetic observation. We find that we require a very strong constraint on the chirp mass independent of the gravitational wave data to mitigate the parameter degeneracy from the p-g mode instability. Here we make a quantitative analysis of how accurate an electromagnetic observer's measurement of the chirp mass would have to be to constrain the chirp mass back to the measurements found in [19].

To do so we consider the joint posterior distribution, $\mathcal{P}(\mathcal{M} \,|\, \mathbf{d_{GW}}, \mathbf{d_{EM}})$, from two statistically independent data sets, the gravitational wave data $\mathbf{d_{GW}}$, and a mock electromagnetic data set $\mathbf{d_{EM}}$. We then define a (hyper) prior on the chirp mass that we believe credible from the joint observation of GW170817 from gravitational wave detectors and a mock electromagnetic observer. The joint posterior distribution on the chirp mass is then

$$\mathcal{P}(\mathcal{M} \,|\, \mathbf{d_{GW}}, \mathbf{d_{EM}}) = \frac{\pi(\mathcal{M})}{\mathcal{Z}(\mathbf{d_{GW}}, \mathbf{d_{EM}})} \mathcal{L}(\mathbf{d_{GW}}, \mathbf{d_{EM}} \,|\, \mathcal{M}). \tag{6.33}$$

We can separate $\mathcal{L}(\mathbf{d_{GW}}, \mathbf{d_{EM}} \,|\, \mathcal{M})$ into $\mathcal{L}(\mathbf{d_{GW}} \,|\, \mathcal{M})\, \mathcal{L}(\mathbf{d_{EM}} \,|\, \mathcal{M})$ since the measurements are statistically independent measurements of the chirp mass of the binaries. This gives

$$\mathcal{P}(\mathcal{M} \,|\, \mathbf{d_{GW}}, \mathbf{d_{EM}}) = \frac{\pi(\mathcal{M})}{\mathcal{Z}(\mathbf{d_{GW}}, \mathbf{d_{EM}})} \mathcal{L}(\mathbf{d_{GW}} \,|\, \mathcal{M})\, \mathcal{L}(\mathbf{d_{EM}} \,|\, \mathcal{M}). \tag{6.34}$$

Here $\mathcal{Z}(\mathbf{d_{GW}}, \mathbf{d_{EM}})$ is the normalizing constant that maintains the equality. Since we only consider one parameter $\mathcal{M}$ we can calculate this normalizing constant using a fine-grid trapezoidal rule. We denote this normalizing constant as c from now on. To find the marginal likelihood of $\mathcal{L}(\mathbf{d_{GW}} \,|\, \mathcal{M})$ we use Bayes' theorem from the available marginal posterior distribution on $\mathcal{M}$. That is, we use $\mathcal{L}(\mathbf{d_{GW}} \,|\, \mathcal{M}) = \mathcal{P}(\mathcal{M} \,|\, \mathbf{d_{GW}})/\pi_{\mathrm{GW}}(\mathcal{M})$ for the properly normalized marginal posterior and prior distributions on the chirp mass. We can now express the joint posterior distribution as

$$\mathcal{P}(\mathcal{M} \,|\, \mathbf{d_{GW}}, \mathbf{d_{EM}}) = \frac{\pi(\mathcal{M})}{c} \times \frac{\mathcal{P}(\mathcal{M} \,|\, \mathbf{d_{GW}})}{\pi_{\mathrm{GW}}(\mathcal{M})} \times \frac{\mathcal{P}(\mathcal{M} \,|\, \mathbf{d_{EM}})}{\pi_{\mathrm{EM}}(\mathcal{M})} \tag{6.35}$$

Now, since our electromagnetic observer is purely hypothetical we let $\mathcal{P}(\mathcal{M}|\mathbf{d_{EM}})$ be a Gaussian distribution whose mean value estimation of the chirp mass is centered at the posterior mode of the standard models marginal chirp mass posterior distribution [19]. We will vary the standard deviation of this Gaussian distribution to see when the mock electromagnetic observer constrains the joint observation to be nearly identical to the measurement of the chirp mass from the standard model of [19]. We specify our (hyper) prior $\pi(\mathcal{M}) = \pi_{\mathrm{GW}}(\mathcal{M}) = \pi_{\mathrm{EM}}(\mathcal{M})$. The prior of $\pi_{\mathrm{GW}}(\mathcal{M})$ was uniform in chirp mass in the detector fram between $\mathcal{M} \in (1.1876, 1.2076)$. The mock estimation procedure can be seen in Fig. 30 where we find that an electromagnetic observer would need a constraint on $\sigma_{\mathcal{M}} < 0.0001 M_{\odot}$. This corresponds to a measurement error on the chirp mass of less than 0.017 %, well outside the realm of current methods.

One might consider an improvement on this approach by using the marginal chirp mass distribution when marginalizing over all p-g mode models and then comparing it to the marginal chirp mass distribution when marginalizing over all models in [19]. The result, however, is qualitatively identical.

6.5.5 Strict Constraints on the p-g mode instability

Given the observed parameter degeneracies and the statistically nonsignificant results of the nonlinear tidal hypotheses that we have tested, we now investigate whether there are regions of the parameter space where nonlinear tidal effects are not degenerate with standard waveforms. We do this by thresholding the results of our Bayesian MCMC analysis p-g waveforms on their fitting factor with standard waveforms. We combine the results of our analysis on the uniform mass, $\delta\phi > 0.1$ rads, narrow f_0 prior distribution model to obtain $22,600$ independent samples. We then examine the fitting factor of every independent sample, from every temperature, with a non-spinning, mass-only template bank of TaylorF2 waveforms with comparable masses to GW170817. For simplicity, we only keep the mass parameters and p-g mode parameters in the overlap calculations, since the correlation between nonlinear tidal dynamics is most apparent in the measured chirp mass. When we examine the fitting factor between nonlinear tidal waveforms and this template bank we observe that there is a very high match between standard templates and nonlinear tidal waveforms when $n = 4/3$. The nonlinear tidal waveforms that least match this template bank tend to be those parameterized by large amplitude and large gravitational-wave phase shift. We then recompute the Bayes factor when discarding samples from the analysis below a particular fitting factor with the template bank. To ensure a robustness of the point-estimate we use a bootstrap method to estimate the Monte Carlo error for this Bayes factor estimate [318]. The bootstrap estimated Monte Carlo error tends to be much larger than the convergence error for this analysis and so we neglect inclusion of convergence error in the estimate. A statistically significant Bayes factor of $\sim 30\,(20)$, against nonlinear tides, is found when the waveform has an overlap less than $98.5\,(98.85)\%$ match with the standard waveform, see Fig. 31. While this metric is insufficient to rule out the p-g mode instability, it is a useful metric in understanding why the evidence is nearly identical to the evidence from [19]. We find that portions of the p-g mode parameter space that most contribute towards the evidence come from regions of the parameter space that have a high overlap with standard waveforms. This occurs either through A being too small to induce a large change in the phase of the waveform or through an associated parameter degeneracy with the chirp mass caused by large A, low f_0, and $n \sim 4/3$.

6.6 Discussion

We have used the observation of GW170817 and the model of [328] to look for evidence of nonlinear tides from p-g mode coupling during the inspiral [63, 64, 332]. Over the broad prior space, we find a Bayes factor of unity which gives an inconclusive result on whether nonlinear tides are favored or disfavored in GW170817, consistent with [18]. This Bayes factor can be interpreted as stating that there is insufficient evidence to change our prior beliefs about the credibility of the p-g mode hypothesis after the observation of GW170817. A closer examination of the posterior distribution lead us to conclude that nonlinear tides are consistent with the signal GW170817 because they either cause very small phase shifts to the waveform, or the nonlinear tides must enter the waveform in a way that is degenerate with the other intrinsic parameters of GW170817. Regions of the nonlinear tide parameter space that have a fitting factor of less than 99% (98.5%) are disfavored by a Bayes factor of 15 (25). we find that waveforms from a *p*-*g* mode instability with overlap > 98.5 %, tend to either induce a very small phase shifts to the waveform or are degenerate with other intrinsic parameters of GW170817. This leads us to conclude that modeling GW170817 with nonlinear tidal parameters may not offer advantages over using a simpler model. We conclude that the consistency of the GW170817 signal with the model of [328] is due to parameter degeneracy and that regions where nonlinear tides produce a measurable effect are strongly disfavored.

In principle, one could improve our analysis by separately parameterizing the amplitude, turn-on frequency, and frequency evolution for each star as in [18]. However, we find our results to be broadly consistent with [18], and so we do not expect these to affect the main conclusion of our paper. Further improvements on the parametric model of *p*-*g* mode instability could include a higher order post-Newtonian expansion of the instability, or further understanding of the instability's interaction with neutron star magnetic fields [64]. Nonlinear tides are poorly understood and the contribution from other stellar oscillation modes may yet contribute to a more accurate picture of the interior dynamics of neutron stars [333]. Current models of the gravitational-wave phase shift caused by nonlinear tides from the *p*-*g* mode instability suffer from parameter degeneracies with the other intrinsic parameters of a neutron star binary. A measurement of the binary's chirp mass that is independent of gravitational-wave observations would break this degeneracy. However, for a system like GW170817, this would require measurement of the binary's chirp mass to a precision greater than $\sim 0.02\%$ using an electromagnetic counterpart, which is implausible. Absent improved theoretical understanding of nonlinear tides from *p*-*g* mode coupling that can excludes degenerate regions of the parameter space *a priori*, we do not expect this situation to improve with future detections.

Finally, we now it will ever be possible to accumulate sufficient evidence to rule-in, or rule-out the presence of nonlinear tides due to a *p*-*g* mode instability. When more binary neutron star events are detected by gravitational wave networks it will be possible to take advantage of the fact that we can accumulate evidence for hypotheses across statistically independent events. The Bayes factor for testing the same hypotheses for many events is the product of the Bayes factor for the hypothesis for each individual observation of the merger of binary neutron stars. As more binary neutron star events are detected we can accumulate evidence for or against p-g mode instability through continuous testing of these hypotheses on these individual events. We can also update our parameter inference on the nonlinear tidal parameters to potentially constrain them more sharply. In Chapter 5 we described how to use the Bayes factor for N events to build a combined Bayes factor via following expression. There are no publicly available binary neutron stars other than the observation of GW170817 to build evidence for the nonlinear tidal hypothesis so we consider the hypothetical case where multiple GW170817-like neutron star events are detected.

To do so we consider the results of our current analysis on GW170817. Here we consider two estimators for the logarithm Bayes factor, the thermodynamic integration method which we found to have a log Bayes factor of $\mu \sim -0.38$, and at worst $\sigma \sim 0.1$, and the logspline estimator with the Savage Dickey density ratio which we found to have a log Bayes factor of $\mu \sim -0.46, \sigma \sim 0.06$. While the log Bayes factor for the thermodynamic integration method is formally log-normal, the Bayes factor estimated from the logspline estimator is not formally log-normal. This discrepancy does not significantly affect our proposed analysis. We also consider the analysis of [18] which found a log Bayes factor of $0.03^{+0.70}_{-0.58}$ at 90% confidence using the Savage-Dickey density ratio. We model this as a Gaussian distribution in the logarithm Bayes factor with $\mu = 0.03, \sigma = 0.4$ so as to have a similar 90% interval width. The hypothesis test of [18] is distinct from our own in that the waveform model parametrizes the nonlinear tidal parameters for each star independently and makes different assumptions about the correlation between the nonlinear tidal parameters and the masses of the binaries. This could be considered a systematic difference in the waveform modeling that could potentially impact Bayesian inference. In Chapter 5 we discussed how a combined Bayes factor for continued testing of a hypothesis over many observations could be attained by multiplying the Bayes factor from each observation. We illustrate this method in Fig. 32 where we show the divergence of the Bayes factors after 15 repeat observations of GW170817 for Bayesian hypothesis test techniques and potentially due to waveform systematics. After 15 repeat observations of GW170817 our thermodynamic integration and logspline Savage-Dickey density ratio estimates give a statistically significant result where we can confidently reject the nonlinear tidal hypothesis. Our model of the analysis of [18] suggests a different decision, i.e. statistical significance is not achieved. The assumptions of this approach are very strong and are not motivated by realistic physics, but they highlight the importance of the need for robust understanding of our systematic and statistical uncertainties when we use Bayesian hypothesis testing.

A more realistic approach would be to consider the a realistic population of binary neutron star mergers. Considerations for the source properties have been discussed in this chapter of this dissertation as well as in chapter 3 of this dissertion. A software injection campaign where simulated binary neutron stars drawn from a model of the population of neutron stars could be considered like in chapter 3 of this dissertation. Simulated signals could be added to different Gaussian noise realizations informed by the expected sensitivity of future observing runs from Advanced LIGO and Virgo. Bayesian inference and hypotheses could be performed on each simulated signal and predictions about future events could be made. The largest contributor to the Bayesian inference will in all likelihood be due to signals with the largest signal to noise ratio *rho*. For isotropically and homogeneously distributed binary neutron star mergers we can expect a power-law distribution on the ρ [346, 347]. More specifically, we can expect that for a network of interferometers with a signal to noise ratio detection threshold of $\rho_{\text{threshold}}$ that our distribution will follow

$$p(\rho) = 3\,\frac{\rho_{\text{threshold}}^{3}}{\rho^{4}}. \tag{6.36}$$

This expression is a normalized probability distribution function in the domain that $\rho > \rho_{\text{threshold}}$. The signal to noise ratio ρ is permitted to go to positive infinity. From Eq. (6.36) we can expect an average ρ to be equal to $\frac{3}{2}\rho_{\text{threshold}}$. If we assume a very conservative $\rho_{\text{threshold}} = 11$, then the probability of observing a gravitational wave neutron star mergers with signal to noise ratio greater than or equal to the signal to noise ratio of GW170817 ($\rho \approx 34$) is ~ 3 %. At a signal to noise ratio of ~ 34 we have found that the *p*-*g* mode instability hypothesis has a Bayes factor of approximately 1. We expect that 97% of neutron star detections will have a lower signal to noise ratio than GW170817 and so we expect that the Bayes factor will be less informative for the p-g mode instability hypothesis than it was for the observation of GW170817. For these 97% of mergers the parameter degeneracies between the nonlinear tidal parameters and the other intrinsic parameters of the binary may be more pronounced and less informative. Moreover, the study of [328] found that the nonlinear tidal parameters are degenerate with the inferred luminosity distance of the binary. While the waveform model of [328] incorrectly used the stationary phase approximation in the construction of the waveform model, c.f. this study and [18], this degeneracy could potentially obscure future studies of the nonlinear tidal hypothesis with future binary neutron stars. With all of this in mind, i.e. the difficulties in parameter estimation and Bayesian hypothesis testing, we may have to wait for hundreds of binary neutron star mergers to accumulate sufficient evidence to make a decision on whether binary neutron stars are contain nonlinear tides from a p-g mode instability. Additional choices on conductin and tuning an MCMC analysis in addition to what choice of prior distributions in the modeling will make this endeavor all the more difficult.

