

MIT Open Access Articles

*Characterization of low-significance
gravitational-wave compact binary sources*

The MIT Faculty has made this article openly available. **Please share**
how this access benefits you. Your story matters.

Citation: Huang, Yiwen, et al. "Characterization of Low-Significance Gravitational-Wave Compact Binary Sources." *Physical Review D*, vol. 98, no. 12, Dec. 2018. © 2018 American Physical Society

As Published: <http://dx.doi.org/10.1103/PhysRevD.98.123021>

Publisher: American Physical Society

Persistent URL: <http://hdl.handle.net/1721.1/120005>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



Characterization of low-significance gravitational-wave compact binary sources

Yiwen Huang,^{1,2,*} Hannah Middleton,^{3,4} Ken K. Y. Ng,^{1,2} Salvatore Vitale,^{1,2,†} and John Veitch⁵

¹*LIGO, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

²*Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

³*Institute of Gravitational Wave Astronomy and School of Physics and Astronomy, University of Birmingham, Birmingham, B15 2TT, United Kingdom*

⁴*OzGrav-Melbourne, University of Melbourne, School of Physics, Parkville, Victoria 3010, Australia*

⁵*Institute for Gravitational Research, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, United Kingdom*



(Received 24 October 2018; published 27 December 2018)

Advanced LIGO and Virgo have so far detected gravitational waves from 10 binary black hole mergers (BBH) and 1 binary neutron star merger (BNS). In the future, we expect the detection of many more marginal sources, since compact binary coalescences detectable by advanced ground-based instruments are roughly distributed uniformly in comoving volume. In this paper we simulate weak signals from compact binary coalescences of various morphologies and optimal network signal-to-noise ratios (henceforth SNRs), and analyze if and to which extent their parameters can be measured by advanced LIGO and Virgo in their third observing run. We show that subthreshold binary neutron stars, with SNRs below 12 (10) yield uncertainties in their sky position larger than 400 (700) deg² (90% credible interval). The luminosity distance, which could be used to measure the Hubble constant with standard sirens, has relative uncertainties larger than 40% for BNSs and neutron star black hole mergers. For sources with SNRs below 8, it is not uncommon that the extrinsic parameters, sky position and distance, cannot be measured. Next, we look at the intrinsic parameters, masses and spins. We show that the detector-frame chirp mass can sometimes be measured with uncertainties below 1% even for sources at SNRs of 6, although multimodality is not uncommon and can significantly broaden the posteriors. The effective inspiral spin is best measured for neutron star black hole mergers, for which the uncertainties can be as low as ~ 0.08 (~ 0.2) at SNR 12 (8). The uncertainty is higher for systems with comparable component masses or lack of spin precession.

DOI: [10.1103/PhysRevD.98.123021](https://doi.org/10.1103/PhysRevD.98.123021)

I. INTRODUCTION

Interferometric ground-based gravitational-wave (GW) detectors, such as Advanced LIGO [1,2] and Advanced Virgo [3] are most sensitive to GWs in the frequency band of [20–2,000] Hz [4]. The merger of compact objects such as neutron stars and stellar-mass black holes in binaries can produce GW signals in this frequency band, and will be the most commonly detected sources by ground-based detectors. At the time of writing this paper, LIGO and Virgo have published the detections of 10 BBHs [5–11] and one BNS, GW170817 [12,13].

These first few detections have already made clear how characterization of compact binary coalescence (CBC) sources will broaden our understanding of the key properties of black holes and neutron stars. As more detections are

made, they will make it possible to constrain the role of stellar wind, rotation, and metallicity in the progenitor stars [14], to measure the merger rate [15], the spin and mass distribution [16–23] and to perform stringent tests of general relativity in its strong-field regime [24–29].

Some of the expected sources of GWs are also luminous in the electromagnetic (EM) band, which allow for the possibility of multimessenger observations. This has been spectacularly shown with the joint detection of GW and EM signals from the BNS source GW170817 [30]. The host of GW170817 was identified, together with radiation in the whole EM spectrum, from radio to γ -rays [30–34]. The science output of this discovery is too rich to be fully described here. We thus only mention a few highlights. GW170817 was used to set constraints on the equation of state of neutron stars [35], search for evidence of p-g modes [36] and put bounds on the component neutron star masses and spins [13]. The EM data confirmed the connection between short gamma-ray

*ywh@mit.edu

†salvatore.vitale@ligo.org

bursts and BNS sources, lead to the observation of the kilonova, and yielded insights on the details of the EM emission [37–40]. Information from both the GW and the EM sides was used to measure the Hubble constant in a way that is independent of the cosmic distance ladder [41]. As more BNSs are detected in the next years, we will be able to gain a more solid understanding of the properties of compact binaries, their progenitors, and the electromagnetic radiation they emit.

Given the current detections, it is possible to estimate the local merger rates of binary neutron stars [12] and binary black holes [6,9,15]. Neutron star black hole mergers (NSBHs) are also promising sources, but have not been detected yet, hence there are only predictions for their merger rate [42]. The measured or predicted merger rates together with the projected improvements in the sensitivity of LIGO and Virgo [43] allow one to estimate the number of detections per year above some pre-defined matched-filter signal-to-noise ratio (SNR) or false alarm rate threshold [43].

As long as the distance distribution of GW sources can be considered uniform in volume (which is a good approximation for sources close enough that cosmological effects can be neglected), the distribution of events' luminosity distances d_L is proportional to d_L^2 . Given that the SNR ρ goes like $1/d_L$, the resulting SNR distribution goes like ρ^{-4} [44]. Even a smaller decrease in the threshold matched-filter SNR should thus correspond to a large increase in the number of detected events. In practice, other factors limit the benefit of lowering the detection threshold. The noise from GW detectors is not perfectly gaussian. Non-gaussian noise artifacts (often referred to as *glitches*) limit the sensitivity of GW searches [45]. To assess the significance of a candidate event, its SNR (or other detection statistics) is compared with the distribution of SNRs from the background, which is usually estimated either by time-sliding the data of different instruments [46] or by constructing the network SNR distribution assuming noise is independent in each detector [47]. Unfortunately, while the distribution of SNRs from CBCs goes like ρ^{-4} , the distribution of SNRs from glitches increases much faster as the SNR decreases [48,49].

This can limit the benefits of lowering the threshold while searching for CBC signals and following them up in the EM band. For example, Ref. [50] has shown that lowering the threshold false alarm rate to 1 per month (week) would result in only 39% (13%) of the BNS candidate to be of astrophysical origin. Observers who decide to follow up marginal events would thus have to deal with a large number of false positives.

In this paper we look at a complementary aspect of marginal events, namely the fact that their characterization can be challenging or inconclusive. Virtually all of the GW literature focusing on characterization of GWs from compact binaries has considered clear detections, generating

simulated GW signals with optimal SNRs above some threshold (often ~ 12). In this paper we reverse that approach, and only consider sources *weaker* than what could be considered a clear detection.

Since the Fisher matrix approximation fails at small SNRs [51], we perform full numerical simulations. We simulate BNS, BBH, and NSBH sources and add them to real interferometric noise of LIGO and Virgo, recolored to have the spectral behavior expected in the next observing run, starting in early 2019 [46]. We analyze the sources at different optimal network SNRs,¹ from 6 to 12, and show exactly how the quality of the parameter estimation process degrades as the optimal network SNR decreases. We find even at optimal network SNR of 12, BNS sources cannot be localized to areas smaller than ~ 400 deg² (90% credible interval). Meanwhile, the luminosity distance is always measured with relative uncertainties larger than 40%. It is not uncommon that the sky position and luminosity distance simply cannot be measured. This reveals the challenges of finding an EM counterpart to a weak GW source, and to make a compelling case that the association is real. The chirp mass, which is usually the best measured parameter, shows signs of multimodality, especially for heavy systems. At optimal network SNR 7 or below, for the majority of the sources we consider, we obtain posterior distributions for the chirp mass which is multimodal, or with very large tails. The uncertainty in the mass ratio is large for heavy systems, whereas it can be as small as 0.07 for NSBH at optimal network SNR 12. Finally, we show that, especially for NSBH, the effective spin [52–55] can be measured at very low optimal network SNRs.

The rest of the paper is structured as follows: in Sec. II, we report the setup of the simulations we performed, including noise and signal generation, and of the analysis. In Sec. III, we describe the main results for a set of simulated weak sources. The conclusions are presented in Sec. IV.

II. METHOD

A. Noise and signal generation

The data in each interferometer in presence of a GW signal can be written as \mathbf{d} :

$$\mathbf{d} = \mathbf{h} + \mathbf{n} \quad (1)$$

where \mathbf{h} is signal and \mathbf{n} is noise.

The necessary ingredients to simulate an end-to-end analysis of a GW signal are thus the generation of a synthetic GW signal, and a stretch of data that the signal

¹The square of the optimal network SNR is defined as the sum of the squares of the single-instrument optimal SNR, for all instruments taking data. The optimal SNR is defined in Eq. (3).

TABLE I. O1 times (in GPS and GMT) used for the analyses. These are the times at which signals are “injected.”

GPS time	Year	Month	Day	Time
1135924088	2016	Jan	04	06:27:51
1135989351	2016	Jan	05	00:35:34
1136267078	2016	Jan	08	05:44:21
1136506663	2016	Jan	11	00:17:26
1136594611	2016	Jan	12	00:43:14

can be added to. In this section we describe our approach to generate these quantities, starting from the noise.

We work with a network made of the two advanced LIGO instruments and the advanced Virgo detector. To make this study immediately relevant to the next observing run (O3, starting in early 2019 [4]), we work with noise streams that have the projected O3 sensitivity for each instrument.² The power spectral density (PSD), $S_n(f)$, of a stretch of data is defined as the average of the noise autocorrelation over the duration of the data segment [56]. It is trivial to produce Gaussian noise colored to have any specific PSD, and this is the approach followed in a significant fraction of the gravitational-wave literature. However, when dealing with marginal events, using real noise seems important, as small noise artifacts, or fluctuations, might have significant impact on weak signals.

