UNIVERSITY OF BIRMINGHAM

Assessing gravitational-wave binary black hole candidates with Bayesian odds

Pratten, Geraint; Vecchio, Alberto

Document Version
Publisher's PDF, also known as Version of record

[Link to publication on Research at Birmingham portal](#)

Assessing gravitational-wave binary black hole candidates with Bayesian odds

Geraint Pratten[*] and Alberto Vecchio[†]

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

Gravitational waves from the coalescence of binary black holes can be distinguished from noise transients in a detector network through Bayesian model selection by exploiting the coherence of the signal across the network. We present a Bayesian framework for calculating the posterior probability that a signal is of astrophysical origin, agnostic to the specific search strategy, pipeline or search domain with which a candidate is identified. We apply this framework under *identical* assumptions to all events reported in the LIGO-Virgo GWTC-1 catalog, GW190412 and numerous event candidates reported by independent search pipelines by other authors. With the exception of GW170818, we find that all GWTC-1 candidates, and GW190412, have odds overwhelmingly in favor of the astrophysical hypothesis, including GW170729, which was assigned significantly different astrophysical probabilities by the different search pipelines used in GWTC-1. GW170818 is *de facto* a single detector trigger, and is therefore of no surprise that it is disfavored as being produced by an astrophysical source in our framework. We find *three* additional event candidates, GW170121, GW170425 and GW170727, that have significant support for the astrophysical hypothesis, with a probability that the signal is of astrophysical origin of 0.53, 0.74 and 0.64, respectively. We carry out a hierarchical population study which includes these three events in addition to those reported in GWTC-1, finding that the main astrophysical results are unaffected.

I. INTRODUCTION

Gravitational-wave (GW) astronomy is having a profound impact on our understanding of the fundamental nature of astrophysial binary black holes and the properties of the underlying population. During the first (O1) and second (O2) observing runs of Advanced LIGO [1] and Advanced Virgo [2], ten unambiguous binary black hole (BBH) mergers were reported as part of the first gravitational-wave transient catalog (GWTC-1) [3]. Since then, over 25 additional BBH candidates have been reported [4–10]. This population is expected to grow considerably starting with the events observed during the third observing run (O3), with two confident detections having already been announced [11,12].

One of the key goals of GW searches is to determine which of the detection candidates are produced by astrophysical compact binaries. This result is often condensed into a single quantity referred to as p_{astro}, the probability that a transient signal is of astrophysical origin [13–15]. One of the key limiting factors in our ability to confidently identify such astrophysical binaries is the presence of instrumental noise transients (glitches) that contaminate

the data. Noise transients can mimic astrophysical signals, impeding the statistical significance to which we can detect GW signals. Advanced LIGO and Advanced Virgo are instruments of exquisite sensitivity with many of the significant detections being observed with signal-to-noise ratios (SNRs), ρ, so large that p_{astro} is effectively indistinguishable from unity. This is, for example, the case for GW150914 [16,17], the very first BBH merger observed, as well as all but one (GW170729) of the events reported in GWTC-1.

Under the assumption that coalescing binary systems are distributed uniformly in volume, and neglecting cosmological evolution of the merger rate, the expected distribution of signal-to-noise ratio scales as $\sim \rho^{-4}$ [18,19], but the distribution of background and transient instrumental signals also rises steeply as one goes to progressively smaller values of ρ to identify event candidates at "threshold" [20–24]. As modeled search pipelines, currently based on the same underlying technique of matched filtering, battle to identify quiet signals overwhelmed by noise, different choices are made in different studies: for example cuts bases on data quality, signal consistency tests, power spectral density estimations and physical search parameter domains (e.g., binary mass range covered by a search).