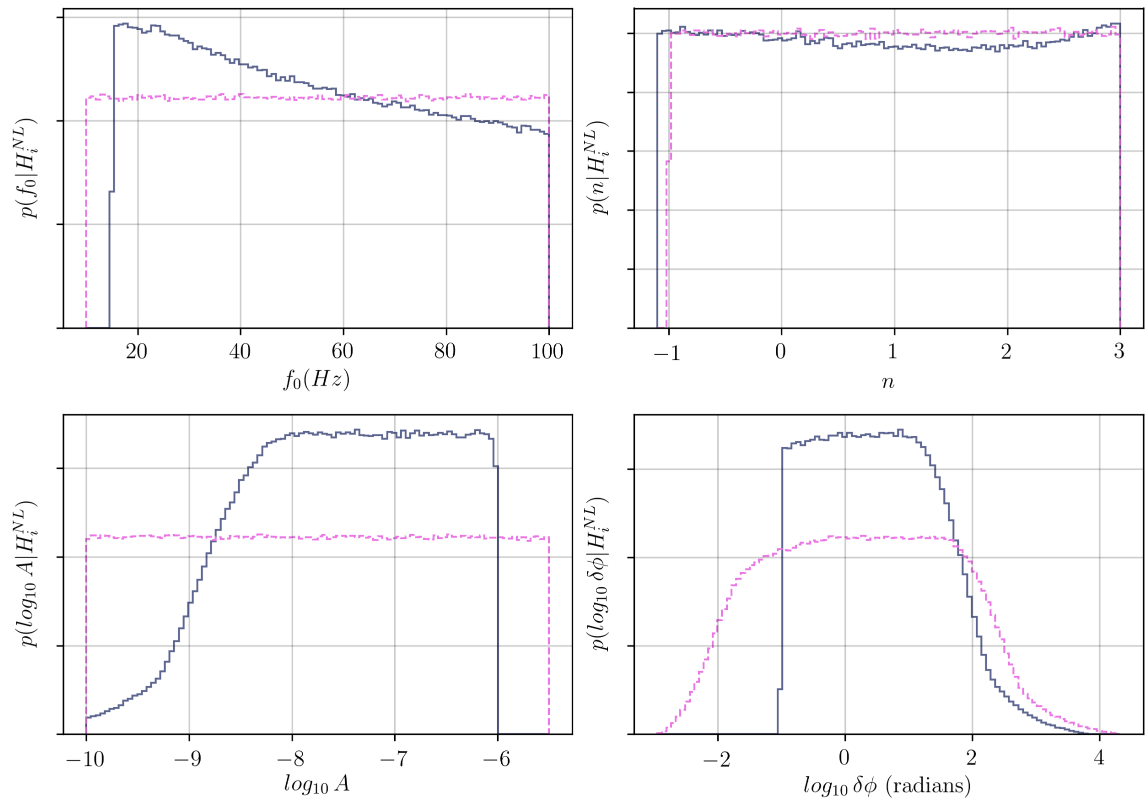


Figure 25: Prior probability distributions on the parameters (f_0, n, A) for the waveform model $\mathrm{H}^{\mathrm{NL}} = \mathrm{H}_{\mathrm{TaylorF2+NL}}$ and the resulting prior on the gravitational wave phase shift $\delta\phi$ shift due to nonlinear tides. The dark blue, solid lines shows the priors when f_0 is drawn from a uniform distribution between 15 and 100 Hz with a $\delta\phi \geq 0.1$ rad constraint restricting some of the prior space. The pink, dotted lines represent prior distributions on the nonlinear tidal parameters similar to [18].

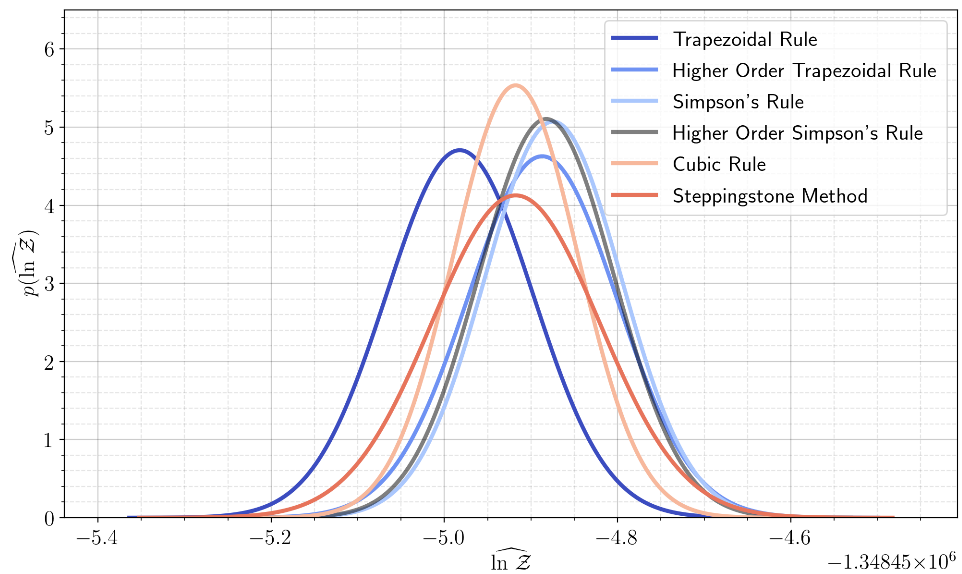


Figure 26: The estimates of the logarithm of the evidence from multi-temper evidence integration methods. We model the logarithm of the evidence as a Gaussian in log-space. These data are for the logarithm of the evidence from the unconstrained $\delta\phi$ prior for the p-g mode instability model. The trapezoidal rule estimates the lowest log evidence for this model, and the cubic rule has the smallest estimated statistical error uncertainty (the smallest confidence interval). The mean values of the higher order quadrature rules appear to be closer together to one another than they are to the trapezoidal rule.

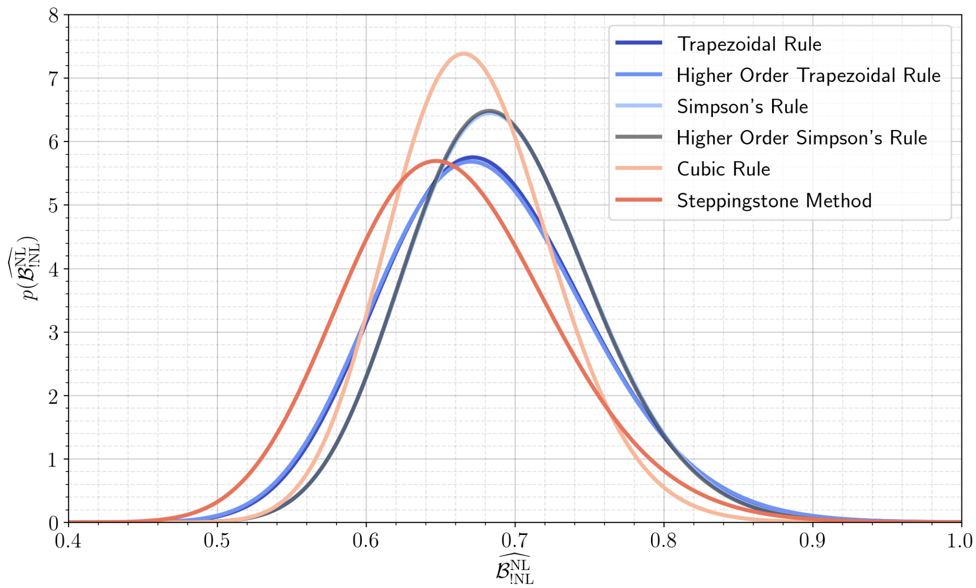


Figure 27: The distribution for the Bayes factor for nonlinear tides from p-g mode instability from the unconstrained $\delta\phi$ prior relative to the uniform mass, common equation of state prior from [19] under the assumption that the logarithm of the evidence for each model is well approximated by a Gaussian distribution. but our method is sufficiently accurate in the high-sample limit. When the uncertainty on the logarithm of the evidences in the Bayes factor estimation are sufficiently small, the Bayes factor distribution is approximately normal in shape, but formally they are log-normal distributions.

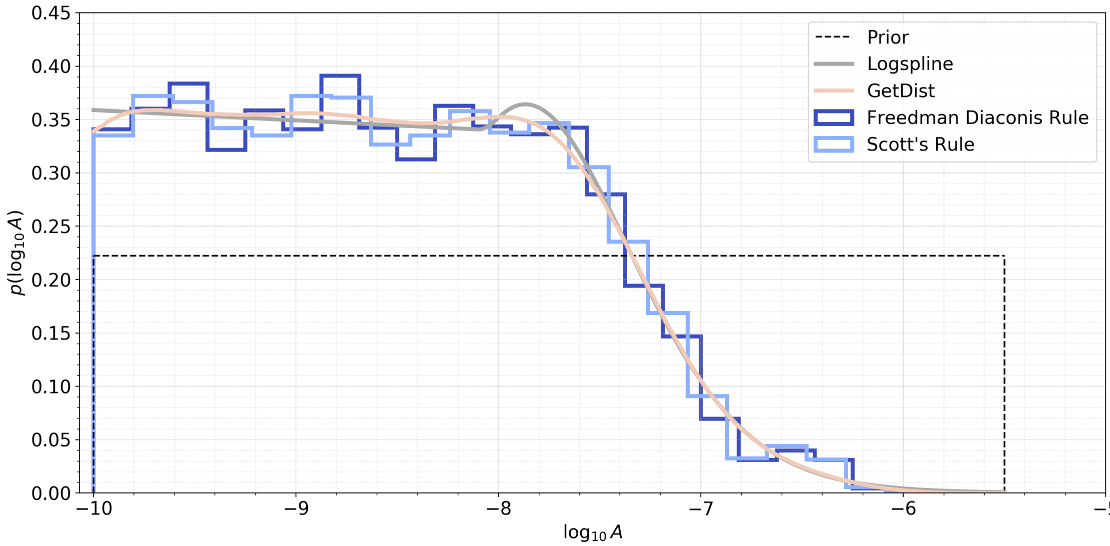


Figure 28: The prior and posterior density estimations from different density estimators for the parameter $\log_{10} A$. The prior density is uniform in $\log_{10}$ and is 0.2 between -10 and -5.5. The Logspline curve (dark grey) is the density estimation under the logspline density estimator. The GetDist (light pink curve) is the Gaussian kernel density estimator described in [20]. The histograms are FD and Scott for the Freedman-Diaconis binning rule and Scott's binning rule, respectively. We can see here that there is some wasted prior space at large $\log_{10} A$. Removing this low-likelihood region from the prior hypothesis model would likely move the p-g mode instability Bayes factor closer to unity.

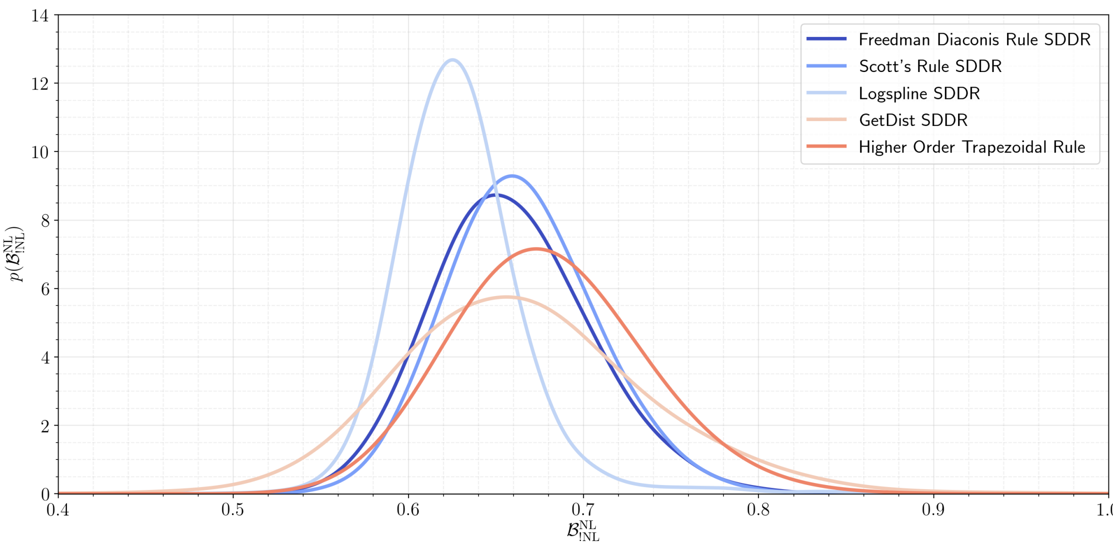


Figure 29: A comparison of the Bayes factor estimates for p-g mode instability with the permissive prior on $\delta\phi$ vs no p-g mode instability from different methods. Here, SDDR refers to the Savage Dickey density ratio test for each corresponding estimator technique. We compare these results to the higher order trapezoidal rule from thermodynamic integration. The other multi-tempered Bayes factors are comparable to the one shown here and so are not displayed. The estimates generally agree as can be seen from comparing values in Table 8 and Table 9.

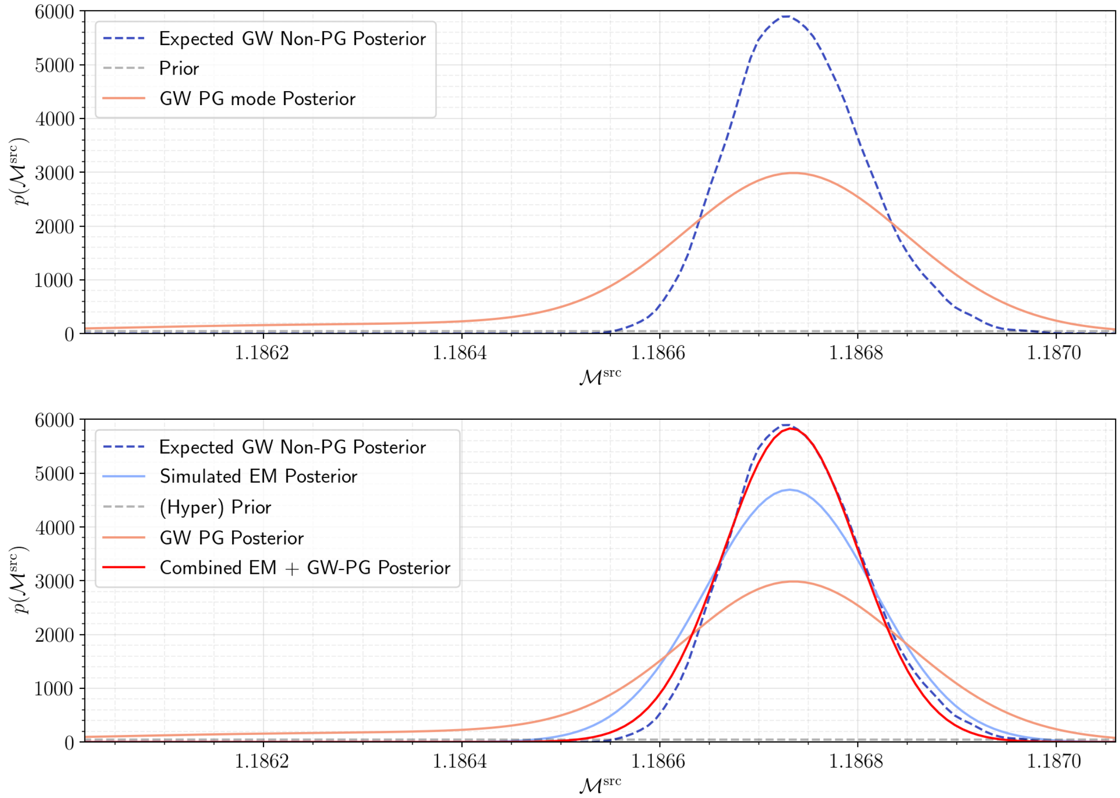


Figure 30: (*Top*) The prior distribution on the chirp mass for two gravitational wave astrophysical hypotheses. The first hypothesis is the uniform mass and constrained equation of state constraint model from [19], while the second model is the p-g mode instability hypothesis with unconstrained $\delta\phi$. The marginal posterior distributions on the chirp mass are in dashed-blue and solid, light-red, respectively. (*Bottom*) Combining a simulated Gaussian electromagnetic posterior on the chirp mass (light-blue) and a prior on the chirp mass we can combine the posterior distributions from the gravitational wave data with the p-g mode instability from the unconstrained $\delta\phi$ model with this electromagnetic posterior to construct a joint posterior distribution (solid, red) that closely matches the inferred chirp mass for GW170817 from [19]. The simulated Gaussian electromagnetic posterior has mean centered at the maximum a posteriori value from [19], $\mu = 1.186731\, M_{\odot}$, and standard deviation, $\sigma = 0.000085\, M_{\odot}$.