We use public real data from the first observing run (O1) [57] and recolor it to have the expected O3 spectral behavior. More specifically, we select 5 GPS times in O1 such that the data around them does not contain any (known) astrophysical events nor significant instrumental artifact that would have resulted in vetoing of that data (Ref. [57] provides list of data segments which are considered clean by the LIGO and Virgo collaboration). The O1 times we used for our analysis are listed in Table I. We download the corresponding 5 data files for each of the LIGO instruments from the Ref. [57] and use a routine of the *gstlal* algorithm [47,58–61] to apply the expected O3 power spectrum (specifically the lower bounds of the “Late” curve (2018–19) for Advanced LIGO and of the “Mid” curve (2018–19) for Advanced Virgo in [4]).

Since Virgo was not running together with the advanced LIGO in O1, no Virgo O1 data is available for those 5 GPS times. We simulate Virgo data as follows. For each GPS time, we pick a Hanford frame corresponding to one of the *other* four GPS times, shift the time stamp to coincide with the desired GPS time, and recolor it to the projected O3 Virgo PSD.

We stress that while recoloring O1 data to a target sensitivity (O3, in this case) gives a way to make prediction

²However, the results we present will be relevant for any situations in which the two LIGO detectors accumulate most of the network SNR, which will likely be the case for the next few years [43].

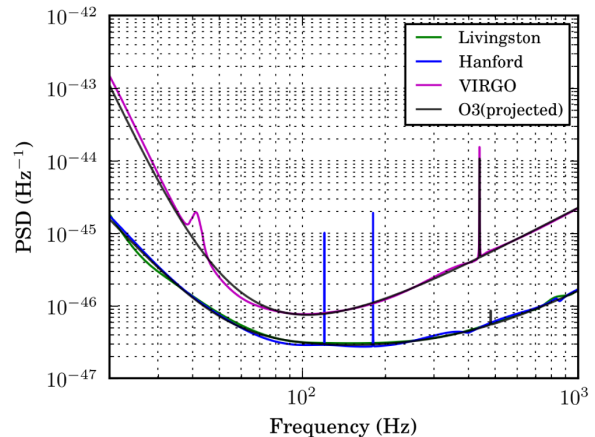


FIG. 1. PSD of the recolored data for advanced LIGO (green for Livingston, blue for Hanford) and advanced Virgo (purple). The black lines are PSD for each instrument at the projected O3 sensitivities [4].

about the performances of future science runs, it also has limitations. In particular, it will not capture new types of instrumental artifacts that might arise as the instruments get more sensitive. On the other hand, recoloring archival data maintains old instrumental artifacts that might very well be solved in the future. However, as our signals do not overlap with major instrumental artifacts, and given the lack of alternatives, we proceed with recolored data, and use Hanford data as time-shifted mock Virgo data.

The PSDs *estimated* with the BayesWave algorithm [62] for one of the data streams produced with this method are shown for each interferometer in Fig. 1 (colored curves), together with the projected O3 curves (black lines) that we use to recolor the O1 data.

We can describe the gravitational waveform emitted by a binary of two point masses in a circular orbit by a set of 15 parameters, θ , including masses and spins, as well as extrinsic parameters such as luminosity distance, orbital orientation, and sky position.³ Different parametrizations are possible for the mass parameters. In this work we will report the uncertainty of the asymmetric mass ratio $q = m_2/m_1$ (m_1, m_2 are the component masses, with $m_1 \geq m_2$ by convention), and the chirp mass \mathcal{M} :

$$\mathcal{M} = (m_1 m_2)^{3/5} (m_1 + m_2)^{-1/5} \quad (2)$$

In general, six parameters are needed to describe the spins of the binary: 2 dimensionless spin magnitudes a_1, a_2 defined as $a_i = |\mathbf{s}_i|/m_i^2$ where \mathbf{s}_i is the spin vector; the tilt angles t_1, t_2 between the spin vectors and the orbital

³BNS would require two additional parameters to model their linear tidal deformability [12,63–66] In this study we neglect tidal effects for BNS and NSBH to keep the computational cost reasonable. Given that tidal deformabilities are hard to measure even for loud events such as GW170817, they would have been unmeasurable with the marginal sources we consider in this work.

angular momentum at some reference frequency (equal to the lower frequency used in the analysis, see below) and two azimuth angles.

To generate the waveform signals, we use the IMRPhenomPv2 waveform family. This is an inspiral-merger-ringdown (IMR) waveform with an effective precessing spin [67–69]. We keep all the phase and amplitude corrections supported by the waveform.

The loudness of a given waveform can be assessed with its optimal signal-to-noise ratio ρ , defined in each interferometer by the inner product of a waveform with itself [70]:

$$\rho^2 = \langle h(\boldsymbol{\theta}, f) | h(\boldsymbol{\theta}, f) \rangle \equiv 4 \int_{f_{\text{low}}}^{f_{\text{high}}} \frac{h(\boldsymbol{\theta}, f) h(\boldsymbol{\theta}, f)^*}{S_n(f)} df \quad (3)$$

where $h(\boldsymbol{\theta}, f)$ is the frequency-domain waveform projected in the detector, and $S_n(f)$ is the PSD of the detector noise.

We stress that the optimal SNR is not exactly the output of search algorithms, which instead calculate a matched-filter SNR (defined as the inner product of the best waveform template with the data [56]). We prefer to work with the optimal SNR since that can be calculated from the waveform being added to the noise, without any knowledge of the exact realization of the noise. The optimal SNR is the expectation of the matched-filter SNR in the limit where the noise can be considered Gaussian. In the rest of the paper, we will just refer to optimal SNR as SNR, unless otherwise specified.

Clear detections typically have matched-filter network SNRs well above 12 for heavy objects [6–11], whereas BNS can be detected with high confidence at lower matched-filter network SNR [12,71]. Since we wish to focus on marginal events, we generate signals with lower network SNRs: [6,7,8,9,10,12]. We consider 4 representative CBC systems, BNSs, NSBHs, stellar-mass BBHs, and heavy stellar-mass BBHs (hBBHs). These representative morphologies capture some of the key features that can be present in CBC signals: long inspirals with little or no spin (BNSs); high mass ratio and visible spin precession (NSBHs) and heavier objects with little inspiral (BBHs and hBBHs). The masses and spins of each system are given in Table II. When the inclination angle is larger than ~ 70 degrees, the cross polarization becomes negligible, affecting the estimation of the extrinsic parameters [72–74]. Thus, we consider two different inclination angles for each

TABLE II. The intrinsic parameters for the 4 morphologies considered in this study.

Type	$m_1(M_\odot)$	a_1	$t_1(^{\circ})$	$m_2(M_\odot)$	a_2	$t_2(^{\circ})$
BNS	1.4	0	0	1.4	0	0
NSBH	8	0.8	46	1.4	0	0
BBH	12	0.6	60	6	0.1	60
hBBH	30	0.6	60	30	0.6	60

source: 30 degrees and 80 degrees, to take into account the effect of polarization. We assign the same geographical coordinate to all of our simulated sources. Specifically, they are overhead of the two LIGO detectors, which is the position where most events should be detected. This results in roughly equal SNR in the two LIGO detectors. We can quantify the sensitivity of each detector to a particular direction and polarization with the square root of the sum of the antenna pattern, F , squared [56]:

$$F^{\text{IFO}} = \sqrt{F_+^{\text{IFO}}(\alpha, \delta, \psi)^2 + F_\times^{\text{IFO}}(\alpha, \delta, \psi)^2}$$

where α is the right ascension, δ is the declination, and ψ is the polarization of the GW signal.

For all of our sources we have the same values: $F^H = F^L = 0.7$, $F^V = 0.3$. Each system has its distance scaled to give the desired SNR and added to the recolored data at the 5 GPS times of Table I. In total, we analyze $4(\text{massbins}) \times 2(\text{inclinations}) \times 6(\text{SNRs}) \times 5(\text{GPS times}) = 240$ simulated events.

B. Bayesian parameter estimation

In this section we describe the method used to measure the unknown parameters of the detected signals. Given that Bayesian inference in the field of gravitational waves is now standard (Refs. [75,76] provide excellent descriptions) we only quickly review it.

Given the data in the frequency domain \mathbf{d} and a signal model hypothesis H , the posterior probability density function (PDF) can be found using Bayes' theorem,

$$p(\boldsymbol{\theta} | \mathbf{d}, H) = \frac{p(\boldsymbol{\theta} | H) p(\mathbf{d} | \boldsymbol{\theta}, H)}{p(\mathbf{d} | H)} \quad (4)$$

where $p(\boldsymbol{\theta} | H)$ is the prior probability density of $\boldsymbol{\theta}$ given the hypothesis H , and $p(\mathbf{d} | \boldsymbol{\theta}, H)$ is the likelihood. Under the assumptions that the noise streams are statistically uncorrelated, the likelihood can be written as a product of individual likelihood from each interferometer as [77]

$$p(\mathbf{d} | \boldsymbol{\theta}, H) = \prod_{\text{IFO}} p(\mathbf{d}^{\text{IFO}} | \boldsymbol{\theta}, H) \quad (5)$$

where the product spans all the instruments in the network. The normalization constant $p(\mathbf{d} | H)$ is called the evidence of the data, Z :

$$Z = p(\mathbf{d} | H) = \int d\theta_1 \dots d\theta_N p(\mathbf{d} | \boldsymbol{\theta}, H) p(\boldsymbol{\theta} | H) \quad (6)$$

Given the multidimensional posteriors in Eq. (4), the posterior PDFs of any specific parameter can be found by marginalizing over all the other parameters:

$$p(\theta_1|\mathbf{d}, H) = \int d\theta_2 \dots d\theta_N p(\boldsymbol{\theta}|\mathbf{d}, H) \quad (7)$$

We use nested sampling implementation of *LALInference* [75] to stochastically explore the parameter space and produce posterior distributions for $\boldsymbol{\theta}$. To reduce the computational cost of the likelihood evaluations, we use the reduced order quadrature (ROQ) approximation [78]. Sampling the parameter space to measure the properties of weak signals in presence of strong priors can be challenging. In Appendix A we report on some sanity checks we have performed to verify the code had properly converged.