The astrophysical probability p_{astro} introduced in [13–15], is a Bayesian odds comparing the astrophysical and

*g.pratten@bham.ac.uk
†av@star.sr.bham.ac.uk

terrestrial hypotheses. This method estimates the joint posterior on the Poisson expected counts for an arbitrary choice of foreground categories, such as astrophysical BBHs. A key limitation, however, is that the method appeals to bootstrap techniques in order to characterize the noise model for a specific search pipeline [25–28]. As a

consequence of different choices made by different search pipelines, the p_{astro} assigned to a same signal can be widely discrepant depending on the pipeline and/or search employed, in particular for low signal-to-noise events. This is the case for GW170729, which was assigned a p_{astro} of 0.98 by GstLAL [29,30] and 0.52 by PyCBC

TABLE I. Table summarizing the astrophysical probability $P(\mathcal{S}|\mathbf{d})$ and the Bayes factors computed according to different assumptions: $\mathcal{O}_{S/N}$ in Eq. (3), $\mathcal{B}_{cr}$ in Eq. (12) using $\{\alpha = 10^{-6}, \beta_k = 10^{-4}\}$, $\mathcal{B}_{s/n}$ defined in Eq. (6), $\mathcal{B}_{c/i}$ in Eq. (8) and $\mathcal{B}_{c/(i.\vee n.)}$ in Eq. (11). For $\mathcal{O}_{S/N}$ we use $P(\mathcal{S}) = 6.43 \times 10^{-7}$ and $P(\mathcal{G}_k|\mathbf{d})$ as estimated from the OMICRON triggers in a 24h window around each event and reported in Table II. For GW190412 [11] we adopt $P(\mathcal{G}_k|\mathcal{N}) = 10^{-4}$ as a fiducial value, though this event is unambiguous with $P(\mathcal{S}|\mathbf{d}) \sim 1$ irrespective of the glitch prior, assuming reasonable values. The purple shaded rows denote the significant events reported in GWTC-1 [3] and the blue shaded rows denote the new event candidates with $P(\mathcal{S}|\mathbf{d}) > 0.5$.

Event	ρ^N_{MF}	p_{astro}	$P(\mathcal{S}|\mathbf{d})$	$\log_{10}\mathcal{O}_{s/n}$	$\log_{10}\mathcal{B}_{cr}$	$\log_{10}\mathcal{B}_{s/n}$	$\log_{10}\mathcal{B}_{c/i}$	$\log_{10}\mathcal{B}_{c/(i\vee n)}$
GW150914	25.01	0.99[b], 1.00[a]	1.00	7.86	8.87	121.10	6.87	6.87
GW151012	9.63	0.97[b], 0.96[a]	0.99	2.02	2.24	8.25	5.51	5.51
GW151226	12.71	0.88[b], 1.00[a]	1.00	6.02	7.19	18.45	8.96	8.96
GW170104	14.01	1.00[b], 1.00[a]	1.00	6.75	8.38	30.19	6.38	6.38
GW170608	15.62	0.92[b], 1.00[a]	1.00	8.32	9.79	34.70	7.79	7.79
GW170729	10.62	0.98[b], 0.52[a]	1.00	3.46	4.65	15.52	2.66	2.66
GW170809	12.82	0.99[b], 1.00[a]	1.00	3.77	4.45	23.44	3.09	3.09
GW170814	16.66	1.00[b], 1.00[a]	1.00	6.86	7.80	46.32	5.80	5.80
GW170818	11.34	0.99[b]	0.10	−0.94	−0.39	15.58	1.56	1.56
GW170823	12.04	0.99[b], 1.00[a]	1.00	5.30	5.89	21.72	4.21	4.21
GW151011	7.46	0.08[c]	0.01	−1.93	−1.71	4.30	1.56	1.56
GW151124	8.67	0.16[d]	0.00	−5.69	−4.66	5.10	−2.38	−2.38
GW151205	6.81	0.53[c]	0.03	−1.53	−1.33	4.67	2.06	2.05
GW151216	8.16	0.18[c]	0.00	−2.35	−2.16	3.84	2.96	2.91
GW151216A	8.48	0.07[c]	0.00	−6.02	−5.82	0.18	0.34	−0.05
GW151217	7.99	0.26[c]	0.00	−6.04	−5.85	0.15	0.22	−0.12
GW151222	10.22	0.03[d]	0.00	−4.68	−3.52	9.34	−1.65	−1.65
GW170104A	7.36	0.12[c]	0.00	−4.82	−4.63	1.37	0.97	0.83
GW170106	9.62	0.01[d]	0.00	−6.35	−4.96	8.19	−2.88	−2.88
GW170121	10.55	1.00[c], 1.00[e]	0.53	0.05	1.78	12.26	3.38	3.38
GW170123	6.20	0.08[c]	0.00	−3.86	−3.66	2.34	1.70	1.61
GW170201	8.25	0.24[c]	0.00	−2.55	−2.36	3.64	2.84	2.77
GW170202	8.50	0.13[c], 0.68[e]	0.12	−0.86	−0.67	5.33	3.54	3.54
GW170220	6.59	0.10[c]	0.00	−4.01	−3.82	2.18	1.06	1.03
GW170304	8.50	0.70[c], 0.99[e]	0.03	−1.52	−0.31	7.53	1.24	1.24
GW170402	9.00	0.68[e]	0.00	−3.55	−2.52	6.16	−0.53	−0.53
GW170403	8.20	0.03[c], 0.56[e]	0.27	−0.42	−0.21	5.80	2.45	2.45
GW170425	7.99	0.21[c], 0.77[e]	0.74	0.46	0.72	6.74	2.44	2.43
GW170620	8.36	0.02[c]	0.00	−2.81	−2.62	3.38	2.94	2.80
GW170629	7.56	0.02[c]	0.00	−6.25	−6.06	−0.06	0.36	−0.20
GW170721	8.55	0.06	0.15	−0.75	−0.39	5.66	2.35	2.35
GW170724	10.76	0.02[d]	0.00	−6.77	−5.68	13.46	−3.48	−3.48
GW170727	9.90	0.99[c], 0.98[e]	0.66	0.28	0.87	9.15	2.71	2.71
GW170729A	11.55	0.05[d]	0.00	−4.95	−4.24	10.50	−2.93	−2.93
GW170801	8.14	…	0.00	−4.13	−3.70	2.69	−1.53	−1.53
GW170817A	10.65	0.86[e]	0.02	−1.66	−1.26	16.62	0.63	0.63
GW190412	18.47	…	1.00	7.94	8.14	59.27	6.14	6.14