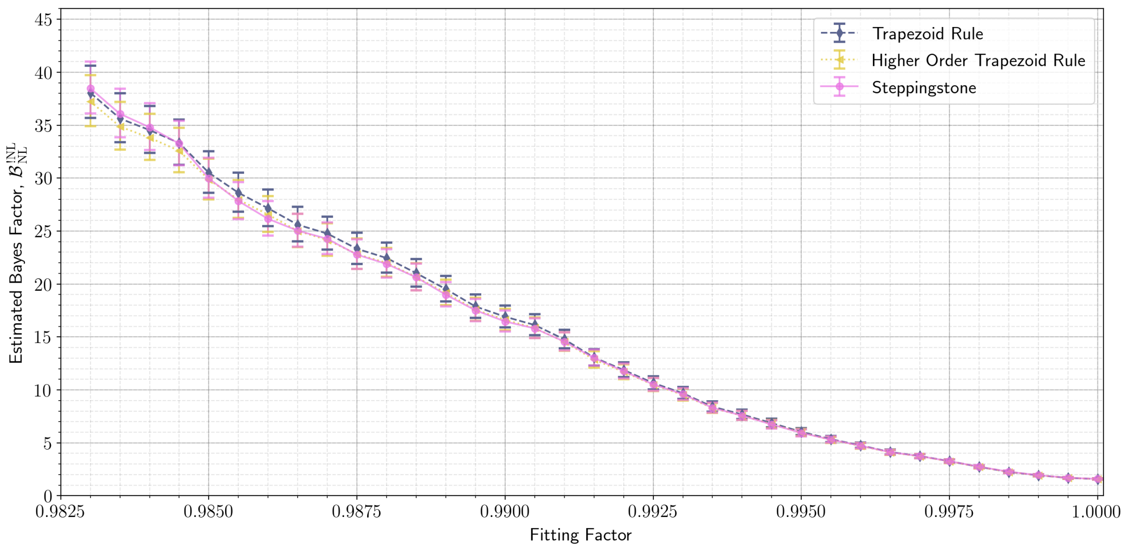


Figure 31: The estimated Bayes factors for nonlinear tidal parameters when the samples are filtered by the fitting factor to a non-spinning, mass-only template bank of TaylorF2 waveforms. The convention in Bayes factor is switched from the main body of the text to represent the Bayes factor for the ratio of evidence for no nonlinear tides, $p\left(\mathbf{d} \,|\, \mathrm{H}_{\mathrm{TaylorF2}}\right)$, to the evidence for nonlinear tides, $p\left(\mathbf{d} \,|\, \mathrm{H}_{\mathrm{TaylorF2+NL}}\right)$. This is abbreviated as $\mathcal{B}^{!\mathrm{NL}}_{\mathrm{NL}}$. The three methods for estimating the Bayes factor are the thermodynamic integration method from trapezoid rule integration (dark grey, dashed line), the thermodynamic integration method from the higher order trapezoid rule (yellow, small-dashed line), and the steppingstone algorithm (dark pink, solid line). A bootstrap method is used to estimate approximate errors on the Bayes Factors. Error bars represent 5th and 95th percentiles. The sampling error becomes large at a fitting factor $\lesssim 99\%$.

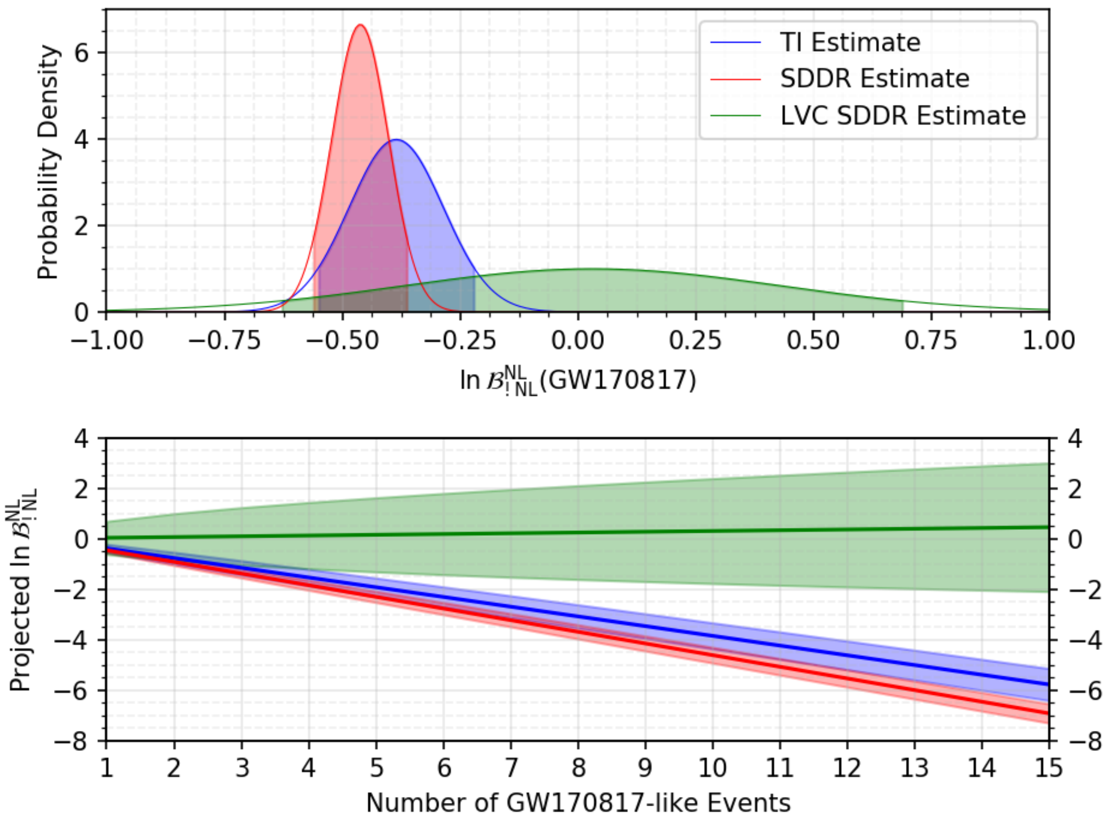


Figure 32: (*Top*) A comparison of Gaussian approximations of the logarithm of the Bayes factor using different estimators or waveform systematics. Note that the LVC estimate here is a rough Gaussian approximation based on the reported bounds in [18]. The 90% confidence regions are shaded in. Positive log Bayes factors are indicative of support for the p-g mode hypothesis, while negative log Bayes factors are indicative of support for the null hypothesis. (*Bottom*) For repeated GW170817-like binary neutron star mergers the cumulative logarithm of the Bayes factor for the p-g mode hypothesis vs the null hypothesis begin to diverge in estimation. The solid lines represent the cumulative median estimates, while the shaded regions represent the cumulative 90% confidence intervals. Waveform systematics or uncontrolled variables in the Bayes factor estimation methods may be the main driver of this divergence and future meta-analyses will have to control for these sorts of uncertainty.

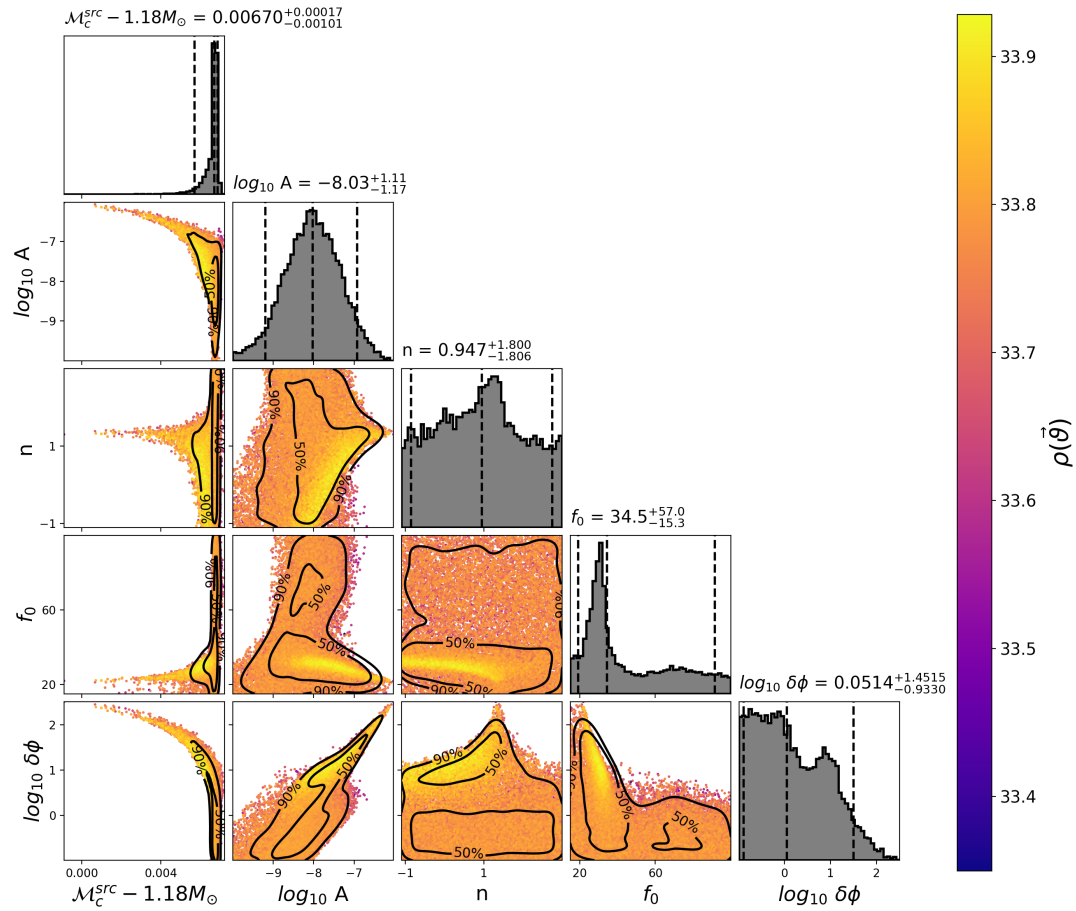


Figure 33: The marginalized posterior distributions for the uniform mass prior and a f_0 restricted to the range 15 and 100 Hz. The vertical lines on the marginalized histograms display the 5th, 50th, and 95th percentiles of the posteriors. The three-detector network signal to noise ratio for each sample is given on the color-bar. The posterior scatter plots show 50% and 90% credible interval contours. The posteriors on n is peaked $n \lesssim 4/3$ and for values of f_0 close to the lower end of the detector's low frequency sensitivity. In this region of parameters space, the effect of nonlinear tides is degenerate with chirp mass, causing a secondary peak in the chirp mass posterior. It can be seen from the $\delta\phi$–$\mathcal{M}$ plot (lower left) that large phase shifts due to nonlinear tides are due to points in parameter space where a value of chirp mass can be found that compensates for the phase shift of the nonlinear tides. These are the combined posteriors from 9 runs. It is notable that the the peaks in the f_0 posterior, at $f_0 \approx 30$ Hz and $f_0 \approx 70$ Hz seem to be reversed from those in Fig 2. of [18]. Note that the marginalized posterior for A is diminished for $A < 10^{-8}$ due to the $\delta\phi$ prior constraint.

Hypothesis Tested	$\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}(A)$	$\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}(B)$	$\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}(C)$	$\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}(D)$	$\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}(E)$	$\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}(F)$
H_1 (Uniform Mass, A, n, $f_0 \in (15, 100)$ Hz), $\delta\phi > 0.1$	$0.63^{+0.08}_{-0.07}$	$0.63^{+0.08}_{-0.07}$	$0.64^{+0.07}_{-0.06}$	$0.64^{+0.05}_{-0.05}$	$0.63^{+0.06}_{-0.06}$	$0.63^{+0.06}_{-0.06}$
H_2 (Gaussian Mass, A, n, $f_0 \in (15, 100)$ Hz), $\delta\phi > 0.1$	$0.71^{+0.08}_{-0.07}$	$0.71^{+0.08}_{-0.07}$	$0.70^{+0.06}_{-0.06}$	$0.70^{+0.04}_{-0.04}$	$0.71^{+0.06}_{-0.06}$	$0.73^{+0.07}_{-0.07}$
H_3 (Uniform Mass, A, n, $f_0 \in (15, 800)$ Hz), $\delta\phi > 0.1$	$0.64^{+0.09}_{-0.08}$	$0.64^{+0.09}_{-0.08}$	$0.64^{+0.08}_{-0.07}$	$0.64^{+0.06}_{-0.06}$	$0.64^{+0.06}_{-0.06}$	$0.63^{+0.09}_{-0.07}$
H_4 (Gaussian Mass, A, n, $f_0 \in (15, 800)$ Hz), $\delta\phi > 0.1$	$0.76^{+0.08}_{-0.07}$	$0.76^{+0.08}_{-0.07}$	$0.75^{+0.06}_{-0.06}$	$0.75^{+0.04}_{-0.04}$	$0.75^{+0.05}_{-0.05}$	$0.76^{+0.07}_{-0.06}$
H_5 (Uniform Mass, A, n, $f_0 \in (10, 100)$ Hz)	$0.68^{+0.12}_{-0.11}$	$0.68^{+0.13}_{-0.11}$	$0.69^{+0.11}_{-0.1}$	$0.69^{+0.1}_{-0.09}$	$0.67^{+0.1}_{-0.08}$	$0.65^{+0.13}_{-0.11}$

Table 8: The various Bayes factors under different multi-tempered integration methods. The column marked with $\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}(A)$ is the Bayes factor under the thermodynamic integration method using the trapezoid quadrature rule. The (B) column is the Bayes factor from the thermodynamic integration method using the higher-order trapezoid quadrature rule. The (C) column is the Bayes factor from the thermodynamic integration method using Simpson's quadrature rule. The (D) column is the Bayes factor for the thermodynamic integration method using Simpson's higher-order quadrature rule. The (E) column is the Bayes factor for the thermodynamic integration method using a cubic polynomial quadrature rule. And (F) is the Bayes factor from the steppingstone method. The 50th percentile with the 5th and 95th percentiles in the plus and minus superscripts and subscripts, respectively, are shown above.

Hypothesis Tested	$\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}$(FD)	$\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}$(Scott)	$\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}$(Gaussian KDE)	$\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}$(Logspline)
H_5 (Uniform Mass, A, n, $f_0 \in (10, 100)$ Hz)	$0.66^{+0.08}_{-0.07}$	$0.66^{+0.08}_{-0.07}$	$0.66^{+0.13}_{-0.1}$	$0.63^{+0.06}_{-0.05}$

Table 9: The various Bayes factors from the Savage-Dickey Density Ratio test under different density estimators. The column marked with $\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}$(FD) is the Bayes factor from the Freedman Diaconis histogram binning rule. The $\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}$(Scott) column is the Bayes factor estimate under Scott's histogram binning rule. The $\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}$(Gaussian KDE) column is the Bayes factor estimate when using the Gaussian kernel density estimator with linear boundary bias corrections as found in the GetDist Python package. The column denoted as $\mathcal{B}^{\mathrm{NL}}_{!\mathrm{NL}}$(Logspline) is the Bayes factor estimate when using the logspline density estimator. The 50th percentile with the 5th and 95th percentiles in the plus and minus superscripts and subscripts, respectively.