For BNS events we start the analysis at 24 Hz, following the LIGO and Virgo collaborations [12], while for all other sources we start at 20 Hz.

C. Choice of priors

Bayesian inference requires explicit priors, in our case $p(\boldsymbol{\theta}|H)$ under the hypothesis that a CBC signal is present in the data. For all parameters except the luminosity distance (see below), we use the same priors used by the LIGO and Virgo collaborations for the CBCs detected so far [7–13,76]. For sky position, orbital orientation and spin orientation, we use isotropic priors. For the dimensionless spin magnitude, we used uniform priors in the range [0, 0.89] for black holes [7–11,76] and [0, 0.05] for neutron stars [12,13]. The priors on the component masses are also uniform, in the range in which the corresponding ROQ basis is valid [78].

The most natural prior for the luminosity distance would be a prior uniform in volume, $p(d_L) \sim d_L^2$, since the detection horizon for CBC is tens or hundreds of megaparsecs, depending on the total mass. This is indeed the prior used in LIGO-Virgo papers.

In this study we prefer to use a *uniform* prior on the luminosity distance, and then reweight the samples to enforce a uniform-in-volume prior. The reason is as follows. The nested sampling algorithm samples the prior [79] to calculate the evidence Z , and obtains the posterior distribution as a by-product. When using a uniform-in-volume distance prior, the nested sampling algorithm will spend a significant fraction of time exploring the region of parameter space where the distance is large. If we could run the algorithm for an arbitrarily large number of steps, it would eventually converge to the correct parameters. But for low-significance events like the ones we are considering, and using a finite number of steps, the algorithm might in practice not explore properly the part of the parameter space where the distance is small, since it would have to overcome a significant prior penalty. By sampling with a uniform-in-distance prior we avoid this issue. The sampler can explore with ease the whole distance range, and the correct prior is applied in post-processing with a standard rejection sampling approach.

III. RESULTS

In this section we report the uncertainty in measuring some of the key parameters of the simulated events. Unless otherwise specified, we used 90% credible intervals (CI). Those are either absolute intervals (in the appropriate units), or relative to the true value (in percentages).

A. Extrinsic parameters

Extrinsic parameters such as sky location and luminosity distance are of fundamental importance for EM follow-ups. Although the details of the follow-up to GW triggers vary with the facility and geographical factors (e.g., altitude of the source) [30], one usually tries to cover all or part of the sky uncertainty area (or sky-distance uncertainty volume). In fact, 3-dimensional uncertainty volumes have been routinely released in low latency by the LIGO and Virgo collaborations for significant events in the first and second observing run. Lowering the threshold for making triggers public would in general imply a significant increase of the false positive fraction [50]. In this section, we show that additionally, even for the events which are of astrophysical origin, an EM follow-up campaign might be prohibitive.

In Fig. 2, we show the 90% credible regions (deg^2) for all the sources we analyzed. Each color represents a different morphology (red for BNS, purple for NSBH, blue for BBH, green for hBBH), whereas the shape of the bullet indicates the inclination angle (circles for 30° , stars for 80°). The optimal network SNR (for the remainder of this section we will drop the “network”) is reported in the horizontal axis. Notice that we have artificially introduced a small horizontal offset while plotting to avoid significant superposition between data from different morphologies. The 5 symbols for each (SNR, morphology, inclination) set are the results from the 5 GPS times.

We see the expected overall trend of decreasing uncertainties with increasing SNRs. At SNRs below 10, we find that some of the sources are localized to uncertainties of 10,000 deg^2 or more. We have verified that the outliers are

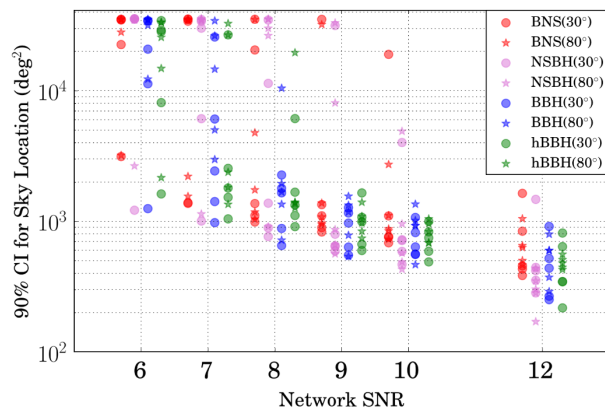


FIG. 2. 90% credible interval of the marginalized posteriors of the sky location vs network SNR.

due to the specific data stream into which events were added. In particular, we find that 1 of the 5 GPS times produces systematically lower uncertainties. At these low SNRs, the actual noise realization can have significant impact on the outcomes of the parameter estimation process. As the SNRs increase, the signals can be more easily distinguished from the noise, and the latter plays less of a role.

The loudest events we consider have SNRs of 12. For most of those, the 90% credible regions of sky localization are of $[200\text{--}1,000]$ deg². Even at SNRs below 10, there are sources that can be localized within areas of $[400\text{--}1,000]$ deg².

For comparison, Ref. [80] simulated a large number of BNS with astrophysically motivated parameters, and found that 50% of the BNS detectable by a LIGO-Virgo network (at their O2 sensitivity) would have 90% CI of 235 deg² or smaller if a detection threshold of matched-filter network SNR above 12 is used. While follow-up will be challenging, these large error areas may be accessible by wide-field/all-sky survey instruments available across the electromagnetic band and neutrinos. Here below we discuss some of the instruments.

The most viable option would be radio [81–83], where wide-field instruments like the Karl G. Jansky Very Large Array and the Very Long Baseline Array [84] of the National Radio Astronomy Observatory, and the Long Wavelength Array [85] can cover a large fraction of the sky. Coincident, high-energy (GeV to EeV) neutrinos might be searched in all-sky neutrino observatories like the IceCube [86] or the Antares [87]. Gamma-ray bursts can be found independently of the GW observation, which could be used to confirm the astrophysical nature of the GW candidate. The Fermi Gamma-ray Burst Monitor [88] aims to localize strong burst near the center of the field of view (FoV) of the Large Area Telescope, which is over 6,000 deg². From Earth, the High-Altitude Water Cherenkov [89] has an instantaneous FoV of over 6,000 deg² (15% of the sky). However, given the large uncertainties in sky position and distance, one might have to deal with a significant background of EM signals, which would make it difficult to claim a solid association. In this scenario, requiring both position and time coincidence (for the prompt EM emission) might increase the chances of success. Even at SNRs of 6, the arrival time of the GW signals at the detectors can be determined within a fraction of a second, see Appendix B. Instruments working at optical frequencies usually have smaller FoV: for example, the Asteroid Terrestrial-impact Last Alert System [90] has a FoV of 60 deg², but it scans a large fraction of the sky every night of operation. There might be more observatories available at the time of O3. For example, the Zwicky Transient Facility [91,92] with a FoV of 47 deg² in the optical, and the Cherenkov Telescope Array [93] with whole sky coverage for gamma rays. Beyond O3, the Large

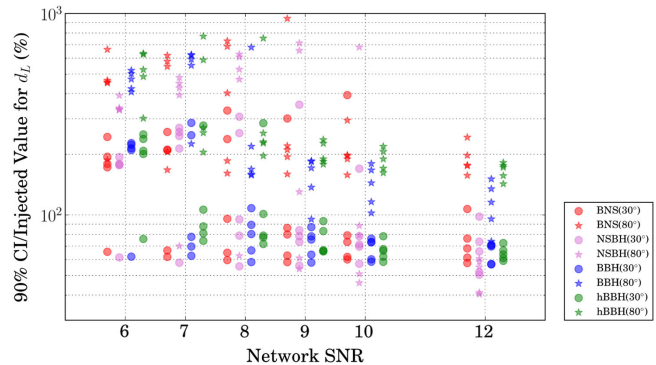


FIG. 3. Relative 90% credible interval of the marginalized posteriors of the luminosity distance d_L vs network SNR.

Synoptic Survey Telescope [94] with a FoV of 9.62 deg² in the optical might come online in 2023.

Similar large variations in precision can be seen for the luminosity distance d_L , Fig. 3. Relative uncertainties are above 40% at all SNRs, and very often above 100%. We observe that NSBH systems typically yield the smallest uncertainties. This is expected and is due to the fact that we allowed spin precession for NSBH systems, as given in Table II. Spin precession, together with the high mass ratios of NSBH, helps breaking the degeneracy between distance and inclination angle, leading to smaller uncertainties [21,73,95].

B. Intrinsic parameters

We now discuss the estimation of intrinsic parameters: mass and spins.

For clear detections, the detector-frame chirp mass [defined in Eq. (2)] is typically measured very well, with relative uncertainties of $\sim 0.01\%$ for BNS sources [13,96,97] and around 15% for heavier objects [6–11,76]. The reason why the chirp mass is measured better for low-mass systems is trivially that it affects the phase evolution at the lowest post-Newtonian order [98], and thus can be measured better for signals with long inspirals. Since the merger frequency decreases as the total mass increases, BNSs are the sources for which the chirp mass can be best measured.

In Fig. 4, we report the relative 90% uncertainty for the chirp mass against the network SNR. Overall, we see that BNS have the smallest relative uncertainties followed by NSBH, BBH and heavy BBH. The different morphologies are thus naturally sorted by total mass, as expected.