[a]The p_{astro} values are taken from the following search pipeline: PyCBC [3].
[b]The p_{astro} values are taken from the following search pipeline: GstLAL [3].
[c]The p_{astro} values are taken from the following search pipeline: [4].
[d]The p_{astro} values are taken from the following search pipeline: [5].
[e]The p_{astro} values are taken from the following search pipeline: [6,10].

FIG. 1. $P(\mathcal{S}|\mathbf{d})$ as a function of $P(\mathcal{S})$. The probability of the glitch hypothesis $P(\mathcal{G}_k|\mathcal{N})$ for each detector is estimated using the trigger rate in 24h of data around each event. The horizontal dashed line denotes the threshold for inclusion in the population analysis, $P(\mathcal{S}|\mathbf{d}) > 0.5$. The vertical dashed lines denote the median and 90% CI for the signal hypothesis $P(\mathcal{S})$ as inferred from the uncertainty in the local astrophysical merger rate $\mathcal{R}_0$.

[25,26], as reported in GWTC-1 [3]. Possibly more disturbing is that the region of parameter space on which a search is performed, i.e., the BBH template bank versus the full template bank covering the whole neutron star and black hole mass range, can lead to different values of p_{astro} for the same event, to the extent that some signals cross the (arbitrary) threshold of detection in one search and not the other. Although the reason for this behavior is understood, it is important to understand whether p_{astro} is a sufficiently informative and robust characterization of the probability that detection candidates identified by a search are of astrophysical origin, and whether or not they are included in further studies, either of the fundamental physics properties of black holes [31–33] or the underlying astrophysical properties of the population [32,34,35].

Here we reconsider the problem of assigning a measure of the likelihood of a signal being of astrophysical origin by considering a different statistical quantity: the posterior odds of a transient candidate being produced by an astrophysical source (specifically a BBH described by general relativity) versus random transient noise in the detectors' data, which could arise from instrumental glitches and/or Gaussian noise fluctuations. From this quantity we can trivially evaluate the posterior probability of a signal being of astrophysical origin. Our fully Bayesian framework provides a single, unified measure of the astrophysical probability that is crucially agnostic to the specific search strategy, pipeline implementation or search domain, with which a candidate is identified. Our approach builds on several studies and preliminary investigations [27,28,36–38] and critically takes advantage of observations carried out with a network of detectors: a signal produced by an astrophysical source must appear in the data of the detectors in the network with consistent astrophysical parameters, arrival times, signal strength and so on.