Chapter 7

Conclusions

In Chapter 2, we discussed the PyCBC search pipeline, a matched-filter search pipeline for the detection of compact binary coalescence, and the results that it gathered during LIGO's first observing run. In the first observing run, the LIGO detectors observed gravitational waves from the merger of two stellar-mass black holes, GW150914. The binary coalescence search detects GW150914 with a significance greater than 5.1σ during this first observing run. Detailed parameter estimation for GW150914 is reported in Ref. [2], the implications for the rate of binary black hole coalescences in Ref. [348], and tests for consistency of the signal with general relativity in Ref. [349]. Ref. [350] discusses the astrophysical implications of this discovery. During the first observing run PyCBC also discovered a second binary merger, GW151226 [136]. A third gravitational wave candidate, LVT151012, was also discovered but was only found with a false alarm rate less than 0.44 yr^{-1}, and could not be confidently claimed as a gravitational wave candidate.

In Chapter 3, we reported the non-detection of binary neutron stars and neutron star-black hole mergers in Advanced LIGO's first observing run. We estimated the sensitive volume of Advanced LIGO to such systems and were able to place 90% confidence upper limits on the rates of binary neutron star and neutron star-black hole mergers, improving upon limits obtained from Initial LIGO and Initial Virgo by roughly an order of magnitude. Specifically, we constrained the merger rate of binary neutron star systems with component masses of $1.35 \pm 0.13 M_\odot$ to be less than 12,600 $\mathrm{Gpc}^{-3}\ \mathrm{yr}^{-1}$. We also constrained the rate of neutron star-black hole systems with neutron star masses of $1.4 M_\odot$ and black hole masses of at least $5 M_\odot$ to be less than 3,210 $\mathrm{Gpc}^{-3}\ \mathrm{yr}^{-1}$ for a population where the component spins are (anti-)aligned with the orbit. Lastly, we constrained the rate of neutron star-black hole systems with isotropic spin distributions in the components of the spin direction to be less than 3,600 $\mathrm{Gpc}^{-3}\ \mathrm{yr}^{-1}$.

We compared these upper limits with existing astrophysical rate models and found that the current upper limits are in conflict with only the most optimistic models of the merger rate for binary systems with neutron stars. For continued non-detections of binary neutron star mergers and neutron star-black hole mergers in the second and third observing runs, we estimated plausible upper limits on the rate of these mergers given estimates of the detector sensitivity during the second and third observing runs. Finally, we have explored the implications of this non-detection of binary neutron stars and neutron star-black hole binaries on the beaming angle of short GRB. We find that, if one assumes that all GRB are produced by binary neutron star mergers, then the opening angle of gamma-ray radiation must be larger than $2.3^{+1.7}_{-1.1}{}^{\circ}$; or larger than $4.3^{+3.1}_{-1.9}{}^{\circ}$ if one assumes all GRB are produced by neutron star-black hole mergers.

In Chapter 4, we presented a full catalog of gravitational-wave events and candidates from a `PyCBC`based, templated, matched-filter search of the LIGO O1 open data. Our analysis improved upon [17, 246] and the analysis of Chapter 2 by using improved ranking of candidates via a phase, amplitude and time delay consistency check, an improved background model, and a template bank targeting a wider range of sources [45, 46, 47]. We verifed the discovery of GW150914 and GW151226 and report an improved statistical significance of the candidate event LVT151012. In the analysis of [17, 246] LVT151012 was found to have a false alarm rate of approximately 1 per 2 years, but in the analysis of 1-OGC we found that LVT151012 could be instead found with a false alarm of 1 per 24 years. If the analysis had restricted itself to a search of the parameter space where binary black holes had been discovered before, the false alarm rate could have been estimated at 1 per 446 years. We also found that in our analysis the probability of LVT151012 being of astrophysical origin is approximately 98%. With these improvements of the statistical significance estimation we confidently claim LVT151012 as a gravitational wave event and designate it GW151012. Apart from the detections of GW150914, GW151012, and GW151226, none of the other candidate events in the 1-OGC analysis were found to be statistically significance. All of these candidates are listed in our catalog available at www.github.com/gwastro/1-ogc.

In Chapter 5 we developed tools for Bayesian hypothesis testing. We discussed how to calculate the Bayes factor, a likelihood ratio used to evaluate the relative statistical significance of hypotheses. We looked at Markov-Chain Monte Carlo methods for calculating Bayes factors such as the thermodynamic integration method and the steppingstone method. We also introduced the Savage-Dickey density ratio method for calculating Bayes factors of nested hypotheses.

In Chapter 6, we examined the detection of GW170817, a binary neutron star merger discovered by LIGO and Virgo during their second observing run [54]. We conducted Bayesian parameter estimation and hypothesis testing to examine whether nonlinear tides from a nonresonant, nonlinear p-g mode instability were compatible with the observation of GW170817. Our resulting analysis showed that nonlinear tides were broadly compatible with the observation of GW170817, although we found that this occurred because the nonlinear tides either did not cause a measurable change to the waveform or the nonlinear tidal parameters were degenerate with the other intrinsic parameters in the signal. We also found that we could rule out nonlinear tides from a p-g mode instability that matched standard waveforms with $< 98.5\%$ match with a Bayes factor of ~ 25.

The field of astrophysics will in the not too distant future be able to answer many long standing questions regarding compact binaries through increased number of gravitational wave detections. In this thesis, we presented methods for investigating astrophysical implications for non-detections of gravitational waves as well as methods for improving the sensitivity of compact binary coalescence searches towards already detected classes of binary systems. We also developed Bayesian hypothesis testing methods for investigating astrophysical models on detected signals.

Bibliography

[1] J. Veitch et al. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library. *Phys. Rev.*, D91(4):042003, 2015.

[2] B. P. Abbott et al. Properties of the binary black hole merger GW150914. 2016. arXiv:1602.03840.

[3] Roberto Trotta. Bayes in the sky: Bayesian inference and model selection in cosmology. *Contemporary Physics*, 49(2):71–104, 2008.

[4] Mehran Kardar. *Statistical physics of particles*. Cambridge University Press, 2007.

[5] Milton Abramowitz and Irene A Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation, 1965.

[6] J. Abadie et al. Predictions for the Rates of Compact Binary Coalescences Observable by Ground-based Gravitational-wave Detectors. *Class.Quant.Grav.*, 27:173001, 2010.

[7] Chunglee Kim, Benetge Bhakthi Pranama Perera, and Maura A. McLaughlin. Implications of PSR J0737-3039B for the Galactic NS-NS Binary Merger Rate. *Mon. Not. Roy. Astron. Soc.*, 448(1):928–938, 2013.

[8] Wen-fai Fong, Edo Berger, Raffaella Margutti, and B. Ashley Zauderer. A Decade of Short-duration Gamma-ray Burst Broad-band Afterglows: Energetics, Circumburst Densities, and Jet Opening Angles. 2015.

[9] K. Siellez, M. Bor, and B. Gendre. Simultaneous event detection rates by electromagnetic and gravitational wave detectors in the Advanced Era of LIGO and Virgo. *Mon. Not. Roy. Astron. Soc.*, 437(1):649–655, 2014.

[10] David Coward, Eric Howell, Tsvi Piran, Giulia Stratta, Marica Branchesi, Omer Bromberg, Bruce Gendre, Ronald Burman, and Dafne Guetta. The Swift short gamma-ray burst rate density: implications for binary neutron star merger rates. *Mon. Not. Roy. Astron. Soc.*, 425:1365, 2012.

[11] Carlo Enrico Petrillo, Alexander Dietz, and Marco Cavaglia. Compact object coalescence rate estimation from short gamma-ray burst observations. *Astrophys. J.*, 767:140, 2013.

[12] Zhi-Ping Jin, Xiang Li, Zach Cano, Stefano Covino, Yi-Zhong Fan, and Da-Ming Wei. The Light Curve of the Macronova Associated With the Longshort Burst GRB 060614. *Astrophys. J.*, 811(2):L22, 2015.

[13] E. Vangioni, S. Goriely, F. Daigne, P. Francois, and K. Belczynski. Cosmic Neutron Star Merger Rate and Gravitational Waves constrained by the R Process Nucleosynthesis. 2015.

[14] S. E. de Mink and K. Belczynski. Merger rates of double neutron stars and stellar origin black holes: The Impact of Initial Conditions on Binary Evolution Predictions. 2015.

[15] M. Dominik, E. Berti, R. O'Shaughnessy, I. Mandel, K. Belczynski, C. Fryer, D. Holz, T. Bulik, and F. Pannarale. Double Compact Objects III: Gravitational Wave Detection Rates. *Astrophys. J.*, 806(2):263, 2015.

[16] J. Aasi et al. Prospects for Localization of Gravitational Wave Transients by the Advanced LIGO and Advanced Virgo Observatories. 2013.

[17] B. P. Abbott et al. Binary Black Hole Mergers in the first Advanced LIGO Observing Run. *Phys. Rev.*, X6(4):041015, 2016. [erratum: Phys. Rev.X8,no.3,039903(2018)].

[18] BP Abbott, R Abbott, TD Abbott, F Acernese, K Ackley, C Adams, T Adams, P Addesso, RX Adhikari, VB Adya, et al. Constraining the p-mode–g-mode tidal instability with gw170817. *Physical review letters*, 122(6):061104, 2019.

[19] Soumi De, Daniel Finstad, James M Lattimer, Duncan A Brown, Edo Berger, and Christopher M Biwer. Tidal deformabilities and radii of neutron stars from the observation of gw170817. *Physical review letters*, 121(9):091102, 2018.

[20] Antony Lewis. Getdist: Kernel density estimation. *Available from download at https://cosmologist.info/notes/GetDist.pdf*, 2015.

[21] B. P. Abbott et al. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, 2016.

[22] B. P. Abbott et al. Observing gravitational-wave transient GW150914 with minimal assumptions. 2016. arXiv:1602.03843.

[23] 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.

[24] A. Goldstein, P. Veres, E. Burns, M. S. Briggs, R. Hamburg, D. Kocevski, C. A. Wilson-Hodge, R. D. Preece, S. Poolakkil, O. J. Roberts, C. M. Hui, V. Connaughton, J. Racusin, A. von Kienlin, T. Dal Canton, N. Christensen, T. Littenberg, K. Siellez, L. Blackburn, J. Broida, E. Bissaldi, W. H. Cleveland, M. H. Gibby, M. M. Giles, R. M. Kippen, S. McBreen, J. McEnery, C. A. Meegan, W. S. Paciesas, and M. Stanbro. An ordinary short gamma-ray burst with extraordinary implications:fermi-GBM detection of GRB 170817a. *The Astrophysical Journal*, 848(2):L14, oct 2017.

[25] B. P. Abbott et. al. Multi-messenger observations of a binary neutron star merger. *The Astrophysical Journal*, 848(2):L12, oct 2017.

[26] K S Thorne. Gravitational radiation. In S. W. Hawking and W. Israel, editors, *Three hundred years of gravitation*, chapter 9, pages 330–458. Cambridge University Press, Cambridge, 1987.

[27] B.S. Sathyaprakash and S.V. Dhurandhar. Choice of filters for the detection of gravitational waves from coalescing binaries. *Phys.Rev.*, D44:3819–3834, 1991.

[28] Curt Cutler et al. The Last three minutes: issues in gravitational wave measurements of coalescing compact binaries. *Phys. Rev. Lett.*, 70:2984–2987, 1993.

[29] Lee S. Finn. Detection, measurement and gravitational radiation. *Phys. Rev.*, D46:5236–5249, 1992.

[30] Lee Samuel Finn and David F. Chernoff. Observing binary inspiral in gravitational radiation: One interferometer. *Phys. Rev.*, D47:2198–2219, 1993.

[31] S.V. Dhurandhar and B.S. Sathyaprakash. Choice of filters for the detection of gravitational waves from coalescing binaries. 2. Detection in colored noise. *Phys.Rev.*, D49:1707–1722, 1994.

[32] R. Balasubramanian, B.S. Sathyaprakash, and S.V. Dhurandhar. Gravitational waves from coalescing binaries: Detection strategies and Monte Carlo estimation of parameters. *Phys.Rev.*, D53:3033–3055, 1996.

[33] Eanna E. Flanagan and Scott A. Hughes. Measuring gravitational waves from binary black hole coalescences: 1. Signal-to-noise for inspiral, merger, and ringdown. *Phys. Rev.*, D57:4535–4565, 1998.

[34] L A Wainstein and V D Zubakov. *Extraction of signals from noise*. Prentice-Hall, Englewood Cliffs, NJ, 1962.

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

[36] Luc Blanchet. Gravitational radiation from post-newtonian sources and inspiralling compact binaries. *Living Reviews in Relativity*, 9(1):4, Jun 2006.

[37] A. Buonanno and T. Damour. Effective one-body approach to general relativistic two-body dynamics. *Phys.Rev.*, D59:084006, 1999.

[38] Alessandra Buonanno and Thibault Damour. Transition from inspiral to plunge in binary black hole coalescences. *Phys. Rev. D*, 62:064015, 2000.

[39] Thibault Damour, Piotr Jaranowski, and Gerhard Schaefer. On the determination of the last stable orbit for circular general relativistic binaries at the third postNewtonian approximation. *Phys. Rev.*, D62:084011, 2000.

[40] Thibault Damour. Coalescence of two spinning black holes: an effective one-body approach. *Phys. Rev.*, D64:124013, 2001.

[41] Parameswaran Ajith et al. Phenomenological template family for black-hole coalescence waveforms. *Class. Quantum Grav.*, 24:S689–S699, 2007.

[42] P. Ajith et al. Inspiral-merger-ringdown waveforms for black-hole binaries with non-precessing spins. *Phys. Rev. Lett.*, 106:241101, 2011.

[43] L. Santamaria, F. Ohme, P. Ajith, B. Bruegmann, N. Dorband, et al. Matching post-Newtonian and numerical relativity waveforms: systematic errors and a new phenomenological model for non-precessing black hole binaries. *Phys.Rev.*, D82:064016, 2010.

[44] LVC O1 BBH PRX. In prep.

[45] Alexander H. Nitz, Thomas Dent, Tito Dal Canton, Stephen Fairhurst, and Duncan A. Brown. Detecting binary compact-object mergers with gravitational waves: Understanding and Improving the sensitivity of the PyCBC search. *Astrophys. J.*, 849(2):118, 2017.

[46] Alexander Harvey Nitz. Distinguishing short duration noise transients in LIGO data to improve the PyCBC search for gravitational waves from high mass binary black hole mergers. *Class. Quant. Grav.*, 35(3):035016, 2018.

[47] Tito Dal Canton and Ian W. Harry. Designing a template bank to observe compact binary coalescences in Advanced LIGO's second observing run. 2017.

[48] Michele Vallisneri, Jonah Kanner, Roy Williams, Alan Weinstein, and Branson Stephens. The LIGO Open Science Center. *J. Phys. Conf. Ser.*, 610(1):012021, 2015.

[49] Nicolas Lartillot and Hervé Philippe. Computing bayes factors using thermodynamic integration. *Systematic biology*, 55(2):195–207, 2006.

[50] Nial Friel and Anthony N Pettitt. Marginal likelihood estimation via power posteriors. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(3):589–607, 2008.