Let us first discuss the high-SNR end of our simulations, as the weakest events deserve a separate discussion. At SNR 10 and 12, the BNSs have relative uncertainties between 0.03% and 0.1%. This is in the ballpark of what one could have guessed using a simple $1/\text{SNR}^2$ scaling argument [70], and GW170817 as a reference point. That simple analysis is not totally accurate, since the variance scales like $1/\text{SNR}^2$ only at large SNRs. For smaller SNRs,

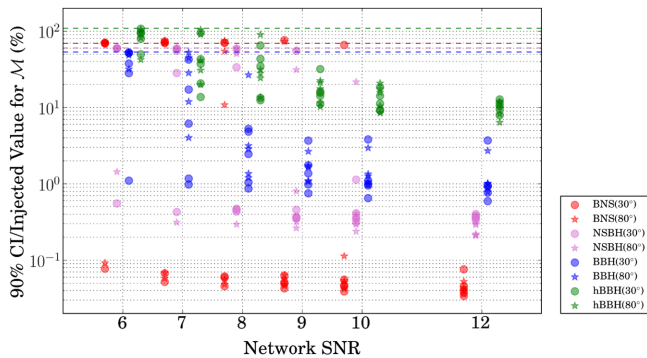
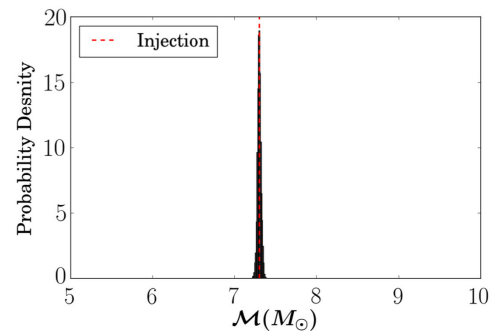


FIG. 4. Relative 90% credible interval of the marginalized posteriors of the detector-frame chirp mass \mathcal{M} vs network SNR. The dashed lines represent the relative 90% width of the prior.

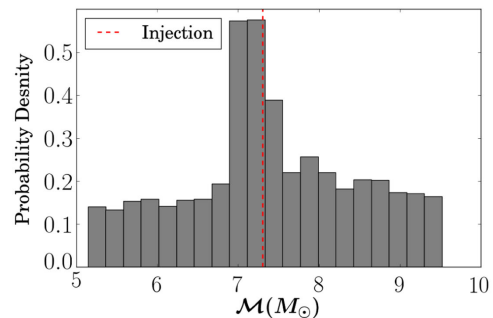
higher order corrections must be taken into account [99,100], and the uncertainties in the plot are indeed well above what a simple scaling argument would suggest. At the high end of our SNR distribution we find that, with very few exceptions, the results from different GPS times and orientations are grouped together, and the uncertainty mostly depends on the morphology. Due to their relatively long inspirals, NSBH are the second-best type of source, with uncertainties between 0.2% and $\sim 2\%$ at SNRs 10 and 12, including all data realizations and orientations. Stellar-mass black hole binaries have uncertainties between 0.5% and $\sim 4\%$, a factor of 2 worse than NSBH, while heavy BBH are considerably worse, with uncertainties between 5% and 30% (again, at SNR 10 and 12).

As the SNRs decrease, we observe 2 distinct populations: one that roughly continues the trend we see at SNR 10 and 12 with some degradation; and another with much larger uncertainties, comparable to the prior width, especially for BNSs. The latter are events for which the posteriors are not unimodal. One could expect that at low SNRs, noise fluctuations can seriously impact the measurement of the chirp mass, since this is obtained by “following” the phase evolution of the waveform signal through thousands of cycles [101]. We get a (rough) classification of the chirp mass posteriors by using the *find_peaks* routine of Scipy [102] using a prominence threshold of 9% of the main peak to trigger the presence of additional modes. While more sophisticated statistical tests could be used, this gives us at least an idea of which fraction of posteriors is clearly unimodal. A distinct type of chirp mass posterior appears often enough to deserve some discussion: those with clearly unimodal posteriors but with tails spanning large enough to cover most or all of the prior range. These are sources for which the likelihood profile is shallow, and does not dominate the underlying prior distribution.

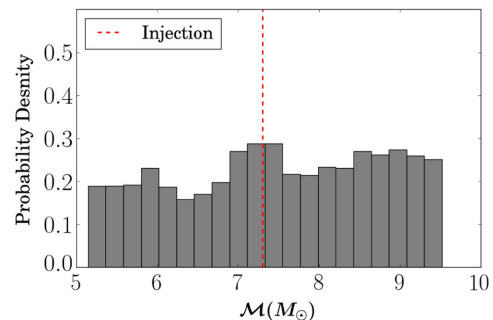
We thus introduce 3 different categories: unimodal; “unimodal-wide”, which we define as posteriors with one clear peak but a 90% uncertainty larger than 50% of the prior range; and multimodal posteriors. Examples of each are shown in Fig. 5.



(a) Unimodal. The network SNR is 12.



(b) “Unimodal-Wide”. The network SNR is 7.



(c) Multimodal. The network SNR is 6.

FIG. 5. Representative posterior distribution for \mathcal{M} . All systems are BBHs with true inclination angle of 80° . The range is the same in all panels in the x axis.

We apply this classification scheme to all events, and show the fraction of events that belong to each category in Fig. 6. At SNR 10 and 12, all morphologies but BBH show unimodal posteriors. Some of the BBH sources at these SNRs have hints of secondary modes which, however, do not significantly broaden the 90% credible intervals, as clear from Fig. 4. On the opposite end, at SNRs of 6 and 7, most events have multimodal posteriors, or a broad posterior filling up most of the prior range with just a hint of peak at the true chirp mass value.

While we directly measure detector-frame chirp mass, the more astrophysically interesting parameter is source-frame chirp mass $\mathcal{M}^{\text{source}}$:

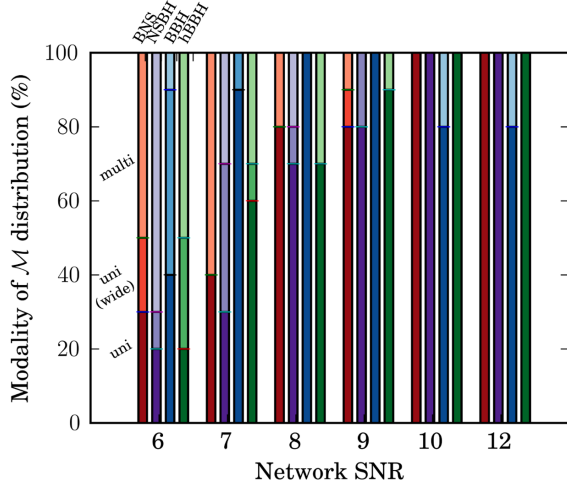


FIG. 6. Proportions for unimodal, unimodal (wide) and multimodal posterior distributions for detector-frame chirp mass \mathcal{M} (%) vs network SNR.

$$\mathcal{M} = (1 + z)\mathcal{M}^{\text{source}}.$$

Since GWs do not provide a direct measurement of the redshift, one must use the measured luminosity distance and a fiducial cosmology to convert that into a redshift [12,13]. Since the uncertainty in the distance are usually large for the events we simulated, the source-frame chirp mass will be measured more poorly than the detector-frame chirp mass, Fig. 7. We note that even though the uncertainties follow similar patterns, the scale is different than what was found in Fig. 4. The best measured systems of BNSs at SNR 6, for example, have relative uncertainties $\sim 5\%$ instead of 0.1% or better. At SNRs 10 and 12, BNSs are measured to uncertainties $\sim [2-5]\%$, NSBHs $\sim [2-7]\%$, BBHs $\sim [5-10]\%$, and hBBHs above 10% . Yet, it would be possible to get a better redshift measurement, thus a better source-frame chirp mass measurement, if the host galaxy is identified by EM follow-up campaigns for BNSs or NSBHs.

Next, we consider the asymmetric mass ratio q , Fig. 8. We expect NSBH to yield the best measurement, since

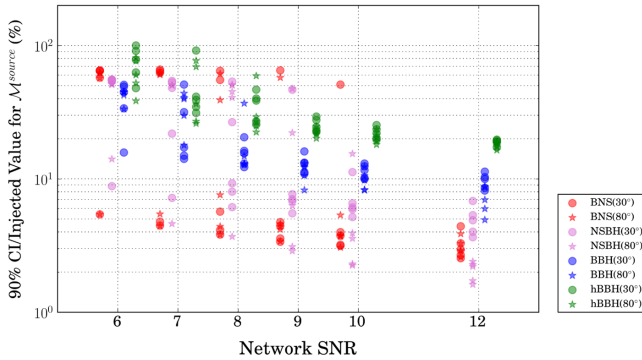


FIG. 7. Relative 90% credible interval of the marginalized posteriors of the source-frame chirp mass $\mathcal{M}^{\text{source}}$ vs network SNR.

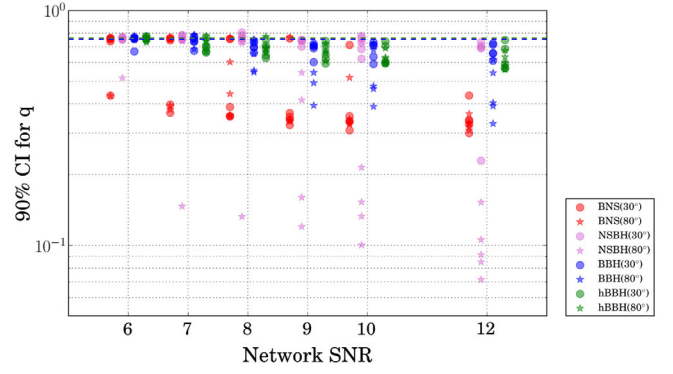


FIG. 8. 90% credible interval of the marginalized posteriors of the mass ratio q vs network SNR. The dashed line represents the 90% CI for q prior.

spin-induced precession breaks the degeneracy between mass ratio and spins [21,73,95,103], improving the measurability of both. For NSBH, we find uncertainties as small as 0.07 for SNR 12 and orientations close to edge-on. For the NSBH with orientation close to face-on, the uncertainty is systematically worse, at all SNRs. The reason is again correlations. When orbital precession is present, inclinations closer to 90° yield better spin measurement [21,22]. The smaller spin uncertainties result in smaller mass ratio uncertainties. The same trend is visible for the BBHs. However, hBBHs have comparable uncertainties regardless of the orientation, because the hBBHs in our simulations are equal-mass, which suppresses spin precession. Owing to their longer inspiral phase, the uncertainties for the BNS are of $[0.3-0.5]$ for most BNSs in the SNR range 7–12, better than for BBHs and hBBHs. These results suggest that, if the real source is a marginal a NSBH with visible spin precession, one might be able to distinguish between a NSBH and a low-mass BBH, but not so easily between a BNS and a low-mass NSBH.