Using *identical* assumptions for each event, we analyze all BBH candidates reported in GWTC-1 [3] and those reported by other search pipelines, notably the PyCBC 2-OGC catalog [4] and the IAS search [10]. We also analyze GW190412, the first BBH detected during the LIGO-Virgo third observing run [11]. The main results are summarized in Table I and Figs. 1–3. We find that all significant BBHs reported in GWTC-1, modulo the single detector event GW170818, and GW190412 have odds overwhelmingly in favor of the astrophysical hypothesis. For all the other event candidates, the p_{astro} reported by the searches is not a monotonic function of the odds. Three candidate events [10], GW170121, GW170727 and GW170425 have odds that favor the astrophysical origin hypothesis and that are sufficient to be included in typical population studies. Interestingly GW170425, which was originally reported with $p_{astro} = 0.21$, has odds comparable to GW170727, which was reported with $p_{astro} = 0.99$. In order to gauge the astrophysical implications of these new events, we perform a hierarchical population analysis incorporating these new binaries, finding that the main astrophysical results reported in [35] are unchanged. Whilst we focus on BBH coalescences in this paper, the framework can be trivially extended to an arbitrary choice of foreground categories.

This paper is organized as follows: in Sec. II we define the model hypotheses and define the two key quantities in our analysis, the posterior odds ratio $\mathcal{O}_{S/N}$ and the probability that the signal is of astrophysical origin $P(\mathcal{S}|\mathbf{d})$. We then discuss the choice of astrophysical signal and glitch priors followed by details regarding the calculation of the evidences and Bayes factors. Section III presents the core results of the paper and Sec. IV discusses the implications for population inference. We conclude in Sec. V.

II. ASTROPHYSICAL SIGNAL ODDS

The end point of our analysis is to evaluate the posterior odds ratio, $\mathcal{O}_{S/N}$, that an event candidate is due to an astrophysical source—a BBH merger—versus a transient noise fluctuation, and as a consequence the posterior probability, $P(\mathcal{S}|\mathbf{d})$, that an astrophysical signal is present in the data.

Hence, given the data $\mathbf{d}$, we wish to compute

$$\mathcal{O}_{S/N} = \frac{P(\mathcal{S}|\mathbf{d})}{P(\mathcal{N}|\mathbf{d})}, \qquad (1)$$

where $P(\mathcal{S}|\mathbf{d})$ and $P(\mathcal{N}|\mathbf{d})$ are the posterior probability of the signal hypothesis, $\mathcal{S}$, and noise hypothesis, $\mathcal{N}$, respectively. The data $\mathbf{d} = \{d_k; k = 1, \ldots N_d\}$ comprise the collection of datasets, d_k, from N_d detectors in the network.

As the signal and noise hypotheses are exhaustive, we can use Eq. (1) to determine the probability that the signal is of astrophysical origin

$$P(\mathcal{S}|\mathbf{d}) = \frac{\mathcal{O}_{S/N}}{1 + \mathcal{O}_{S/N}}, \qquad (2)$$

where Eq. (2) is the quantity returned by our analysis that plays a similar role to the p_{astro} computed from the distribution of background and foreground events produced by the search pipelines [6,13,14,17]. However, $\mathcal{O}_{S/N}$ and $P(\mathcal{S}|\mathbf{d})$ are completely independent of the search pipeline used to identify a candidate, and only depend on the model hypotheses and the associated prior probabilities. They are derived by naturally using a *coherent* analysis of the data— as opposed to searches, which rely on identifying triggers in *coincidence*—and can implement the best signal and noise model at one's disposal.

We note that similar strategies for computing Eq. (2) have been pursued in [27,28]. However, the signal odds, as defined here, does not rely on the marginalization over a glitch hyper model using contextual data [27,28] nor is it treated as a traditional detection statistic to obtain a frequentist estimate of the significance of an event given the measured background, as in [37].