[51] Wangang Xie, Paul O Lewis, Yu Fan, Lynn Kuo, and Ming-Hui Chen. Improving marginal likelihood estimation for bayesian phylogenetic model selection. *Systematic biology*, 60(2):150–160, 2010.

[52] Ward Edwards, Harold Lindman, and Leonard J Savage. Bayesian statistical inference for psychological research. *Psychological review*, 70(3):193, 1963.

[53] James M Dickey. The weighted likelihood ratio, linear hypotheses on normal location parameters. *The Annals of Mathematical Statistics*, pages 204–223, 1971.

[54] B.P. Abbott et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Phys. Rev. Lett.*, 119(16):161101, 2017.

[55] I. Tews, J. Margueron, and S. Reddy. How well does GW170817 constrain the equation of state of dense matter? 2018.

[56] Elias R. Most, L. Jens Papenfort, Veronica Dexheimer, Matthias Hanauske, Stefan Schramm, Horst Stcker, and Luciano Rezzolla. Signatures of quark-hadron phase transitions in general-relativistic neutron-star mergers. 2018.

[57] Carolyn Raithel, Feryal zel, and Dimitrios Psaltis. Tidal deformability from GW170817 as a direct probe of the neutron star radius. *Astrophys. J.*, 857(2):L23, 2018.

[58] B. P. Abbott et al. GW170817: Measurements of neutron star radii and equation of state. 2018.

[59] B. P. Abbott et al. Properties of the binary neutron star merger GW170817. 2018.

[60] David Radice and Liang Dai. Multimessenger Parameter Estimation of GW170817. *Eur. Phys. J.*, A55(4):50, 2019.

[61] Benjamin P Abbott et al. Model comparison from LIGO-Virgo data on GW170817's binary components and consequences for the merger remnant. 2019.

[62] Collin D. Capano, Ingo Tews, Stephanie M. Brown, Ben Margalit, Soumi De, Sumit Kumar, Duncan A. Brown, Badri Krishnan, and Sanjay Reddy. GW170817: Stringent constraints on neutron-star radii from multimessenger observations and nuclear theory. 2019.

[63] Nevin N. Weinberg, Phil Arras, and Joshua Burkart. An instability due to the nonlinear coupling of p-modes to g-modes: Implications for coalescing neutron star binaries. *Astrophys. J.*, 769:121, 2013.

[64] Nevin N. Weinberg. Growth rate of the tidal p-mode g-mode instability in coalescing binary neutron stars. *Astrophys. J.*, 819(2):109, 2016.

[65] Tito Dal Canton et al. Implementing a search for aligned-spin neutron star-black hole systems with advanced ground based gravitational wave detectors. *Phys. Rev.*, D90(8):082004, 2014.

[66] Samantha A. Usman et al. An improved pipeline to search for gravitational waves from compact binary coalescence. 2015.

[67] Alexander H. Nitz, Ian W. Harry, Joshua L. Willis, Christopher M. Biwer, Duncan A. Brown, Larne P. Pekowsky, T. Dal Canton, Andrew R. Williamson, Thomas Dent, Collin D. Capano, Thomas J. Massinger, Amber K. Lenon, Alex B. Nielsen, and Miriam Cabero. PyCBC Software. https://github.com/gwastro/pycbc, 2018.

[68] Kipp Cannon, Romain Cariou, Adrian Chapman, Mireia Crispin-Ortuzar, Nickolas Fotopoulos, et al. Toward Early-Warning Detection of Gravitational Waves from Compact Binary Coalescence. *Astrophys.J.*, 748:136, 2012.

[69] Stephen Privitera, Satyanarayan R. P. Mohapatra, Parameswaran Ajith, Kipp Cannon, Nickolas Fotopoulos, Melissa A. Frei, Chad Hanna, Alan J. Weinstein, and John T. Whelan. Improving the sensitivity of a search for coalescing binary

black holes with nonprecessing spins in gravitational wave data. *Phys. Rev.*, D89(2):024003, 2014.

[70] C. Messick, K. Blackburn, P. Brady, P. Brockill, K. Cannon, S. Caudill, S. J. Chamberlin, J. D. E. Creighton, R. Everett, C. Hanna, R. N. Lang, T. G. F. Li, D. Meacher, C. Pankow, S. Privitera, H. Qi, S. Sachdev, L. Sadeghian, B. Sathyaprakash, L. Singer, E. G. Thomas, L. Wade, M. Wade, and A. Weinstein. Analysis Framework for the Prompt Discovery of Compact Binary Mergers in Gravitational-wave Data. April 2016.

[71] Andrea Taracchini, Alessandra Buonanno, Yi Pan, Tanja Hinderer, Michael Boyle, et al. Effective-one-body model for black-hole binaries with generic mass ratios and spins. *Phys.Rev.*, D89:061502, 2014.

[72] Michael Pürrer. Frequency domain reduced order model of aligned-spin effective-one-body waveforms with generic mass-ratios and spins. 2015.

[73] Collin Capano, Thomas Dent, Yi-Ming Hu, Martin Hendry, Chris Messenger, and John Veitch. Systematic errors in estimation of gravitational-wave candidate significance. 2016.

[74] 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. *Phys.Rev.*, D85:122006, 2012.

[75] Kipp Cannon, Adrian Chapman, Chad Hanna, Drew Keppel, Antony C. Searle, et al. Singular value decomposition applied to compact binary coalescence gravitational-wave signals. *Phys.Rev.*, D82:044025, 2010.

[76] B. P. Abbott et al. Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914. 2016. arXiv:1602.03844.

[77] Bruce Allen. A chi**2 time-frequency discriminator for gravitational wave detection. *Phys.Rev.*, D71:062001, 2005.

[78] Kipp Cannon, Chad Hanna, and Jacob Peoples. Likelihood-Ratio Ranking Statistic for Compact Binary Coalescence Candidates with Rate Estimation. 2015.

[79] B. P. Abbott et al. GW150914: The Advanced LIGO Detectors in the Era of First Discoveries. *Phys. Rev. Lett.*, 116(13):131103, 2016.

[80] J. Aasi et al. Advanced LIGO. *Class. Quant. Grav.*, 32:074001, 2015.

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

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

[83] Luc Blanchet, Thibault Damour, Bala R. Iyer, Clifford M. Will, and Alan.G. Wiseman. Gravitational radiation damping of compact binary systems to second postNewtonian order. *Phys. Rev. Lett.*, 74:3515–3518, 1995.

[84] Lawrence E. Kidder, Clifford M. Will, and Alan G. Wiseman. Spin effects in the inspiral of coalescing compact binaries. *Phys.Rev.*, D47:4183–4187, 1993.

[85] Lawrence E. Kidder. Coalescing binary systems of compact objects to post-Newtonian 5/2 order. 5. Spin effects. *Phys.Rev.*, D52:821–847, 1995.

[86] 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.

[87] Krzysztof Belczynski, Alessandra Buonanno, Matteo Cantiello, Daniel E. Holz, Chris L. Fryer, et al. The Formation and Gravitational-Wave Detection of Massive Stellar Black-Hole Binaries. 2014.

[88] Marta Burgay et al. An increased estimate of the merger rate of double neutron stars from observations of a highly relativistic system. *Nature.*, 426:531–533, 2003.

[89] D. R. Lorimer. Binary and Millisecond Pulsars. *Living Rev. Rel.*, 11:8, 2008.

[90] Jeffrey E. McClintock, Ramesh Narayan, and James F. Steiner. Black Hole Spin via Continuum Fitting and the Role of Spin in Powering Transient Jets. *Space Sci. Rev.*, pages 1–28, July 2013.

[91] Benjamin J. Owen. Search templates for gravitational waves from inspiraling binaries: Choice of template spacing. *Phys.Rev.*, D53:6749–6761, 1996.

[92] Benjamin J. Owen and B.S. Sathyaprakash. Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement. *Phys.Rev.*, D60:022002, 1999.

[93] S. Babak, R. Balasubramanian, D. Churches, T. Cokelaer, and B.S. Sathyaprakash. A Template bank to search for gravitational waves from inspiralling compact binaries. I. Physical models. *Class.Quant.Grav.*, 23:5477–5504, 2006.

[94] Thomas Cokelaer. Gravitational waves from inspiralling compact binaries: Hexagonal template placement and its efficiency in detecting physical signals. *Phys.Rev.*, D76:102004, 2007.

[95] Drew Keppel. Metrics for multi-detector template placement in searches for short-duration nonprecessing inspiral gravitational-wave signals. 2013.

[96] Ian W. Harry, Bruce Allen, and B.S. Sathyaprakash. A Stochastic template placement algorithm for gravitational wave data analysis. *Phys. Rev. D*, 80:104014, 2009.

[97] Duncan A. Brown, Ian Harry, Andrew Lundgren, and Alexander H. Nitz. Detecting binary neutron star systems with spin in advanced gravitational-wave detectors. *Phys.Rev.*, D86:084017, 2012.

[98] Theocharis A. Apostolatos. Construction of a template family for the detection of gravitational waves from coalescing binaries. *Phys. Rev.*, D54:2421–2437, 1996.

[99] B. et al. Abbott. Search for gravitational waves from galactic and extra-galactic binary neutron stars. *Phys. Rev. D*, 72:082001, Oct 2005.

[100] A. Effler, R. M. S. Schofield, V. V. Frolov, G. Gonzlez, K. Kawabe, J. R. Smith, J. Birch, and R. McCarthy. Environmental Influences on the LIGO Gravitational Wave Detectors during the 6th Science Run. *Class. Quant. Grav.*, 32(3):035017, 2015.

[101] Kipp Cannon, Adrian Chapman, Chad Hanna, Drew Keppel, Antony C. Searle, and Alan J. Weinstein. Singular value decomposition applied to compact binary coalescence gravitational-wave signals. *Phys. Rev. D*, 82(4):044025, Aug 2010.

[102] Serge Droz, Daniel J. Knapp, Eric Poisson, and Benjamin J. Owen. Gravitational waves from inspiraling compact binaries: Validity of the stationary phase approximation to the Fourier transform. *Phys. Rev.*, D59:124016, 1999.

[103] Luc Blanchet, Thibault Damour, Gilles Esposito-Farese, and Bala R. Iyer. Gravitational radiation from inspiralling compact binaries completed at the third post-Newtonian order. *Phys. Rev. Lett.*, 93:091101, 2004.

[104] Michael Pürrer. Frequency domain reduced order models for gravitational waves from aligned-spin compact binaries. *Class. Quantum Grav.*, 31(19):195010, 2014.

[105] Louis Lyons. Open statistical issues in Particle Physics. *Ann. Appl. Stat.*, 2:887–915, 2008.

[106] Duncan A. Brown et al. Searching for Gravitational Waves from Binary Inspirals with LIGO. *Class. Quantum Grav.*, 21:S1625–S1633, 2004.

[107] Duncan A Brown. Using the INSPIRAL program to search for gravitational waves from low-mass binary inspiral. *Class. Quant. Grav.*, 22:S1097–S1108, 2005.

[108] S. Babak, R. Biswas, P.R. Brady, D.A. Brown, K. Cannon, et al. Searching for gravitational waves from binary coalescence. *Phys.Rev.*, D87:024033, 2013.

[109] D. A. Brown et al. A case study on the use of workflow technologies for scientific analysis: Gravitational wave data analysis. In Ian J. Taylor, Ewa Deelman, Dennis Gannon, and Matthew S. Shields, editors, *Workflows for e-Science*, chapter 5, pages 41–61. Springer-Verlag, 2006.

[110] Ewa Deelman, Gurmeet Singh, Mei-Hui Su, James Blythe, Yolanda Gil, Carl Kesselman, Gaurang Mehta, Karan Vahi, G. Bruce Berriman, John Good, Anastasia Laity, Joseph C. Jacob, and Daniel S. Katz. Pegasus: a framework

for mapping complex scientific workflows onto distributed systems. *Scientific Programming Journal*, 13(3):219–237, 2005.

[111] Ewa Deelman, Karan Vahi, Gideon Juve, Mats Rynge, Scott Callaghan, Philip J Maechling, Rajiv Mayani, Weiwei Chen, Rafael Ferreira da Silva, Miron Livny, and Kent Wenger. Pegasus: a workflow management system for science automation. *Future Generation Computer Systems*, 46:17–35, 2015. Funding Acknowledgements: NSF ACI SDCI 0722019, NSF ACI SI2-SSI 1148515 and NSF OCI-1053575.

[112] Douglas Thain, Todd Tannenbaum, and Miron Livny. Distributed computing in practice: the condor experience. *Concurrency - Practice and Experience*, 17(2-4):323–356, 2005.

[113] Peter Couvares, Tevfik Kosar, Alain Roy, Jeff Weber, and Kent Wenger. Workflow management in condor. In Ian J. Taylor, Ewa Deelman, Dennis Gannon, and Matthew S. Shields, editors, *Workflows for e-Science*, chapter 22, pages 357–375. Springer-Verlag, 2006.

[114] B. Abbott et al. Analysis of LIGO data for gravitational waves from binary neutron stars. *Phys.Rev.*, D69:122001, 2004.

[115] B. Abbott et al. Search for gravitational waves from galactic and extra-galactic binary neutron stars. *Phys.Rev.*, D72:082001, 2005.

[116] B. Abbott et al. Search for gravitational waves from primordial black hole binary coalescences in the galactic halo. *Phys. Rev.*, D72:082002, 2005.

[117] B. Abbott et al. Search for gravitational waves from binary black hole inspirals in LIGO data. *Phys.Rev.*, D73:062001, 2006.

[118] B. Abbott et al. Search for gravitational waves from binary inspirals in S3 and S4 LIGO data. *Phys.Rev.*, D77:062002, 2008.

[119] B. Abbott et al. Search of S3 LIGO data for gravitational wave signals from spinning black hole and neutron star binary inspirals. *Phys.Rev.*, D78:042002, 2008.

[120] B.P. Abbott et al. Search for Gravitational Waves from Low Mass Binary Coalescences in the First Year of LIGO's S5 Data. *Phys.Rev.*, D79:122001, 2009.

[121] B.P. Abbott et al. Search for Gravitational Waves from Low Mass Compact Binary Coalescence in 186 Days of LIGO's fifth Science Run. *Phys.Rev.*, D80:047101, 2009.

[122] J. Abadie et al. Search for Gravitational Waves from Compact Binary Coalescence in LIGO and Virgo Data from S5 and VSR1. *Phys. Rev.*, D82:102001, 2010.

[123] J. Abadie et al. Search for Gravitational Waves from Low Mass Compact Binary Coalescence in LIGO's Sixth Science Run and Virgo's Science Runs 2 and 3. *Phys.Rev.*, D85:082002, 2012.

[124] J. Aasi et al. Search for gravitational waves from binary black hole inspiral, merger, and ringdown in LIGO-Virgo data from 20092010. *Phys. Rev.*, D87(2):022002, 2013.

[125] J. Aasi et al. Search for gravitational wave ringdowns from perturbed intermediate mass black holes in LIGO-Virgo data from 20052010. *Phys. Rev.*, D89(10):102006, 2014.

[126] *Intel Math Kernel Library.* Intel Corporation, 2015.

[127] Matteo Frigo and Steven G. Johnson. The design and implementation of FFTW3. *Proceedings of the IEEE*, 93(2):216–231, 2005. Special issue on "Program Generation, Optimization, and Platform Adaptation".

[128] C Cutler and E Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform? *Phys. Rev. D*, 49:2658, 1994.

[129] F. Robinet. Omicron: An Algorithm to Detect and Characterize Transient Noise in Gravitational-Wave Detectors. *https://tds.ego-gw.it/ql/?c=10651*, 2015.