We now look at the estimation of the spins. Since individual spins are hard to measure even for loud sources [6–11,13,21,22,76,103] we focus on the effective inspiral spin, χ_{eff} :

$$\chi_{\text{eff}} = \frac{c}{G(m_1 + m_2)} \left(\frac{S_1}{m_1} + \frac{S_2}{m_2} \right) \cdot \frac{\mathbf{L}}{|\mathbf{L}|} \quad (8)$$

where \mathbf{L} is the orbital angular momentum. χ_{eff} is the mass-weighted projection of the total spin onto the orbital angular momentum and takes values between -1 and 1 [52–55]. While individual spins will not often be measurable, the effective spin is usually measured well [6–11,13,22,76,104]. Importantly, the effective spin can also be used to distinguish between different astrophysical formation channels [96,105], instead of relying on the measurement of the individual spins [16,106].

We report the 90% CI for χ_{eff} in Fig. 9, where the horizontal dashed lines represent the 90% of the prior

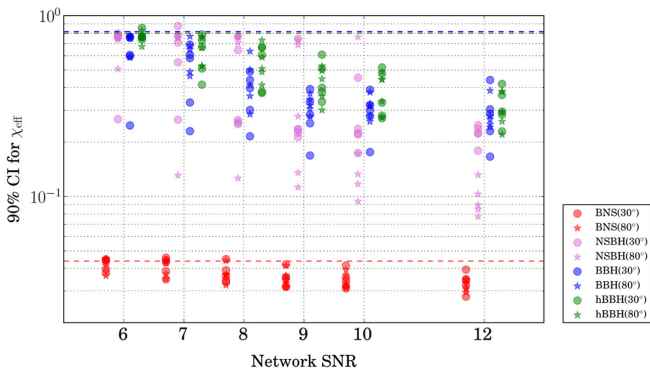


FIG. 9. 90% credible interval of the marginalized posteriors of the effective spin χ_{eff} vs network SNR. The dashed line represents the 90% CI for χ_{eff} prior.

width. For all sources that include at least one black hole, the uncertainties at SNR 12 are between ~ 0.08 and ~ 0.3 , that is, a factor of 2–10 narrower than the prior. At SNR 10, the uncertainties for these systems are between ~ 0.1 and ~ 0.4 . For the same reasons that we described while discussing the mass ratio results, NSBH oriented edge-on are the events for which the effective spin estimation is best. Heavy BBH are the worst since they have short inspirals, and mass ratio of unity. As the SNRs decrease, for some sources the χ_{eff} posterior is still informative. Conversely, for the BNS sources, even at SNR 12 the posterior is only marginally narrower than the prior.

To quantify the amount of information gained about χ_{eff} after analyzing the data, we calculate the Kullback-Leibler (K-L) divergence [107] of the posterior Q over the prior P :

$$D_{\text{KL}}(P||Q) = \sum_i P(i) \ln \frac{Q(i)}{P(i)} \quad (9)$$

where the index i spans the samples. This is shown in Fig. 10. For most of the sources with at least one black hole at SNR 12, the data yield more than 1 nat of information. NSBHs are the systems for which most information is gained. The data is usually less informative for BNSs. At SNR 12, the typical information gained for BNSs is ~ 0.1 nats, while at low SNR no information is gained,

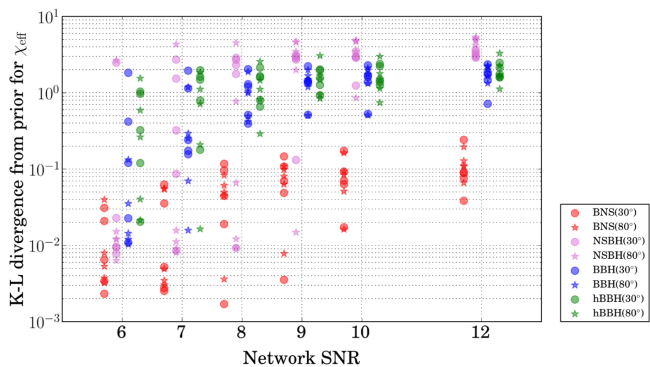


FIG. 10. K-L divergence for χ_{eff} vs network SNR.

which implies the posterior is basically the prior. This is consistent with Fig. 9. As the SNR decreases, the data is not allowed to significantly update the prior, and the K-L divergence can get values below 0.1 nats.

The effective spin parameter related to precession, χ_p [69,108], is not measurable at these SNRs. Only for NSBH and at high inclinations the data provides some information about χ_p . In Appendix C, we report the median K-L divergence over the 5 GPS times for all parameters and all signal morphologies.

IV. CONCLUSIONS

Compact binary coalescences, the most common source of gravitational waves detectable by ground-based detectors, are expected to be distributed uniformly in volume. This implies that their SNR ρ should be distributed as ρ^{-4} : for each loud detection there should be many more marginal signals in the data.

In practice, lowering the threshold matched-filter network SNR (or other detection statistics) will not increase the number of detections with the fourth power, since the background from instrumental and environmental sources increases more steeply. As the detection threshold is lowered, candidate events will be competing with background events of non-astrophysical origin [50,71,109,110].

It is still the case, though, that a significant fraction of detections made in the next few years will be marginal. In this paper we have explored what kind of astrophysical information can be extracted from these weak signals. This is a topic that was not extensively explored in the literature, since most existing work focuses on clear detections (but see Ref. [110]).

We have simulated GWs from various CBC sources with different network SNRs, from 6 to 12, and added them into real interferometric data. We have considered a network made by the two advanced LIGOs and the advanced Virgo detector at their expected sensitivity for the third observing run (planned for early 2019 [4]).

We have shown that the 90% credible regions in the sky localization of the sources are of $[200 - 1,000] \text{ deg}^2$ for network SNR of 12, for all signal morphologies. As the network SNR decreases, the uncertainty increases, and a larger spread between different noise realizations is present, which confirms that the specific noise realization can seriously affect the outcome of the analysis at very low SNRs. At the lowest SNRs we consider, most sources are simply not localized. It is important to notice that for those events it is not systematically the case that other parameters are not measured (e.g., the chirp mass). The reason is that a single detector is enough to measure an intrinsic parameter like the chirp mass, while at least two are needed to constrain the sky position, through time triangulation as well as amplitude and phase consistency [111–113].

The measurement of the luminosity distance, which could be used to infer the Hubble constant if a counterpart

is found [41,73,114–116] (or statistically, in absence of counterparts [117,118]), is also challenging for similar reasons. We have found relative uncertainties for the luminosity distance above 40% even for the loudest event we considered, and above 100% for a significant fraction of events. The large sky areas within which these marginal events are localized could be searched by wide-field/all-sky survey instruments available across the electromagnetic band and neutrinos. Requiring time coincidence between the gravitational and the prompt EM emission might further increase the significance of a joint detection.

We have verified that intrinsic parameters, such as detector-frame chirp mass and the effective spin, can usually be constrained, and do not simply yield the prior. We have found that the detector-frame chirp mass can be estimated to be better than 0.1% for BNS, 2% for NSBH and 4% for stellar-mass BBHs, at SNR of 10 and 12. As the SNR decreases, the uncertainties increase gradually for the bulk of source, but for some noise realizations the posteriors start being multimodal, or develop fat tails, which dramatically increases the uncertainty. For the BBHs, the bulk of the simulations yields uncertainties in the range $\sim[0.5\text{--}10]\%$. Owing to their shorter inspiral, heavy BBH have larger uncertainties, above 5% at SNR 12 and above 8% at SNR 10 or lower.

The source-frame chirp mass is measured with uncertainties that are an order of magnitude larger for the BNSs, NSBHs and BBHs, and a factor of two larger for hBBHs, than those for the detector-frame chirp mass, due to large uncertainties on the luminosity distance measurements (thus redshift measurements). These uncertainties would decrease if an EM counterpart is identified, and provides a better redshift measurement.

The mass ratio can only be significantly constrained for sources with visible spin precession (NSBHs) or long inspirals (BNSs). At SNR 12, the 90% can be as small as ~ 0.07 for our NSBHs, and $[0.3\text{--}0.5]$ for BNS.

Finally, the effective inspiral spins can be constrained, obtaining posterior distributions a factor of many narrower than the prior for systems with large mass ratios and spin precession. For those, at SNR 10 the uncertainty can be as small as ~ 0.1 , although ~ 0.4 is more common (the prior width is 0.89). Using the K-L divergence we quantified the amount of information provided by the data. For systems with at least one BH, that number is usually above 1 nat. For BNS, we have used a more restrictive prior (the prior width is 0.05), and obtained that the data usually yield only some information. At SNR 12, the typical K-L divergence we have found for BNSs is 0.1 nats. As the SNR decreases, the data does not allow to significantly update the prior and the K-L divergence can get values below 0.1 nats.

Based on the uncertainties mentioned above, it might be occasionally possible to associate a marginal source to a specific astrophysical class (e.g., BBH rather than BNS). It is important to remember that even if each event

individually is not particularly informative, the whole population of marginal events can be used and contribute to the astrophysical inference of the underlying population. This can be done even if one is not certain about their astrophysical origin, since the probability that the event is astrophysical can be folded in the analysis [110].

The posterior samples produced for this study have been made available in a public repository [119].

ACKNOWLEDGMENTS

Y. H., K. N. and S. V. acknowledge support of the MIT physics department through the Solomon Buchsbaum Research Fund, the National Science Foundation, and the LIGO Laboratory. H. M. acknowledges support from the Science and Technology Facilities Council (STFC) and the Australian Research Council Center of Excellence for Gravitational Wave Discovery (OzGrav) through Project No. CE170100004, as well as travel support from the Royal Astronomical Society and the Institute of Physics. J. V. is supported by STFC Grant No. ST/K005014/1. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0757058. The authors would like to thank Hsin-Yu Chen, Thomas Dent, Carl-Johan Haster, Erik Katsavounidis, and Ryan Lynch for useful discussions, and Chad Hanna and Kipp Cannon for help with running the recoloring routines of *gstlal*. The authors acknowledge usage of LIGO Data Grid clusters. This is LIGO Document Number P1800309.