A. Notation and assumptions

Here we summarize the assumptions for the computation of Eqs. (1) and (2) and summarize our notation, which closely follows [28,36–38], to which we refer the reader for further details.

First, we define the models that we consider:

(i) Signal model or hypothesis, $\mathcal{S}$: *there is an astrophysical signal due to a BBH coalescence in the detector network.* We also make the additional assumption that no transient of instrumental nature—a "glitch"—takes place at the same time of the GW signal. This is an excellent approximation

in the case of BBHs, the focus of this study, at current instrument sensitivity, as GWs from BBH coalescences are in the instruments' bandwidth for at best a few seconds. The signal model corresponds therefore to $\mathcal{S} = \mathcal{S}_1 \wedge \ldots \wedge \mathcal{S}_{N_d}$, where $\wedge$ is the logical "and."

(ii) Noise model or hypothesis, $\mathcal{N}$: *there is no astrophysical signal in the data, just noise.* To an excellent approximation, we can further assume that the noise between two detectors is uncorrelated, see [39] and references therein. We assume that the noise in each of the instruments could be due to a glitch—hypothesis $\mathcal{G}_k$—or simply Gaussian stationary noise, a model that we identify with $\mathscr{G}_k$. Therefore $\mathcal{G}_k$ and $\mathscr{G}_k$ are exhaustive and disjoint noise hypotheses. We also make an important additional assumption: the glitch model ($\mathcal{G}_k$) for each individual detector is the most conservative one in which glitches are modeled as having the same functional form as a GW from a BBH coalescence with uncorrelated parameters in each detector, $\mathcal{G}_k = \mathcal{S}_k$ [36]. Under this assumption, the noise model in each instrument is $\mathcal{N}_k = \mathscr{S}_k \wedge (\mathcal{G}_k \vee \mathscr{G}_k)$, where $\vee$ is the logical "or." The noise hypothesis can then be written as $\mathcal{N} = \mathcal{N}_1 \wedge \ldots \wedge \mathcal{N}_{N_d}$. We note that whilst a more exhaustive treatment of noise would account for both non-Gaussianity and nonstationarity in the data [7,40], the search pipelines do use more complex noise models that account for such behavior in calculating the false alarm rate (FAR) of event candidates and hence in determining the list of events analyzed here [5,7,26,30,41].

We can now express the odds ratio, Eq. (1), in terms of the prior probability of the signal $P(\mathcal{S})$ and noise $P(\mathcal{N})$ hypotheses, and of the marginal likelihoods or evidences, $P(\mathbf{d}|\mathcal{S})$ and $P(\mathbf{d}|\mathcal{N})$, of the same models [36,37]:

$$\begin{aligned} \mathcal{O}_{S/N} &= \left[\frac{P(\mathcal{S})}{P(\mathcal{N})}\right] \frac{P(\mathbf{d}|\mathcal{S})}{P(\mathbf{d}|\mathcal{N})}, \\ &= \left[\frac{P(\mathcal{S})}{P(\mathcal{N})}\right] \frac{\mathcal{B}_{s/n}}{\prod_k \{P(\mathcal{G}_k|\mathcal{N})\mathcal{B}_{s/n}^{(k)} + [1 - P(\mathcal{G}_k|\mathcal{N})]\}}. \end{aligned}$$
$$(3)$$

Here we have defined the single-detector Bayes factor as

$$\mathcal{B}_{s/n}^{(k)} \equiv \frac{P(d_k|\mathcal{S}_k)}{P(d_k|\mathscr{G}_k)}, \qquad (4)$$

and the Bayes factor for a coherent GW signal across the network embedded in stationary Gaussian noise as

$$\mathcal{B}_{s/n} \equiv \frac{P(\mathbf{d}|\mathcal{S})}{P(\mathbf{d}|\mathcal{G})}, \qquad (5)$$

where $P(\mathbf{d}|\mathcal{G}) = \prod_k P(d_k|\mathscr{G}_k)$.