[130] S. Babak, H. Grote, M. Hewitson, H. Luck, and K. A. Strain. Signal based vetoes for the detection of gravitational waves from inspiralling compact binaries. *Phys. Rev. D*, 72:022002, 2005.

[131] Alexander Nitz. *THE EFFECT OF COMPACT OBJECT SPIN ON THE SEARCH FOR GRAVITATIONAL WAVES FROM BINARY NEUTRON STAR AND NEUTRON STAR-BLACK HOLE MERGERS*. PhD thesis, Syracuse U.

[132] Yi Pan, Alessandra Buonanno, Luisa T. Buchman, Tony Chu, Lawrence E. Kidder, Harald P. Pfeiffer, and Mark A. Scheel. Effective-one-body waveforms calibrated to numerical relativity simulations: coalescence of non-precessing, spinning, equal-mass black holes. *Phys. Rev.*, D81:084041, 2010.

[133] Mark Hannam, Patricia Schmidt, Alejandro Bohé, Leila Haegel, Sascha Husa, et al. Twist and shout: A simple model of complete precessing black-hole-binary gravitational waveforms. 2013.

[134] Sebastian Khan, Sascha Husa, Mark Hannam, Frank Ohme, Michael Prrer, Xisco Jimnez Forteza, and Alejandro Boh. Frequency-domain gravitational waves from non-precessing black-hole binaries. II. A phenomenological model for the advanced detector era. 2015.

[135] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016.

[136] B. P. Abbott et al. GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 116(24):241103, 2016.

[137] Clifford E. Rhoades, Jr. and Remo Ruffini. Maximum mass of a neutron star. *Phys. Rev. Lett.*, 32:324–327, 1974.

[138] B. P. Abbott et al. Properties of the binary black hole merger GW150914. 2016.

[139] B. P. Abbott et al. The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914. 2016.

[140] B. P. Abbott et al. Astrophysical Implications of the Binary Black-Hole Merger GW150914. *Astrophys. J.*, 818(2):L22, 2016.

[141] B. P. Abbott et al. Tests of general relativity with GW150914. 2016.

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

[143] F Acernese et al. Virgo status. *Classical and Quantum Gravity*, 25(18):184001, 2008.

[144] Atnf pulsar catalogue. http://www.atnf.csiro.au/people/pulsar/psrcat/. Lists 2536 pulsars as of 2016-05-31.

[145] R N Manchester, G B Hobbs, A Teoh, and M Hobbs. The Australia Telescope National Facility pulsar catalogue. *Astron. J.*, 129:1993, 2005.

[146] Observed neutron star masses. https://stellarcollapse.org/nsmasses. Accessed: 2016-04-11.

[147] James M. Lattimer. The nuclear equation of state and neutron star masses. *Ann. Rev. Nucl. Part. Sci.*, 62:485–515, 2012.

[148] Feryal Ozel and Paulo Freire. Masses, Radii, and Equation of State of Neutron Stars. *Annual Review of Astronomy and Astrophysics*, 54(1):null, 2016.

[149] M. Falanga, E. Bozzo, A. Lutovinov, J. M. Bonnet-Bidaud, Y. Fetisova, and J. Puls. Ephemeris, orbital decay, and masses of ten eclipsing high-mass X-ray binaries. *Astron. Astrophys.*, 577:A130, 2015.

[150] Paulo Cesar Carvalho Freire, Scott M. Ransom, Steve Begin, Ingrid H. Stairs, Jason W. T. Hessels, Lucille H. Frey, and Fernando Camilo. Eight New Millisecond Pulsars in NGC 6440 and NGC 6441. *Astrophys. J.*, 675:670, 2008.

[151] John Antoniadis et al. A Massive Pulsar in a Compact Relativistic Binary. *Science*, 340:6131, 2013.

[152] Vassiliki Kalogera and Gordon Baym. The maximum mass of a neutron star. *Astrophys. J.*, 470:L61–L64, 1996.

[153] Feryal Ozel, Dimitrios Psaltis, Ramesh Narayan, and Antonio Santos Villarreal. On the Mass Distribution and Birth Masses of Neutron Stars. *Astrophys.J.*, 757:55, 2012.

[154] Blent Kiziltan, Athanasios Kottas, Maria De Yoreo, and Stephen E. Thorsett. The Neutron Star Mass Distribution. *Astrophys. J.*, 778:66, 2013.

[155] Bulent Kiziltan, Athanasios Kottas, and Stephen E. Thorsett. The Neutron Star Mass Distribution. *Astrophys.J.*, 778:66, 2013.

[156] Jason W.T. Hessels, Scott M. Ransom, Ingrid H. Stairs, Paulo Cesar Carvalho Freire, Victoria M. Kaspi, et al. A radio pulsar spinning at 716-hz. *Science*, 311:1901–1904, 2006.

[157] M. Kramer and N. Wex. The double pulsar system: A unique laboratory for gravity. *Class.Quant.Grav.*, 26:073001, 2009.

[158] Ryan S. Lynch, Paulo C. C. Freire, Scott M. Ransom, and Bryan A. Jacoby. The Timing of Nine Globular Cluster Pulsars. *Astrophys. J.*, 745:109, 2012.

[159] Alexander Harvey Nitz. *THE EFFECT OF COMPACT OBJECT SPIN ON THE SEARCH FOR GRAVITATIONAL WAVES FROM BINARY NEUTRON STAR AND NEUTRON STAR-BLACK HOLE MERGERS*. PhD thesis, Syracuse U.

[160] Jonathan Grindlay, Simon Portegies Zwart, and Stephen McMillan. Short gamma-ray bursts from binary neutron star mergers in globular clusters. *Nature Physics*, 2:116–119, 2006.

[161] Aleksander Sadowski, Krzysztof Belczynski, Tomasz Bulik, Natalia Ivanova, Frederic A. Rasio, et al. The Total Merger Rate of Compact Object Binaries In The Local Universe. *Astrophys.J.*, 2007.

[162] William H. Lee, Enrico Ramirez-Ruiz, and Glenn van de Ven. Short gamma-ray bursts from dynamically-assembled compact binaries in globular clusters: pathways, rates, hydrodynamics and cosmological setting. *Astrophys.J.*, 720:953–975, 2010.

[163] Matthew J. Benacquista and Jonathan M.B. Downing. Relativistic Binaries in Globular Clusters. *Living Rev.Rel.*, 16:4, 2013.

[164] Krzysztof Belczynski, Tomasz Bulik, Ilya Mandel, B.S. Sathyaprakash, Andrzej A. Zdziarski, et al. Cyg X-3: a Galactic double black hole or black hole-neutron star progenitor. *Astrophys.J.*, 764:96, 2013.

[165] Will M. Farr, Niharika Sravan, Andrew Cantrell, Laura Kreidberg, Charles D. Bailyn, et al. The Mass Distribution of Stellar-Mass Black Holes. *Astrophys.J.*, 741:103, 2011.

[166] Feryal Ozel, Dimitrios Psaltis, Ramesh Narayan, and Jeffrey E. McClintock. The Black Hole Mass Distribution in the Galaxy. *Astrophys.J.*, 725:1918–1927, 2010.

[167] Andrea Merloni. Cosmological evolution of supermassive black holes and AGN: a synthesis model for accretion and feedback. *Mem. Soc. Astron. Italiana*, 79:1310, 2008.

[168] Grzegorz Wiktorowicz, Krzysztof Belczynski, and Thomas J. Maccarone. Black Hole X-ray Transients: The Formation Puzzle. 2013.

[169] K. Belczynski, G. Wiktorowicz, C. Fryer, D. Holz, and V. Kalogera. Missing Black Holes Unveil The Supernova Explosion Mechanism. *Astrophys.J.*, 757:91, 2012.

[170] Chris L. Fryer, Krzysztof Belczynski, Grzegorz Wiktorowicz, Michal Dominik, Vicky Kalogera, et al. Compact Remnant Mass Function: Dependence on the Explosion Mechanism and Metallicity. *Astrophys.J.*, 749:91, 2012.

[171] Richard W. O'Shaughnessy, J. Kaplan, V. Kalogera, and K. Belczynski. Bounds on expected black hole spins in inspiraling binaries. *Astrophys. J.*, 632:1035–1041, 2005.

[172] Krzysztof Belczynski, Tomasz Bulik, Chris L. Fryer, Ashley Ruiter, Jorick S. Vink, et al. On The Maximum Mass of Stellar Black Holes. *Astrophys.J.*, 714:1217–1226, 2010.

[173] Michal Dominik, Krzysztof Belczynski, Christopher Fryer, Daniel Holz, Emanuele Berti, et al. Double Compact Objects I: The Significance of the Common Envelope on Merger Rates. *Astrophys.J.*, 759:52, 2012.

[174] J.M. Miller, C.S. Reynolds, A.C. Fabian, G. Miniutti, and L.C. Gallo. Stellar-mass Black Hole Spin Constraints from Disk Reflection and Continuum Modeling. *Astrophys.J.*, 697:900–912, 2009.

[175] Rebecca Shafee, Jeffrey E. McClintock, Ramesh Narayan, Shane W. Davis, Li-Xin Li, et al. Estimating the spin of stellar-mass black holes via spectral fitting of the x-ray continuum. *Astrophys.J.*, 636:L113–L116, 2006.

[176] Jeffrey E. McClintock, Rebecca Shafee, Ramesh Narayan, Ronald A. Remillard, Shane W. Davis, et al. The Spin of the Near-Extreme Kerr Black Hole GRS 1915+105. *Astrophys.J.*, 652:518–539, 2006.

[177] Jifeng Liu, Jeffery McClintock, Ramesh Narayan, Shane Davis, and Jerome Orosz. Precise Measurement of the Spin Parameter of the Stellar-Mass Black Hole M33 X-7. *Astrophys.J.*, 679:L37–L40, 2008.

[178] Lijun Gou, Jeffrey E. McClintock, Jifeng Liu, Ramesh Narayan, James F. Steiner, et al. A Determination of the Spin of the Black Hole Primary in LMC X-1. *Astrophys.J.*, 701:1076–1090, 2009.

[179] Shane W. Davis, Chris Done, and Omer M. Blaes. Testing accretion disk theory in black hole x-ray binaries. *Astrophys.J.*, 647:525–538, 2006.

[180] Li-Xin Li, Erik R. Zimmerman, Ramesh Narayan, and Jeffrey E. McClintock. Multi-temperature blackbody spectrum of a thin accretion disk around a Kerr black hole: Model computations and comparison with observations. *Astrophys.J.Suppl.*, 157:335–370, 2005.

[181] A. C. Fabian, D. R. Wilkins, J. M. Miller, R. C. Reis, C. S. Reynolds, E. M. Cackett, M. A. Nowak, G. G. Pooley, K. Pottschmidt, J. S. Sanders, R. R. Ross, and J. Wilms. On the determination of the spin of the black hole in Cyg X-1 from X-ray reflection spectra. *Mon. Not. Roy. Astron. Soc.*, 424:217–223, July 2012.

[182] Lijun Gou, Jeffrey E. McClintock, Mark J. Reid, Jerome A. Orosz, James F. Steiner, Ramesh Narayan, Jingen Xiang, Ronald A. Remillard, Keith A. Arnaud, and Shane W. Davis. The Extreme Spin of the Black Hole in Cygnus X-1. *Astrophys. J.*, 742:85, 2011.

[183] Jeffrey E. McClintock, Ramesh Narayan, Shane W. Davis, Lijun Gou, Akshay Kulkarni, et al. Measuring the Spins of Accreting Black Holes. *Class.Quant.Grav.*, 28:114009, 2011.

[184] Charles W Misner, Kip S Thorne, and John Archibald Wheeler. *Gravitation*. Princeton University Press, 2017.

[185] James F. Steiner and Jeffrey E. McClintock. Modeling the Jet Kinematics of the Black Hole Microquasar XTE J1550-564: A Constraint on Spin-Orbit Alignment. *Astrophys.J.*, 745:136, 2012.

[186] T. Fragos, M. Tremmel, E. Rantsiou, and K. Belczynski. Black Hole Spin-Orbit Misalignment in Galactic X-ray Binaries. *ApJ*, 719:L79–L83, August 2010.

[187] Cody Messick et al. Analysis Framework for the Prompt Discovery of Compact Binary Mergers in Gravitational-wave Data. 2016.

[188] Ian Harry, Stephen Privitera, Alejandro Boh, and Alessandra Buonanno. Searching for Gravitational Waves from Compact Binaries with Precessing Spins. 2016.

[189] Tito Dal Canton, Andrew P Lundgren, and Alex B Nielsen. Impact of precession on aligned-spin searches for neutron-star–black-hole binaries. *Physical Review D*, 91(6):062010, 2015.

[190] L A Wainstein and V D Zubakov. *Extraction of Signals from Noise*. Prentice-Hall, Englewood Cliffs, 1962.

[191] David Eichler, Mario Livio, Tsvi Piran, and David N. Schramm. Nucleosynthesis, Neutrino Bursts and Gamma-Rays from Coalescing Neutron Stars. *Nature*, 340:126–128, 1989.

[192] Brad M. S. Hansen and Maxim Lyutikov. Radio and x-ray signatures of merging neutron stars. *Mon. Not. Roy. Astron. Soc.*, 322:695, 2001.

[193] Ramesh Narayan, Bohdan Paczynski, and Tsvi Piran. Gamma-ray bursts as the death throes of massive binary stars. *Astrophys. J.*, 395:L83–L86, 1992.

[194] Li-Xin Li and Bohdan Paczynski. Transient events from neutron star mergers. *Astrophys. J.*, 507:L59, 1998.

[195] Ehud Nakar. Short-Hard Gamma-Ray Bursts. *Phys. Rept.*, 442:166–236, 2007.

[196] B.D. Metzger and E. Berger. What is the Most Promising Electromagnetic Counterpart of a Neutron Star Binary Merger? *Astrophys.J.*, 746:48, 2012.

[197] Ehud Nakar and Tsvi Piran. Radio Remnants of Compact Binary Mergers - the Electromagnetic Signal that will follow the Gravitational Waves. *Nature*, 478:82–84, 2011.

[198] Edo Berger. Short-Duration Gamma-Ray Bursts. *Ann. Rev. Astron. Astrophys.*, 52:43–105, 2014.

[199] Bing Zhang. A possible connection between Fast Radio Bursts and Gamma-Ray Bursts. *Astrophys. J. Lett.*, 780:L21, 2014.

[200] B. Abbott et al. A Search for gravitational waves associated with the gamma ray burst GRB030329 using the LIGO detectors. *Phys. Rev.*, D72:042002, 2005.

[201] B. Abbott et al. Implications for the Origin of GRB 070201 from LIGO Observations. *Astrophys. J.*, 681:1419–1428, 2008.

[202] J. Abadie et al. Implications For The Origin Of GRB 051103 From LIGO Observations. *Astrophys. J.*, 755:2, 2012.

[203] J. Abadie et al. Search for gravitational waves associated with gamma-ray bursts during LIGO science run 6 and Virgo science runs 2 and 3. *Astrophys. J.*, 760:12, 2012.

[204] B. P. Abbott et al. Localization and broadband follow-up of the gravitational-wave transient GW150914. *Submitted to: Astrophys. J. Lett.*, 2016.

[205] B. P. Abbott et al. Calibration of the Advanced LIGO detectors for the discovery of the binary black-hole merger GW150914. 2016.

[206] B. P. Abbott et al. Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914. 2016.

[207] Collin Capano, Ian Harry, Stephen Privitera, and Alessandra Buonanno. Implementing a search for gravitational waves from non-precessing, spinning binary black holes. 2016.