APPENDIX A: SAMPLING CONVERGENCE

As mentioned in the body of the paper, sampling algorithms are only guaranteed to return the correct posteriors in the limit in which they run for an infinite amount of time. Especially at low SNRs, there is the risk that the sampler does not reach the right part of the parameter space, or gets stuck in a local and unrelated maximum. We use the maximum value of the recovered log likelihood as a probe of the convergence of the run, by comparing it with the log likelihood corresponding to the exact waveform parameters.

For a well converged run, the difference:

$$\Delta \log L \equiv \log L_{\text{injected}} - \max(\log L_{\text{recovered}})$$

should be slightly negative (and not exactly zero since the component of the noise that happens to be correlated across the network can contribute to the maximum recovered log likelihood. Conversely, positive values would suggest that the code failed to collect all the evidence that was available, which would be indicative of a problem with convergence.

We show $\Delta \log L$ for all sources in Fig. 11. We find the expected behavior for all sources except for BNS, for which

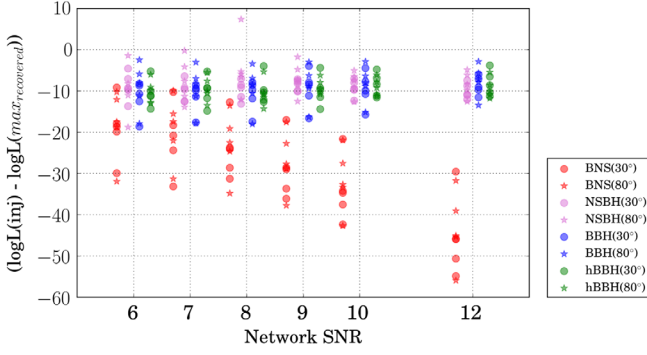


FIG. 11. Log likelihood for injection minus the maximum recovered log likelihood vs network SNR.

a downward trend is visible. We explain this difference with the fact that the current implementation of the ROQ likelihood in the *lscsoft* algorithm repository has a known issue for long signals [104]. Imperfections in the waveform reconstruction used in the ROQ basis can explain the different behavior of the BNS points.

To verify that the uncertainties we obtained for the BNS runs are meaningful, we have checked that the BNS runs do recover the correct SNR. Furthermore, we have run a small numbers of BNS sources (the ones used in Ref. [104]) both with and without the ROQ likelihood and found that the uncertainties are similar.

APPENDIX B: ARRIVAL TIME

Here we report the 90% CI for the arrival GPS time, Fig. 12. Time coincidence may be helpful in making an

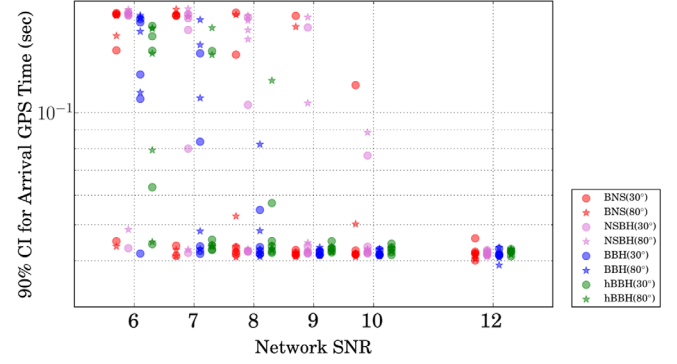


FIG. 12. Relative 90% credible interval of the marginalized posteriors of the arrival GPS time vs network SNR.

association of the GW signal with an EM counterpart. We find that the arrival time can be determined with uncertainties below 0.05 seconds for all sources at SNR of 12, and even at SNRs as low as 6 for some sources.

APPENDIX C: K-L DIVERGENCE

Here we report the median of K-L divergence (in nats) over 5 GPS times for luminosity distance d_L , arrival GPS time t , right ascension α , declination δ , detector-frame chirp mass \mathcal{M} , source-frame chirp mass $\mathcal{M}^{\text{source}}$, mass ratio q , effective spin χ_{eff} , and effective precessing spin χ_p , for the four morphologies and the two inclinations, Tables III–VI. Numbers close to zero imply the data is not informative about that parameter (at a given SNR). Conversely, large K-L divergence implies the prior and posteriors are significantly different.

TABLE III. Median of K-L divergence (in nats) over 5 GPS times for BNS.

Network SNR	6	7	8	9	10	12
d_L (inc = 30°)	0.01	0.06	4.43	5.28	5.83	6.67
d_L (inc = 80°)	0.01	0.02	3.86	5.17	5.72	6.68
Time (inc = 30°)	0.01	0.03	2.61	2.78	2.82	2.85
Time (inc = 80°)	0.00	0.02	2.13	2.83	2.86	2.92
α (inc = 30°)	0.01	0.02	0.39	0.46	0.45	0.90
α (inc = 80°)	0.00	0.01	0.41	0.41	0.53	0.76
δ (inc = 30°)	0.01	0.02	0.28	0.59	0.53	0.98
δ (inc = 80°)	0.00	0.01	0.46	0.35	0.36	0.54
\mathcal{M} (inc = 30°)	0.02	0.04	5.61	10.45	10.46	11.21
\mathcal{M} (inc = 80°)	0.01	0.03	4.54	10.68	10.82	10.98
$\mathcal{M}^{\text{source}}$ (inc = 30°)	0.01	0.02	3.94	4.43	4.65	4.97
$\mathcal{M}^{\text{source}}$ (inc = 80°)	0.01	0.03	3.37	4.40	4.60	4.91
q (inc = 30°)	0.00	0.01	1.70	1.79	1.81	1.85
q (inc = 80°)	0.00	0.00	1.48	1.83	1.89	1.93
χ_{eff} (inc = 30°)	0.01	0.01	0.05	0.08	0.07	0.09
χ_{eff} (inc = 80°)	0.01	0.01	0.06	0.08	0.08	0.12
χ_p (inc = 30°)	0.01	0.00	0.04	0.06	0.06	0.06
χ_p (inc = 80°)	0.00	0.01	0.06	0.06	0.06	0.06

TABLE IV. Median of K-L divergence (in nats) over 5 GPS times for NSBH.

Network SNR	6	7	8	9	10	12
d_L (inc = 30°)	0.01	0.67	4.23	5.20	5.83	6.65
d_L (inc = 80°)	0.01	0.02	0.80	5.98	7.28	8.77
Time (inc = 30°)	0.00	0.27	2.57	2.88	2.88	2.95
Time (inc = 80°)	0.00	0.01	0.22	2.70	2.81	3.37
α (inc = 30°)	0.00	0.14	0.41	0.67	0.75	0.90
α (inc = 80°)	0.00	0.00	0.04	0.68	0.70	1.45
δ (inc = 30°)	0.00	0.21	0.48	0.68	0.73	0.92
δ (inc = 80°)	0.00	0.01	0.07	0.70	0.68	1.62
\mathcal{M} (inc = 30°)	0.01	0.53	4.99	7.35	7.39	7.49
\mathcal{M} (inc = 80°)	0.01	0.04	0.58	5.71	7.52	8.05
$\mathcal{M}^{\text{source}}$ (inc = 30°)	0.01	0.29	2.21	2.61	2.82	3.08
$\mathcal{M}^{\text{source}}$ (inc = 80°)	0.02	0.03	0.29	2.79	3.48	4.13
q (inc = 30°)	0.01	0.06	0.13	0.08	0.39	0.30
q (inc = 80°)	0.01	0.00	0.03	0.55	1.24	1.74
χ_{eff} (inc = 30°)	0.01	0.29	2.30	2.85	2.89	3.20
χ_{eff} (inc = 80°)	0.01	0.01	0.07	3.52	4.71	5.03
χ_p (inc = 30°)	0.00	0.01	0.18	0.20	0.19	0.34
χ_p (inc = 80°)	0.00	0.00	0.00	1.32	1.57	1.76

TABLE V. Median of K-L divergence (in nats) over 5 GPS times for BBH.

Network SNR	6	7	8	9	10	12
d_L (inc = 30°)	1.92	4.69	5.11	5.60	6.15	6.89
d_L (inc = 80°)	0.02	2.09	5.43	5.71	6.60	7.76
Time (inc = 30°)	0.95	2.19	2.53	2.94	3.11	3.02
Time (inc = 80°)	0.01	1.05	2.60	2.90	2.94	2.97
α (inc = 30°)	0.08	0.18	0.38	0.70	0.71	0.89
α (inc = 80°)	0.00	0.14	0.38	0.62	0.55	0.61
δ (inc = 30°)	0.15	0.45	0.41	0.52	0.59	0.68
δ (inc = 80°)	0.00	0.15	0.56	0.67	0.69	0.82
\mathcal{M} (inc = 30°)	1.55	3.28	4.18	4.94	5.48	5.58
\mathcal{M} (inc = 80°)	0.03	1.24	4.26	4.76	5.09	5.44
$\mathcal{M}^{\text{source}}$ (inc = 30°)	0.78	2.21	2.46	3.09	3.64	4.27
$\mathcal{M}^{\text{source}}$ (inc = 80°)	0.01	0.90	2.72	3.26	4.06	5.02
q (inc = 30°)	0.13	0.27	0.47	0.57	0.60	0.63
q (inc = 80°)	0.02	0.14	0.24	0.51	0.45	0.66
χ_{eff} (inc = 30°)	0.12	0.27	1.25	1.44	1.65	1.79
χ_{eff} (inc = 80°)	0.01	0.26	1.01	1.52	1.78	2.05
χ_p (inc = 30°)	0.01	0.07	0.06	0.08	0.10	0.09
χ_p (inc = 80°)	0.00	0.01	0.25	0.75	0.79	1.38

TABLE VI. Median of K-L divergence (in nats) over 5 GPS times for hBBH.