There are a number of limiting cases of Eq. (3) that have been discussed in the literature and used in GW data analysis. First, if we assume the case of a perfect instrument, where the noise is Gaussian and stationary and no glitches occur, $\mathcal{N} = \mathcal{G}$, hence $P(\mathcal{G}_k|\mathcal{N}) = 0$. Consequentially, Eq. (3) reduces to

$$\mathcal{O}_{S/N} = \left[\frac{P(S)}{P(\mathcal{N})}\right] \mathcal{B}_{s/n}, \qquad (6)$$

which reduces to the signal vs Gaussian noise Bayes factor in the limit $P(S) = P(\mathcal{N})$. If we now assume a fundamentally flawed detector, in which the glitch rate is sufficiently high that it can be well approximated by $P(\mathcal{G}_k|\mathcal{N}) = 1$, then Eq. (3) becomes

$$\mathcal{O}_{S/N} = \left[\frac{P(S)}{P(\mathcal{N})}\right] \frac{\mathcal{B}_{s/n}}{\prod_k \mathcal{B}_{s/n}^{(k)}}, \qquad (7)$$

which reduces to the "coherent vs incoherent" Bayes factor [36]

$$\mathcal{B}_{c/i} \equiv \frac{\mathcal{B}_{s/n}}{\prod_k \mathcal{B}_{s/n}^{(k)}}, \qquad (8)$$

under the assumption that $P(S) = P(\mathcal{N})$. If we next assume that

$$P(\mathcal{G}_k|\mathcal{N}) = 0.5 \quad \forall\, k, \qquad (9)$$

then the odds ratio reduces to

$$\mathcal{O}_{S/N} = \left[\frac{P(S)}{P(\mathcal{N})}\right] \frac{2\mathcal{B}_{s/n}}{\prod_k (\mathcal{B}_{s/n}^{(k)} + 1)}, \qquad (10)$$

which is related to the "coherent vs incoherent or noise" Bayes factor

$$\mathcal{B}_{c/(i\vee n)} \equiv \frac{\mathcal{B}_{s/n}}{\prod_k (\mathcal{B}_{s/n}^{(k)} + 1)}. \qquad (11)$$

Finally, if one sets $\alpha \equiv P(S)/P(\mathcal{N}) \simeq P(S) = 10^{-6}$ and $\beta = P(\mathcal{G}_k|\mathcal{N}) = 10^{-4}$ ($\forall\, k$), the odds ratio reduces to the "Bayesian coherence ratio" introduced in [37],

$$\mathcal{B}_{cr} = \alpha \frac{\mathcal{B}_{s/n}}{\prod_k \{\beta \mathcal{B}_{s/n}^{(k)} + [1 - \beta]\}}, \qquad (12)$$

where the values of α and β were chosen through an injection campaign aimed at separating the foreground and background populations.

B. Priors for signal and glitch models

An important aspect of our framework is the choice of model priors, which affect the results in a straightforward way, see Eq. (3). The signal odds scales linearly with the prior belief of an astrophysical signal being present in a given segment of data, $P(S)/P(\mathcal{N})$, while the glitch probability acts as a weighting factor for the presence of an uncorrelated transient in each of the detectors in the network.

As discussed in Sec. II C, for each of the event candidates we select an 8s data segment containing the putative signal. We set a uniform prior on the coalescence time, t_c, to search for a BBH merger occurring in an interval $\Delta t_c = 0.2$ s around the GPS time of the reported event candidate. As BBH mergers are rare [3], we can assume that $P(S) \ll 1$. Leveraging on prior knowledge of the BBH merger rate, we set $P(S)$ to be the probability of a coalescence occurring in an interval Δt_c and be produced by a binary within the sensitive spacetime volume for all signals that yield a single interferometer SNR $\rho \geq 7$ for the typical detector sensitivities throughout the observing period. The SNR threshold used here is slightly lower than the more conventional $\rho > 8$ [42] as we are particularly interested in signals at the detection threshold, whilst avoiding the steep rise in the background of glitches. Using these assumptions, we have

$$P(S) = 6.43 \times 10^{-7} \left(\frac{\mathcal{R}_0}{53.2 \text{ Gpc}^{-3} \text{ y}^{-1}}\right) \left(\frac{\Delta t_c}{0.2 \text{ s}}\right), \quad (13)$$

where $\mathcal{R}_0$ is the local merger rate of BBHs; we provide further details about the derivation of this result in the Appendix A.