[208] K.G. Arun, Alessandra Buonanno, Guillaume Faye, and Evan Ochsner. Higher-order spin effects in the amplitude and phase of gravitational waveforms emitted by inspiraling compact binaries: Ready-to-use gravitational waveforms. *Phys. Rev. D*, 79:104023, 2009.

[209] Alejandro Boh, Sylvain Marsat, and Luc Blanchet. Next-to-next-to-leading order spinorbit effects in the gravitational wave flux and orbital phasing of compact binaries. *Class. Quant. Grav.*, 30:135009, 2013.

[210] Luc Blanchet. Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries. *Living Rev.Rel.*, 17:2, 2014.

[211] Alejandro Boh, Guillaume Faye, Sylvain Marsat, and Edward K. Porter. Quadratic-in-spin effects in the orbital dynamics and gravitational-wave energy flux of compact binaries at the 3PN order. *Class. Quant. Grav.*, 32(19):195010, 2015.

[212] Chandra Kant Mishra, Aditya Kela, KG Arun, and Guillaume Faye. Ready-to-use post-newtonian gravitational waveforms for binary black holes with non-precessing spins: An update. *Physical Review D*, 93(8):084054, 2016.

[213] D. V. Martynov et al. Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy. D93(11):112004, 2016.

[214] B. P. Abbott et al. GW150914: First results from the search for binary black hole coalescence with Advanced LIGO. 2016.

[215] Will M. Farr, Jonathan R. Gair, Ilya Mandel, and Curt Cutler. Counting And Confusion: Bayesian Rate Estimation With Multiple Populations. *Phys. Rev.*, D91(2):023005, 2015.

[216] Alessandra Buonanno, Bala Iyer, Evan Ochsner, Yi Pan, and B.S. Sathyaprakash. Comparison of post-Newtonian templates for compact binary inspiral signals in gravitational-wave detectors. *Phys.Rev.*, D80:084043, 2009.

[217] Yi Pan, Alessandra Buonanno, Andrea Taracchini, Lawrence E. Kidder, Abdul H. Mroue, et al. Inspiral-merger-ringdown waveforms of spinning, precessing black-hole binaries in the effective-one-body formalism. 2013.

[218] Rahul Biswas, 2, Patrick R. Brady, Jolien D. E. Creighton, and Stephen Fairhurst. The Loudest event statistic: General formulation, properties and applications. *Class. Quant. Grav.*, 26:175009, 2009. [Erratum: Class. Quant. Grav.30,079502(2013)].

[219] Stanislav Babak, Andrea Taracchini, and Alessandra Buonanno. Validating the effective-one-body model of spinning, precessing binary black holes against numerical relativity. *arXiv preprint arXiv:1607.05661*, 2016.

[220] K. Belczynski, S. Repetto, D. Holz, R. O'Shaughnessy, T. Bulik, E. Berti, C. Fryer, and M. Dominik. Comparison of LIGO/Virgo upper limits with predicted compact binary merger rates. 2015.

[221] Matthew D. Duez. Numerical relativity confronts compact neutron star binaries: a review and status report. *Class. Quant. Grav.*, 27:114002, 2010.

[222] Bruno Giacomazzo, Rosalba Perna, Luciano Rezzolla, Eleonora Troja, and Davide Lazzati. Compact Binary Progenitors of Short Gamma-Ray Bursts. *Astrophys. J.*, 762:L18, 2013.

[223] Francesco Pannarale, Emanuele Berti, Koutarou Kyutoku, Benjamin D. Lackey, and Masaru Shibata. Gravitational-wave cutoff frequencies of tidally disruptive neutron star-black hole binary mergers. *Phys. Rev.*, D92(8):081504, 2015.

[224] B. P. Abbott et al. Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817a. *The Astrophysical Journal*, 848(2):L13, oct 2017.

[225] M. Soares-Santos at 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*, 848(2):L16, oct 2017.

[226] D. Williams, J. A. Clark, A. R. Williamson, and I. S. Heng. Constraints on short, hard gamma-ray burst beaming angles from gravitational wave observations. *The Astrophysical Journal*, 858(2):79, may 2018.

[227] Hsin-Yu Chen, Salvatore Vitale, and Ramesh Narayan. Viewing angle of binary neutron star mergers. *Phys. Rev. X*, 9:031028, Aug 2019.

[228] J. M. Lattimer and D. N. Schramm. Black-hole-neutron-star collisions. *Astrophysical Journal, Letters*, 192:L145–L147, September 1974.

[229] Y. Z. Qian and G. J. Wasserburg. Where, oh where has the r-process gone? *Phys. Rept.*, 442:237–268, 2007.

[230] Andreas Bauswein, R. Ardevol Pulpillo, H. T. Janka, and S. Goriely. Nucleosynthesis constraints on the neutron star-black hole merger rate. *Astrophys. J.*, 795(1):L9, 2014.

[231] Hao Wang, Fu-Wen Zhang, Yuan-Zhu Wang, Zhao-Qiang Shen, Yun-Feng Liang, Xiang Li, Neng-Hui Liao, Zhi-Ping Jin, Qiang Yuan, Yuan-Chuan Zou, Yi-Zhong Fan, and Da-Ming Wei. The GW170817/GRB 170817a/AT 2017gfo association: Some implications for physics and astrophysics. *The Astrophysical Journal*, 851(1):L18, dec 2017.

[232] Daniel M. Siegel. GW170817–the first observed neutron star merger and its kilonova: implications for the astrophysical site of the r-process. *Eur. Phys. J.*, A55(11):203, 2019.

[233] Hsin-Yu Chen and Daniel E. Holz. Gamma-Ray-Burst Beaming and Gravitational-Wave Observations. *Phys. Rev. Lett.*, 111(18):181101, 2013.

[234] Ehud Nakar, Avishay Gal-Yam, and Derek B. Fox. The Local Rate and the Progenitor Lifetimes of Short-Hard Gamma-Ray Bursts: Synthesis and Predictions for LIGO. *Astrophys. J.*, 650:281–290, 2006.

[235] Derek B. Fox et al. The afterglow of grb050709 and the nature of the short-hard gamma-ray bursts. *Nature*, 437:845–850, 2005.

[236] Dirk Grupe, David N. Burrows, Sandeep K. Patel, Chryssa Kouveliotou, Bing Zhang, Peter Meszaros, Ralph A. M. Wijers, and Neil Gehrels. Jet breaks in short gamma-ray bursts. 1. the uncollimated afterglow of grb 050724. *Astrophys. J.*, 653:462–467, 2006.

[237] Alicia M. Soderberg et al. The afterglow and host galaxy of the energetic short-hard gamma-ray burst 051221. *Astrophys. J.*, 650:261–271, 2006.

[238] T. Sakamoto, E. Troja, K. Aoki, S. Guiriec, M. Im, G. Leloudas, D. Malesani, A. Melandri, A. de Ugarte Postigo, Y. Urata, D. Xu, P. D'Avanzo, J. Gorosabel, Y. Jeon, R. Sánchez-Ramírez, M. I. Andersen, J. Bai, S. D. Barthelmy, M. S. Briggs, S. Foley, A. S. Fruchter, J. P. U. Fynbo, N. Gehrels, K. Huang, M. Jang, N. Kawai, H. Korhonen, J. Mao, J. P. Norris, R. D. Preece, J. L. Racusin, C. C. Thöne, K. Vida, and X. Zhao. Identifying the Location in the Host Galaxy of the Short GRB 111117A with the Chandra Subarcsecond Position. *Astrophysical Journal*, 766:41, March 2013.

[239] R. Margutti, E. Berger, W. Fong, B. A. Zauderer, S. B. Cenko, J. Greiner, A. M. Soderberg, A. Cucchiara, A. Rossi, S. Klose, S. Schmidl, D. Milisavljevic, and N. Sanders. The Afterglow and Environment of the Short GRB 111117A. *Astrophysical Journal*, 756:63, September 2012.

[240] A. Nicuesa Guelbenzu, S. Klose, A. Rossi, D. A. Kann, T. Krühler, J. Greiner, A. Rau, F. Olivares E., P. M. J. Afonso, R. Filgas, A. Küpcü Yoldaş, S. McBreen, M. Nardini, P. Schady, S. Schmidl, A. C. Updike, and A. Yoldaş. GRB 090426: Discovery of a jet break in a short burst afterglow. *Astronomy and Astrophysics*, 531:L6, July 2011.

[241] Alexander Dietz. Estimation of Compact Binary Coalescense Rates from Short Gamma-Ray Burst Redshift Measurements. *Astron. Astrophys.*, 529:A97, 2011.

[242] T. Regimbau, K. Siellez, D. Meacher, B. Gendre, and M. Bor. Revisiting coincidence rate between Gravitational Wave detection and short Gamma-Ray Burst for the Advanced and third generation. *Astrophys. J.*, 799(1):69, 2015.

[243] J. Clark, H. Evans, S. Fairhurst, I. W. Harry, E. Macdonald, D. Macleod, P. J. Sutton, and A. R. Williamson. Prospects for joint gravitational wave and short gamma-ray burst observations. *Astrophys. J.*, 809(1):53, 2015.

[244] B. P. et al. Abbott. Gwtc-1: A gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs. *Phys. Rev. X*, 9:031040, Sep 2019.

[245] Fergus Hayes, Ik Siong Heng, John Veitch, and Daniel Williams. Comparing Short Gamma-Ray Burst Jet Structure Models. 2019.

[246] Benjamin P. Abbott et al. Upper Limits on the Rates of Binary Neutron Star and Neutron Starblack Hole Mergers From Advanced Ligos First Observing run. *Astrophys. J.*, 832(2):L21, 2016.

[247] G. Ashton, E. Burns, T. Dal Canton, T. Dent, H. B Eggenstein, A. B. Nielsen, R. Prix, M. Was, and S. J. Zhu. Coincident detection significance in multimessenger astronomy. *Astrophys. J.*, 860(1):6, 2018.

[248] E. Burns et al. A Fermi Gamma-ray Burst Monitor Search for Electromagnetic Signals Coincident with Gravitational-Wave Candidates in Advanced LIGO's First Observing Run. 2018.

[249] B P Abbott et al. Effects of data quality vetoes on a search for compact binary coalescences in Advanced LIGOs first observing run. *Class. Quant. Grav.*, 35(6):065010, 2018.

[250] Benjamin P. Abbott et al. Search for intermediate mass black hole binaries in the first observing run of Advanced LIGO. *Phys. Rev.*, D96(2):022001, 2017.

[251] Benjamin P. Abbott et al. GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2. *Phys. Rev. Lett.*, 118(22):221101, 2017. [Erratum: Phys. Rev. Lett.121,no.12,129901(2018)].

[252] B.. P.. Abbott et al. GW170608: Observation of a 19-solar-mass Binary Black Hole Coalescence. *Astrophys. J.*, 851(2):L35, 2017.

[253] B. P. Abbott et al. GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 119(14):141101, 2017.

[254] 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 signals. 2018.

[255] Etienne Racine. Analysis of spin precession in binary black hole systems including quadrupole-monopole interaction. *Phys. Rev.*, D78:044021, 2008.

[256] A. Taracchini, A. Buonanno, Y. Pan, T. Hinderer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, A. H. Mroué, H. P. Pfeiffer, M. A. Scheel, B. Szilágyi, N. W. Taylor, and A. Zenginoglu. Effective-one-body model for black-hole binaries with generic mass ratios and spins. *Phys. Rev. D*, 89(6):061502, March 2014.

[257] Alejandro Boh et al. An improved effective-one-body model of spinning, non-precessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors. 2016.

[258] Luc Blanchet. Gravitational radiation from post-Newtonian sources and inspiralling compact binaries. *Living Rev. Rel.*, 5:3, 2002.

[259] Guillaume Faye, Sylvain Marsat, Luc Blanchet, and Bala R. Iyer. The third and a half post-Newtonian gravitational wave quadrupole mode for quasi-circular inspiralling compact binaries. *Class. Quant. Grav.*, 29:175004, 2012.

[260] Alexander H. Nitz, Tito Dal Canton, Derek Davis, and Steven Reyes. Rapid detection of gravitational waves from compact binary mergers with PyCBC Live. *Phys. Rev.*, D98(2):024050, 2018.

[261] L. K. Nuttall et al. Improving the Data Quality of Advanced LIGO Based on Early Engineering Run Results. 2015.

[262] Benjamin P. Abbott et al. GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2. *Phys. Rev. Lett.*, 118(22):221101, 2017.

[263] Benjamin P. Abbott et al. GW170814: A three-detector observation of gravitational waves from a binary black hole coalescence. *Phys. Rev. Lett.*, 119:141101, 2017.

[264] C. Capano, T. Dent, C. Hanna, M. Hendry, C. Messenger, Y.-M. Hu, and J. Veitch. Systematic errors in estimation of gravitational-wave candidate significance. *Phys. Rev. D*, 96(8):082002, October 2017.

[265] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A Practical and powerful approach to multiple testing. *J. Roy. Statist. Soc.*, 57:289–300, 1995.

[266] F. Guglielmetti, R. Fischer, and V. Dose. Background-source separation in astronomical images with Bayesian probability theory - I. The method. *Mon. Not. Roy. Astr. Soc.*, 396:165–190, June 2009.

[267] W. M. Farr, J. R. Gair, I. Mandel, and C. Cutler. Counting and confusion: Bayesian rate estimation with multiple populations. *Phys. Rev. D*, 91(2):023005, January 2015.

[268] Laura Kreidberg, Charles D. Bailyn, Will M. Farr, and Vassiliki Kalogera. Mass Measurements of Black Holes in X-Ray Transients: Is There a Mass Gap? *Astrophys.J.*, 757:36, 2012.

[269] Christopher S. Reynolds. Measuring Black Hole Spin using X-ray Reflection Spectroscopy. *Space Sci. Rev.*, 183(1-4):277–294, 2014.

[270] K. Belczynski et al. The origin of low spin of black holes in LIGO/Virgo mergers. 2017.

[271] Daniel Wysocki, Davide Gerosa, Richard O'Shaughnessy, Krzysztof Belczynski, Wojciech Gladysz, Emanuele Berti, Michael Kesden, and Daniel E. Holz. Explaining LIGOs observations via isolated binary evolution with natal kicks. *Phys. Rev.*, D97(4):043014, 2018.

[272] Bernard F. Schutz. Networks of gravitational wave detectors and three figures of merit. *Class. Quant. Grav.*, 28:125023, 2011.

[273] Hsin-Yu Chen and Daniel E. Holz. The Loudest Gravitational Wave Events. 2014.

[274] Jasper Vivian Wall and Charles R Jenkins. *Practical statistics for astronomers*. Cambridge University Press, 2012.

[275] Michael P Hobson, Andrew H Jaffe, Andrew R Liddle, Pia Mukherjee, and David Parkinson. *Bayesian methods in cosmology*. Cambridge University Press, 2010.

[276] Campbell B Read and Brani Vidakovic. *Encyclopedia of statistical sciences*, volume 2. John Wiley & Sons, 2006.

[277] BP Abbott, R Abbott, TD Abbott, F Acernese, K Ackley, C Adams, T Adams, P Addesso, RX Adhikari, VB Adya, et al. Model comparison from ligo-virgo data on gw170817's binary components and consequences for the merger remnant. *arXiv preprint arXiv:1908.01012*, 2019.

[278] Robert E Kass and Adrian E Raftery. Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795, 1995.

[279] Walter Del Pozzo, Tjonnie GF Li, Michalis Agathos, Chris Van Den Broeck, and Salvatore Vitale. Demonstrating the feasibility of probing the neutron-star equation of state with second-generation gravitational-wave detectors. *Physical review letters*, 111(7):071101, 2013.