Network SNR	6	7	8	9	10	12
d_L (inc = 30°)	1.25	4.40	5.06	5.66	6.21	6.94
d_L (inc = 80°)	1.18	3.87	4.84	5.41	6.06	7.22
Time (inc = 30°)	0.49	2.62	2.69	2.93	2.88	3.43
Time (inc = 80°)	0.55	2.20	2.63	2.97	2.86	3.17
α (inc = 30°)	0.11	0.29	0.39	0.57	0.74	1.25
α (inc = 80°)	0.09	0.25	0.40	0.60	0.47	1.02
δ (inc = 30°)	0.04	0.40	0.49	0.75	0.98	1.44
δ (inc = 80°)	0.04	0.26	0.30	0.50	0.52	1.19
\mathcal{M} (inc = 30°)	0.53	1.63	1.74	2.90	3.25	3.35
\mathcal{M} (inc = 80°)	0.57	1.45	2.11	2.66	2.79	3.40
$\mathcal{M}^{\text{source}}$ (inc = 30°)	0.18	1.16	1.47	1.94	2.26	2.91
$\mathcal{M}^{\text{source}}$ (inc = 80°)	0.20	0.85	1.29	1.72	2.14	3.05
q (inc = 30°)	0.13	0.54	0.60	0.67	0.79	0.91
q (inc = 80°)	0.08	0.21	0.41	0.58	0.66	0.81
χ_{eff} (inc = 30°)	0.33	1.43	1.60	1.52	1.46	1.75
χ_{eff} (inc = 80°)	0.27	0.72	1.10	1.47	1.51	1.71
χ_p (inc = 30°)	0.02	0.08	0.10	0.10	0.10	0.11
χ_p (inc = 80°)	0.01	0.03	0.09	0.19	0.21	0.49

χ_p is notoriously hard to measure, even for louder events [6,8–11,76,120,121]. The tables below show that for weak events we nearly always recover the prior, which is why we have not reported χ_p uncertainties in the body of the paper.

-
- [1] G. M. Harry (LIGO Scientific Collaboration), *Classical Quantum Gravity* **27**, 084006 (2010).
- [2] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 074001 (2015).
- [3] F. Acernese *et al.* (Virgo Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).
- [4] B. P. Abbott *et al.* (LIGO Scientific Collaboration), *Living Rev. Relativity* **19**, 1 (2016).
- [5] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), arXiv:1811.12940.
- [6] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **6**, 041015 (2016).
- [7] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 061102 (2016).
- [8] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari *et al.*, *Phys. Rev. Lett.* **116**, 241103 (2016).
- [9] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya *et al.*, *Phys. Rev. Lett.* **118**, 221101 (2017).
- [10] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya *et al.*, *Astrophys. J.* **851**, L35 (2017).
- [11] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya *et al.*, *Phys. Rev. Lett.* **119**, 141101 (2017).
- [12] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya *et al.*, *Phys. Rev. Lett.* **119**, 161101 (2017).
- [13] B. P. Abbott *et al.* (Virgo and LIGO Scientific Collaborations), arXiv:1805.11579.
- [14] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J.* **818**, L22 (2016).
- [15] LIGO Scientific and Virgo Collaborations, *Astrophys. J. Lett.* **833**, L1 (2016).
- [16] S. Vitale, R. Lynch, R. Sturani, and P. Graff, *Classical Quantum Gravity* **34**, 03LT01 (2017).
- [17] B. Farr, D. E. Holz, and W. M. Farr, *Astrophys. J.* **854**, L9 (2018).
- [18] W. M. Farr, S. Stevenson, M. C. Miller, I. Mandel, B. Farr, and A. Vecchio, *Nature (London)* **548**, 426 (2017).
- [19] S. Stevenson, C. P. L. Berry, and I. Mandel, *Mon. Not. R. Astron. Soc.* **471**, 2801 (2017).
- [20] I. Mandel, W. M. Farr, A. Colonna, S. Stevenson, P. Tiño, and J. Veitch, *Mon. Not. R. Astron. Soc.* **465**, 3254 (2017).
- [21] S. Vitale, R. Lynch, J. Veitch, V. Raymond, and R. Sturani, *Phys. Rev. Lett.* **112**, 251101 (2014).
- [22] S. Vitale, R. Lynch, V. Raymond, R. Sturani, J. Veitch, and P. Graff, *Phys. Rev. D* **95**, 064053 (2017).
- [23] M. van der Sluys, I. Mandel, V. Raymond, V. Kalogera, C. Röver, and N. Christensen, *Classical Quantum Gravity* **26**, 204010 (2009).

- [24] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari *et al.*, *Phys. Rev. Lett.* **116**, 221101 (2016).
- [25] P. D. Lasky, E. Thrane, Y. Levin, J. Blackman, and Y. Chen, *Phys. Rev. Lett.* **117**, 061102 (2016).
- [26] A. Ghosh, A. Ghosh, N. K. Johnson-McDaniel, C. K. Mishra, P. Ajith, W. Del Pozzo, D. A. Nichols, Y. Chen, A. B. Nielsen, C. P. L. Berry, and L. London, *Phys. Rev. D* **94**, 021101 (2016).
- [27] N. Yunes, K. Yagi, and F. Pretorius, *Phys. Rev. D* **94**, 084002 (2016).
- [28] M. Agathos, W. Del Pozzo, T. G. F. Li, C. Van Den Broeck, J. Veitch, and S. Vitale, *Phys. Rev. D* **89**, 082001 (2014).
- [29] T. G. F. Li, W. Del Pozzo, S. Vitale, C. Van Den Broeck, M. Agathos, J. Veitch, K. Grover, T. Sidery, R. Sturani, and A. Vecchio, *Phys. Rev. D* **85**, 082003 (2012).
- [30] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya *et al.*, *Astrophys. J.* **848**, L12 (2017).
- [31] D. A. Coulter, R. J. Foley, C. D. Kilpatrick, M. R. Drout, A. L. Piro, B. J. Shappee, M. R. Siebert, J. D. Simon, N. Ulloa, D. Kasen, B. F. Madore, A. Murguia-Berthier, Y.-C. Pan, J. X. Prochaska, E. Ramirez-Ruiz, A. Rest, and C. Rojas-Bravo, *Science* **358**, 1556 (2017).
- [32] LIGO Scientific and Virgo Collaborations, *Astrophys. J. Lett.* **848**, L13 (2017).
- [33] M. Soares-Santos *et al.* (Dark Energy Survey and Dark Energy Camera GW-EM Collaboration), *Astrophys. J.* **848**, L16 (2017).
- [34] G. Hallinan *et al.*, *Science* **358**, 1579 (2017).
- [35] B. P. Abbott *et al.* (Virgo and LIGO Scientific Collaborations), *Phys. Rev. Lett.* **121**, 161101 (2018).
- [36] N. N. Weinberg (Virgo and LIGO Scientific Collaborations), [arXiv:1808.08676](https://arxiv.org/abs/1808.08676).
- [37] M. R. Drout *et al.*, *Science* **358**, 1570 (2017).
- [38] B. D. Metzger, T. A. Thompson, and E. Quataert, *Astrophys. J.* **856**, 101 (2018).
- [39] V. A. Villar, J. Guillochon, E. Berger, B. D. Metzger, P. S. Cowperthwaite, M. Nicholl, K. D. Alexander, P. K. Blanchard, R. Chornock, T. Eftekhari, W. Fong, R. Margutti, and P. K. G. Williams, *Astrophys. J.* **851**, L21 (2017).
- [40] P. A. Evans *et al.*, *Science* **358**, 1565 (2017).
- [41] LIGO Scientific, Virgo, 1M2H Collaboration, *Nature (London)* **551**, 85 (2017).
- [42] LIGO Scientific and Virgo Collaborations, *Astrophys. J. Lett.* **832**, L21 (2016).
- [43] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari *et al.*, *Living Rev. Relativity* **19** (2016).
- [44] B. F. Schutz, *Classical Quantum Gravity* **28**, 125023 (2011).
- [45] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Classical Quantum Gravity* **33**, 134001 (2016).
- [46] S. A. Usman *et al.*, *Classical Quantum Gravity* **33**, 215004 (2016).
- [47] K. Cannon, C. Hanna, and D. Keppel, *Phys. Rev. D* **88**, 024025 (2013).
- [48] J. Abadie *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **85**, 082002 (2012).
- [49] A. H. Nitz, T. Dent, T. Dal Canton, S. Fairhurst, and D. A. Brown, *Astrophys. J.* **849**, 118 (2017).
- [50] R. Lynch, M. Coughlin, S. Vitale, C. W. Stubbs, and E. Katsavounidis, *Astrophys. J.* **861**, L24 (2018).
- [51] M. Vallisneri, *Phys. Rev. D* **77**, 042001 (2008).
- [52] T. Damour, *Phys. Rev. D* **64**, 124013 (2001).
- [53] E. Racine, *Phys. Rev. D* **78**, 044021 (2008).
- [54] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, E. L. Robinson, J. Seiler, and B. Krishnan, *Phys. Rev. D* **82**, 064016 (2010).
- [55] P. Ajith, M. Hannam, S. Husa, Y. Chen, B. Brügmann, N. Dorband, D. Müller, F. Ohme, D. Pollney, C. Reisswig, L. Santamaría, and J. Seiler, *Phys. Rev. Lett.* **106**, 241101 (2011).
- [56] B. S. Sathyaprakash and B. F. Schutz, *Living Rev. Relativity* **12**, 2 (2009).
- [57] LIGO and Virgo Collaborations, <https://www.gw-openscience.org/events/> (2016).
- [58] K. Cannon, A. Chapman, C. Hanna, D. Keppel, A. C. Searle, and A. J. Weinstein, *Phys. Rev. D* **82**, 044025 (2010).
- [59] K. Cannon, C. Hanna, and D. Keppel, *Phys. Rev. D* **84**, 084003 (2011).
- [60] K. Cannon, C. Hanna, and J. Peoples, [arXiv:1504.04632](https://arxiv.org/abs/1504.04632).
- [61] C. Messick *et al.*, *Phys. Rev. D* **95**, 042001 (2017).
- [62] N. J. Cornish and T. B. Littenberg, *Classical Quantum Gravity* **32**, 135012 (2015).
- [63] É. É. Flanagan and T. Hinderer, *Phys. Rev. D* **77**, 021502 (2008).
- [64] L. Wade, J. D. E. Creighton, E. Ochsner, B. D. Lackey, B. F. Farr, T. B. Littenberg, and V. Raymond, *Phys. Rev. D* **89**, 103012 (2014).
- [65] J. M. Lattimer, *Annu. Rev. Nucl. Part. Sci.* **62**, 485 (2012).
- [66] K. Yagi and N. Yunes, *Phys. Rev. D* **89**, 021303 (2014).
- [67] P. Ajith, M. Hannam, S. Husa, Y. Chen, B. Brügmann, N. Dorband, D. Müller, F. Ohme, D. Pollney, C. Reisswig, L. Santamaría, and J. Seiler, *Phys. Rev. Lett.* **106**, 241101 (2011).
- [68] J. Aasi, J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. Abernathy, T. Accadia, F. Acernese, C. Adams, T. Adams *et al.*, *Phys. Rev. D* **88**, 062001 (2013).
- [69] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, *Phys. Rev. Lett.* **113**, 151101 (2014).
- [70] C. Cutler and É. E. Flanagan, *Phys. Rev. D* **49**, 2658 (1994).
- [71] A. H. Nitz, T. Dent, T. D. Canton, S. Fairhurst, and D. A. Brown, *Astrophys. J.* **849**, 118 (2017).
- [72] H.-Y. Chen, S. Vitale, and R. Narayan, [arXiv:1807.05226](https://arxiv.org/abs/1807.05226).
- [73] S. Vitale and H.-Y. Chen, *Phys. Rev. Lett.* **121**, 021303 (2018).
- [74] S. A. Usman, J. C. Mills, and S. Fairhurst, [arXiv:1809.10727](https://arxiv.org/abs/1809.10727).
- [75] J. Veitch *et al.*, *Phys. Rev. D* **91**, 042003 (2015).
- [76] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari *et al.*, *Phys. Rev. Lett.* **116**, 241102 (2016).