[280] Alexander Ly, Alexander Etz, Maarten Marsman, and Eric-Jan Wagenmakers. Replication bayes factors from evidence updating. *Behavior research methods*, pages 1–11.

[281] Coryn A.L. Bailer-Jones and Kester Smith. Combining probabilities. *GAIA-C8-TN-MPIA-CBJ-053*, 2011.

[282] Paul Bromiley. Products and convolutions of gaussian probability density functions. *Tina-Vision Memo*, 3(4):1, 2003.

[283] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman. emcee: The MCMC Hammer. *Publ. Astron. Soc. Pac.*, 125:306, March 2013.

[284] W. D. Vousden, W. M. Farr, and I. Mandel. Dynamic temperature selection for parallel tempering in Markov chain Monte Carlo simulations. *Monthly Notices of the Royal Astronomical Society*, 455(2):1919–1937, 2016.

[285] Jian-Sheng Wang and Robert H. Swendsen. Replica monte carlo simulation (revisited). *Progress of Theoretical Physics Supplement*, 157:317–323, 2005.

[286] David J. Earl and Michael W. Deem. Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.*, 7:3910–3916, 2005.

[287] Jeffrey Annis, Nathan J Evans, Brent J Miller, and Thomas J Palmeri. Thermodynamic integration and steppingstone sampling methods for estimating bayes factors: A tutorial. *Journal of mathematical psychology*, 89:67–86, 2019.

[288] David Carlson, Patrick Stinson, Ari Pakman, and Liam Paninski. Partition functions from rao-blackwellized tempered sampling. In *International Conference on Machine Learning*, pages 2896–2905, 2016.

[289] IS Gradshteyn and IM Ryzhik. Table of integrals, series, and products, 8th edn., ed. by d. zwillinger, 2015.

[290] Nial Friel, Merrilee Hurn, and Jason Wyse. Improving power posterior estimation of statistical evidence. *Statistics and Computing*, 24(5):709–723, 2014.

[291] Colin H LaMont and Paul A Wiggins. Correspondence between thermodynamics and inference. *Physical Review E*, 99(5):052140, 2019.

[292] Peigui Liu, Ahmed S Elshall, Ming Ye, Peter Beerli, Xiankui Zeng, Dan Lu, and Yuezan Tao. Evaluating marginal likelihood with thermodynamic integration method and comparison with several other numerical methods. *Water Resources Research*, 52(2):734–758, 2016.

[293] Anita de Ruiter and Chris Oostenbrink. Free energy calculations of protein–ligand interactions. *Current opinion in chemical biology*, 15(4):547–552, 2011.

[294] Anita de Ruiter, Stefan Boresch, and Chris Oostenbrink. Comparison of thermodynamic integration and bennett acceptance ratio for calculating relative protein-ligand binding free energies. *Journal of computational chemistry*, 34(12):1024–1034, 2013.

[295] Chris J Oates, Mark Girolami, and Nicolas Chopin. Control functionals for monte carlo integration. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(3):695–718, 2017.

[296] Nathan J Evans and Jeffrey Annis. Thermodynamic integration via differential evolution: A method for estimating marginal likelihoods. *Behavior research methods*, 51(2):930–947, 2019.

[297] Viggo Brun. A generalization of the formula of simpson for non-equidistant ordinates. *Nordisk Matematisk Tidskrift*, 1(1):10–15, 1953.

[298] Ernts S. Selmer. Numerical integration by non-equidistant ordinates. *Nordisk Matematisk Tidskrift*, pages 97–108, 1958.

[299] Said M Easa. Area of irregular region with unequal intervals. *Journal of Surveying Engineering*, 114(2):50–58, 1988.

[300] Daniel W Chambers. Estimating pit excavation volume using unequal intervals. *Journal of Surveying Engineering*, 115(4):390–401, 1989.

[301] James F Epperson. On the runge example. *The American Mathematical Monthly*, 94(4):329–341, 1987.

[302] William H Press and Saul A Teukolsky. Padé approximants. *Computers in Physics*, 6(1):82–83, 1992.

[303] CM Biwer, Collin D Capano, Soumi De, Miriam Cabero, Duncan A Brown, Alexander H Nitz, and Vivien Raymond. Pycbc inference: A python-based parameter estimation toolkit for compact binary coalescence signals. *Publications of the Astronomical Society of the Pacific*, 131(996):024503, 2019.

[304] Patricio Maturana Russel, Renate Meyer, John Veitch, and Nelson Christensen. The stepping-stone sampling algorithm for calculating the evidence of gravitational wave models. 2018.

[305] Ben Calderhead and Mark Girolami. Estimating bayes factors via thermodynamic integration and population mcmc. *Computational Statistics & Data Analysis*, 53(12):4028–4045, 2009.

[306] Persi Diaconis. Bayesian numerical analysis. *Statistical decision theory and related topics IV*, 1:163–175, 1988.

[307] François-Xavier Briol, Chris J Oates, Mark Girolami, Michael A Osborne, and Dino Sejdinovic. Probabilistic integration: A role for statisticians in numerical analysis. *arXiv preprint arXiv:1512.00933*, 1293, 2015.

[308] Michael A Newton and Adrian E Raftery. Approximate bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society: Series B (Methodological)*, 56(1):3–26, 1994.

[309] Gary W Oehlert. A note on the delta method. *The American Statistician*, 46(1):27–29, 1992.

[310] Eric-Jan Wagenmakers, Tom Lodewyckx, Himanshu Kuriyal, and Raoul Grasman. Bayesian hypothesis testing for psychologists: A tutorial on the savage–dickey method. *Cognitive psychology*, 60(3):158–189, 2010.

[311] Isabella Verdinelli and Larry Wasserman. Computing bayes factors using a generalization of the savage-dickey density ratio. *Journal of the American Statistical Association*, 90(430):614–618, 1995.

[312] David W Scott. On optimal and data-based histograms. *Biometrika*, 66(3):605–610, 1979.

[313] David Freedman and Persi Diaconis. On the histogram as a density estimator:l2 theory. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57(4):453–476, Dec 1981.

[314] Charles J Stone, Mark H Hansen, Charles Kooperberg, Young K Truong, et al. Polynomial splines and their tensor products in extended linear modeling: 1994 wald memorial lecture. *The Annals of Statistics*, 25(4):1371–1470, 1997.

[315] Hirotugu Akaike. Likelihood of a model and information criteria. *Journal of econometrics*, 16(1):3–14, 1981.

[316] Gideon Schwarz et al. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.

[317] Henry Deng and Hadley Wickham. Density estimation in r. 2011.

[318] Bradley Efron. Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer, 1992.

[319] Bradley Efron and Robert Tibshirani. Improvements on cross-validation: the 632+ bootstrap method. *Journal of the American Statistical Association*, 92(438):548–560, 1997.

[320] Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and computing*, 27(5):1413–1432, 2017.

[321] K. S. Thorne. THREE HUNDRED YEARS OF GRAVITATION. 1987.

[322] H. Wahlquist. The Doppler Response to Gravitational Waves From a Binary Star Source. *Gen. Rel. Grav.*, 19:1101–1113, 1987.

[323] Nelson Christensen and Renate Meyer. Using Markov chain Monte Carlo methods for estimating parameters with gravitational radiation data. *Phys.Rev.*, D64:022001, 2001.

[324] Mark Hannam, Duncan A. Brown, Stephen Fairhurst, Chris L. Fryer, and Ian W. Harry. When can gravitational-wave observations distinguish between black holes and neutron stars? *Astrophys.J.*, 766:L14, 2013.

[325] LIGO Scientific Collaboration. LSC Algorithm Library.

[326] Christopher Michael Biwer. *The detection and parameter estimation of binary black hole mergers* . PhD thesis, Syracuse U.

[327] Tejaswi Venumadhav, Aaron Zimmerman, and Christopher M. Hirata. The stability of tidally deformed neutron stars to three- and four-mode coupling. *Astrophys. J.*, 781:23, 2014.

[328] Reed Essick, Salvatore Vitale, and Nevin N. Weinberg. Impact of the tidal p-g instability on the gravitational wave signal from coalescing binary neutron stars. *Phys. Rev.*, D94(10):103012, 2016.

[329] T. A. Apostolatos. Search templates for gravitational waves from precessing, inspiraling binaries. *Phys. Rev.*, D52:605–620, 1995.

[330] Walter Del Pozzo, Tjonnie G. F. Li, Michalis Agathos, Chris Van Den Broeck, and Salvatore Vitale. Demonstrating the feasibility of probing the neutron star equation of state with second-generation gravitational wave detectors. *Phys.Rev.Lett.*, 111(7):071101, 2013.

[331] Lee Lindblom, Benjamin J. Owen, and Duncan A. Brown. Model Waveform Accuracy Standards for Gravitational Wave Data Analysis. *Phys. Rev.*, D78:124020, 2008.

[332] Yixiao Zhou and Fan Zhang. Equation of State Dependence of Nonlinear Mode-tide Coupling in Coalescing Binary Neutron Stars. 2018. [Astrophys. J.849,114(2017)].

[333] N. Andersson and W. C. G. Ho. Using gravitational-wave data to constrain dynamical tides in neutron star binaries. *Phys. Rev.*, D97(2):023016, 2018.

[334] Sylvain Marsat, Alejandro Boh, Luc Blanchet, and Alessandra Buonanno. Next-to-leading tail-induced spinorbit effects in the gravitational radiation flux of compact binaries. *Class. Quant. Grav.*, 31:025023, 2014.

[335] Balazs Mikoczi, Matyas Vasuth, and Laszlo A. Gergely. Self-interaction spin effects in inspiralling compact binaries. *Phys. Rev.*, D71:124043, 2005.

[336] Eanna E. Flanagan and Tanja Hinderer. Constraining neutron star tidal Love numbers with gravitational wave detectors. *Phys. Rev.*, D77:021502, 2008.

[337] Justin Vines, Eanna E. Flanagan, and Tanja Hinderer. Post-1-Newtonian tidal effects in the gravitational waveform from binary inspirals. *Phys. Rev.*, D83:084051, 2011.

[338] Michele Maggiore. *Gravitational waves: Volume 1: Theory and experiments*, volume 1. Oxford university press, 2008.

[339] Alex Nitz, Ian Harry, Duncan Brown, Christopher M. Biwer, Josh Willis, Tito Dal Canton, Larne Pekowsky, Thomas Dent, Collin Capano, Andrew R. Williamson, Soumi De, Miriam Cabero, Bernd Machenschalk, Prayush Kumar, Steven Reyes, Thomas Massinger, Amber Lenon, Stephen Fairhurst, Alex Nielsen, Duncan Macleod, shasvath, Francesco Pannarale, Leo Singer, Stanislav Babak, Hunter Gabbard, dfinstad, John Veitch, CBC Sugar, Sebastian Khan, and Lorena Magaa Zertuche. ligo-cbc/pycbc: Post-o2 release 8, March 2018.

[340] Lee Samuel Finn. Aperture synthesis for gravitational-wave data analysis: Deterministic sources. *Phys. Rev. D*, 63:102001, 2001.

[341] Christian Rover, Renate Meyer, and Nelson Christensen. Coherent Bayesian inference on compact binary inspirals using a network of interferometric gravitational wave detectors. *Phys. Rev.*, D75:062004, 2007.

[342] M. Soares-Santos 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. *Astrophys. J.*, 848(2):L16, 2017.

[343] Michele Cantiello et al. A Precise Distance to the Host Galaxy of the Binary Neutron Star Merger GW170817 Using Surface Brightness Fluctuations. *Astrophys. J.*, 854(2):L31, 2018.

[344] Kent Blackburn et al. *LOSC CLN Data Products for GW170817*, 2017.

[345] Harold Jeffreys. *The theory of probability*. OUP Oxford, 1998.

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

[347] Hsin-Yu Chen and Daniel E Holz. The loudest gravitational wave events. *arXiv preprint arXiv:1409.0522*, 2014.

[348] B. P. Abbott et al. The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914. 2016. arXiv:1602.03842.

[349] B. P. Abbott et al. Tests of general relativity with GW150914. 2016. arXiv:1602.03841.

[350] B. P. Abbott et al. Astrophysical Implications of the Binary Black-Hole Merger GW150914. *Astrophys. J.*, 818(2):L22, 2016. arXiv:1602.03846.

Curriculum Vitae

Steven Reyes

Personal details

Citizenship	United States of America
Birth	May 4, 1992
Phone	(773) 315-2296
Mail	stevereyes01@gmail.com

Education

Syracuse University **2014-Present**

Doctor of Philosophy in Physics

Advisor: Dr. Duncan Brown.

Expected Graduation date: 2019

University of Chicago **2010-2014**

Bachelor of Arts in Physics with Astrophysics specialization

Employment

Graduate Student Research Assistant **2015-Present**

Syracuse University, Advisor: Dr. Duncan Brown

CARMA data analyst **2012-2013**

University of Chicago, Advisor: Dr. John Carlstrom

Honors and Awards

Gruber Cosmology Prize **2016**

The Gruber Foundation

Shared with the LSC

Breakthrough Prize in Fundamental Physics **2016**

Shared with the LSC

STEM Fellowship **2014, 2016**

Syracuse University

Odyssey Scholarship **2010-2014**

University of Chicago

Chicago Public School (CPS) Scholarship **2010-2014**

University of Chicago

Publications in which Steven Reyes made a significant contribution

Nitz, Alexander H and Capano, Collin and Nielsen, Alex B. and Reyes, Steven and White, Rebecca and Brown, Duncan A. and Krishnan, Badri. "1-OGC: The first open gravitational-wave catalog of binary mergers from analysis of public Advanced

LIGO data". In: (2018). arXiv:1811.01921 [gr-qc]

Steven Reyes and Duncan A. Brown "Constraints on nonlinear tides due to *p*-*g* mode coupling from the neutron-star merger GW170817". In: (2018). arXiv: 1808.07013 [astro-ph.HE]

Nitz, Alexander H. and Dal Canton, Tito and Davis, Derek and Reyes, Steven "Rapid detection of gravitational waves from compact binary mergers with PyCBC Live". In: *Phys. Rev.* D98.2 (2018), p. 024050. DOI: 10.1103/PhysRevD.98.024050. arXiv: 1805.11174 [gr-c]

Benjamin P. Abbott et al. "Upper Limits on the Rates of Binary Neutron Star and Neutron Star Black Hole Mergers From Advanced LIGO's First Observing Run". In: *Astrophys. J.* 832.2 (2016), p. L21. DOI: 10.3847/2041-8205/832/2/L21. arXiv: 1607.07456 astro-ph.HE

Contributed Talks and Posters

Searching for p-g mode coupling in GW170817 **2019**

Talk at STAG Research Centre (South Hampton, United Kingdom)

Model Selection with Mult-Tempering Techniques **2019**

Talk at PyCBC Inference Conference (Portsmouth, United Kingdom)

Searching for p-g mode coupling in GW170817 **2018**

Poster at GWPAW 2018 (College Park, Maryland)

The Search for Gravitational Waves from Binaries with Neutron Stars 2017

Talk at GWPAW 2017 (Annecy, France)

First Results from the Search for Binary Black Hole Coalescence with Advanced LIGO **2016**

Poster at Relativity and Gravitation: Contemporary Research and Teaching of Einstein's Physics, Gordon Research Conference.

Outreach

Lead Tour Guide at Holden Observatory **2015-Present**

Adopt-a-Physicist participant **2015, 2016**

Scholastic Dinosaur 13 Webinar Interview **2014**