- [77] J. Veitch and A. Vecchio, *Phys. Rev. D* **81**, 062003 (2010).
- [78] R. Smith, S. E. Field, K. Blackburn, C.-J. Haster, M. Prrer, V. Raymond, and P. Schmidt, *Phys. Rev. D* **94**, 044031 (2016).
- [79] J. Skilling, in *American Institute of Physics Conference Series*, American Institute of Physics Conference Series, edited by R. Fischer, R. Preuss, and U. V. Toussaint, AIP Conference Proceedings Vol. 735 (American Institute of Physics, Melville, NY, 2004), pp. 395–405.
- [80] L. P. Singer, L. R. Price, B. Farr, A. L. Urban, C. Pankow, S. Vitale, J. Veitch, W. M. Farr, C. Hanna, K. Cannon, T. Downes, P. Graff, C.-J. Haster, I. Mandel, T. Sidery, and A. Vecchio, [arXiv:1404.5623](https://arxiv.org/abs/1404.5623).
- [81] K. Hotokezaka, S. Nissanke, G. Hallinan, T. J. W. Lazio, E. Nakar, and T. Piran, *Astrophys. J.* **831**, 190 (2016).
- [82] E. Nakar and T. Piran, *Nature (London)* **478**, 82 (2011).
- [83] N. T. Palliyaguru, A. Corsi, M. M. Kasliwal, S. B. Cenko, D. A. Frail, D. A. Perley, N. Mishra, L. P. Singer, A. Gal-Yam, P. E. Nugent, and J. A. Surace, *Astrophys. J.* **829**, L28 (2016).
- [84] P. J. Napier, *Very high angular resolution imaging: Proceedings of the 158th International Astronomical Union (IAU) Symposium, Australia, 1993*, edited by J. G. Robertson and W. J. Tango (Kluwer, Dordrecht, 1994), Vol. 158, p. 117.
- [85] S. W. Ellingson, T. E. Clarke, A. Cohen, J. Craig, N. E. Kassim, Y. Pihlstrom, L. J. Rickard, and G. B. Taylor, *Proc. IEEE* **97**, 1421 (2009).
- [86] M. G. Aartsen *et al.* (IceCube Collaboration), *J. Instrum.* **12**, P03012 (2017).
- [87] M. Ageron, J. A. Aguilar, I. Al Samarai, A. Albert, F. Ameli, M. André, M. Anghinolfi, G. Anton, S. Anvar, M. Ardid *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **656**, 11 (2011).
- [88] C. Meegan *et al.*, *Astrophys. J.* **702**, 791 (2009).
- [89] A. U. Abeyssekara *et al.*, *Astrophys. J.* **843**, 39 (2017).
- [90] J. L. Tonry, L. Denneau, A. N. Heinze, B. Stalder, K. W. Smith, S. J. Smartt, C. W. Stubbs, H. J. Weiland, and A. Rest, *Publ. Astron. Soc. Pac.* **130**, 064505 (2018).
- [91] R. M. Smith, R. G. Dekany, C. Bebek, E. Bellm, K. Bui, J. Cromer, P. Gardner, M. Hoff, S. Kaye, S. Kulkarni, A. Lambert, M. Levi, and D. Reiley, *Proc. SPIE Int. Soc. Opt. Eng.* **9147**, 914779 (2014).
- [92] R. Dekany, R. M. Smith, J. Belicki, A. Delacroix, G. Duggan, M. Feeney, D. Hale, S. Kaye, J. Milburn, P. Murphy, M. Porter, D. J. Reiley, R. L. Riddle, H. Rodriguez, and E. C. Bellm, *Proc. SPIE Int. Soc. Opt. Eng.* **9908**, 99085M (2016).
- [93] M. Actis, G. Agnetta, F. Aharonian, A. Akhperjanian, J. Aleksić, E. Aliu, D. Allan, I. Allekotte, F. Antico, L. A. Antonelli *et al.*, *Exp. Astron.* **32**, 193 (2011).
- [94] J. Najita *et al.*, [arXiv:1610.01661](https://arxiv.org/abs/1610.01661).
- [95] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, *Phys. Rev. D* **49**, 6274 (1994).
- [96] B. Farr, C. P. L. Berry, W. M. Farr, C.-J. Haster, H. Middleton, K. Cannon, P. B. Graff, C. Hanna, I. Mandel, C. Pankow, L. R. Price, T. Sidery, L. P. Singer, A. L. Urban, A. Vecchio, J. Veitch, and S. Vitale, *Astrophys. J.* **825**, 116 (2016).
- [97] J. Veitch, I. Mandel, B. Aylott, B. Farr, V. Raymond, C. Rodriguez, M. van der Sluys, V. Kalogera, and A. Vecchio, *Phys. Rev. D* **85**, 104045 (2012).
- [98] L. Blanchet, *Living Rev. Relativity* **9**, 4 (2006).
- [99] S. Vitale and M. Zanolin, *Phys. Rev. D* **82**, 124065 (2010).
- [100] M. Zanolin, S. Vitale, and N. Makris, *Phys. Rev. D* **81**, 124048 (2010).
- [101] K. G. Arun, B. R. Iyer, B. S. Sathyaprakash, and P. A. Sundararajan, *Phys. Rev. D* **71**, 084008 (2005).
- [102] SciPy.org, https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html (2018).
- [103] M. Pürrer, M. Hannam, and F. Ohme, *Phys. Rev. D* **93**, 084042 (2016).
- [104] K. K. Y. Ng, S. Vitale, A. Zimmerman, K. Chatziioannou, D. Gerosa, and C.-J. Haster, *Phys. Rev. D* **98**, 083007 (2018).
- [105] C. Talbot and E. Thrane, *Phys. Rev. D* **96**, 023012 (2017).
- [106] D. Gerosa, M. Kesden, E. Berti, R. O’Shaughnessy, and U. Sperhake, *Phys. Rev. D* **87**, 104028 (2013).
- [107] D. J. MacKay, *Information Theory, Inference, and Learning Algorithms*, 1st ed. (Cambridge University Press, Cambridge, England, 2003), p. 34.
- [108] P. Schmidt, F. Ohme, and M. Hannam, *Phys. Rev. D* **91**, 024043 (2015).
- [109] LIGO Scientific and Virgo Collaborations, *Classical Quantum Gravity* **35**, 065010 (2018).
- [110] S. M. Gaebel, J. Veitch, T. Dent, and W. M. Farr, [arXiv:1809.03815](https://arxiv.org/abs/1809.03815).
- [111] S. Fairhurst, *New J. Phys.* **11**, 123006 (2009).
- [112] L. P. Singer, L. R. Price, B. Farr, A. L. Urban, C. Pankow, S. Vitale, J. Veitch, W. M. Farr, C. Hanna, K. Cannon, T. Downes, P. Graff, C.-J. Haster, I. Mandel, T. Sidery, and A. Vecchio, *Astrophys. J.* **795**, 105 (2014).
- [113] T. Sidery *et al.*, *Phys. Rev. D* **89**, 084060 (2014).
- [114] D. E. Holz and S. A. Hughes, *Astrophys. J.* **629**, 15 (2005).
- [115] S. Nissanke, D. E. Holz, N. Dalal, S. A. Hughes, J. L. Sievers, and C. M. Hirata, [arXiv:1307.2638](https://arxiv.org/abs/1307.2638).
- [116] H.-Y. Chen, M. Fishbach, and D. E. Holz, *Nature (London)* **562**, 545 (2018).
- [117] B. F. Schutz, *Nature (London)* **323**, 310 (1986).
- [118] W. Del Pozzo, *Phys. Rev. D* **86**, 043011 (2012).
- [119] Y. Huang, H. Middleton, K. K. Y. Ng, S. Vitale, and J. Veitch, Data release for characterization of low-significance gravitational-wave compact binary sources (2018), DOI: [10.5281/zenodo.1601151](https://doi.org/10.5281/zenodo.1601151).
- [120] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **6**, 041014 (2016).
- [121] S. Vitale, D. Gerosa, C.-J. Haster, K. Chatziioannou, and A. Zimmerman, *Phys. Rev. Lett.* **119**, 251103 (2017).