One may wonder whether using the value of $\mathcal{R}_0$ derived from the analysis of all the data in GWTC-1 is formally consistent with determining $P(S)$ when analyzing all other event candidates identified during the same observing period. A more formally correct strategy, which we plan to implement in future analyses, would be to divide the data into segments, e.g., 1-week long, and based on the identified candidates simultaneously fit both the local merger rate and the astrophysical probability of each event candidate. In practice, however, the results presented here would be unaffected. The estimate of the BBH merger rate determined using the initial 16 days surrounding GW150914, which would be a natural starting point for any iterative analysis, is consistent with $\mathcal{R}_0$ measured at the end of O2 to within a factor ≈ 2, though the uncertainty is significantly reduced. Our results, as described in Sec. III, are robust against such variations in $P(S)$, see in particular Fig. 1.

As the models S and $\mathcal{N}$ are exhaustive hypotheses, $P(S) + P(\mathcal{N}) = 1$, the prior ratio in Eq. (3) can be written as

$$\frac{P(\mathcal{S})}{P(\mathcal{N})} = \frac{P(\mathcal{S})}{1 - P(\mathcal{S})} \simeq P(\mathcal{S}), \qquad (14)$$

to a very good approximation. The effect of $P(\mathcal{S})$ on $\mathcal{O}_{S/N}$ is to act as an overall normalization, with the odds, and hence the posterior astrophysical signal probability, scaling linearly with the total astrophysical merger rate integrated over the entire binary population, as in Eq. (13).

In order to compute Eq. (3), we also need an estimate of the probability that a glitch takes place within the same prior coalescence interval, $P(\mathcal{G}_k|\mathcal{N})$. This is a quantity whose value changes during the course of a run and can differ between each of the detectors in the network. Here we determine the glitch prior by considering all glitches identified by the OMICRON pipeline [43–45]. Such an approach was used in O1 to estimate the rate of glitches with SNR > 5 [21] and more recently in [46]. For each detector, we take a 24 hour period centered around the time of the putative candidate and estimate the number of triggers, $N_k^{(\mathrm{OM})}$ occurring within the frequency range covered by our analysis, 20–1024 Hz, and which also yield a SNR ≥ 7, consistent with our choice for the signal prior. We set the prior glitch probability for each instrument to

$$P(\mathcal{G}_k|\mathcal{N}) = N_k^{(\mathrm{OM})} \frac{\Delta t_c}{T_k^{(\mathrm{live})}}, \qquad (15)$$

where $T_k^{(\mathrm{live})}$ is the live time for instrument k during the 24-hour period. Note that transients producing a given signal-to-noise ratio when analyzed with the OMICRON pipeline would in general yield a lower signal-to-noise ratio when using a BBH waveform; our approach is therefore conservative with respect to the chance that the glitches mimic the GW signal from an astrophysical BBH. The values of $P(\mathcal{G}_k|\mathcal{N})$ adopted in our analysis are reported in Table II and are consistent with previous studies of the glitch rates, e.g., [20–22,24]. Our choice of the signal $P(\mathcal{S})$ and glitch prior $P(\mathcal{G}_k|\mathcal{N})$ is also broadly consistent with values for related quantities considered in other studies [37,46].

Finally, we assume that each data segment can be approximated by stationary, Gaussian noise. A more exhaustive treatment of noise would account for both nonstationarity and non-Gaussianity, but this is beyond the scope of the paper. We note, however, that the list of triggers we analyze is generated by search pipelines that implement more complex noise models [7,25,26,29,47], allowing the pipelines to down-rank triggers accordingly. As such, the final list of triggers has a higher purity level than if we were to perform a full Bayesian search using our framework alone. As noise triggers are typically uncorrelated between detectors, we expect that non-Gaussian transients will be penalized in our framework. The

robustness of our framework to noise transients is further discussed in Sec. III D. Our framework will explicitly benefit from further improvements to noise modeling in search pipelines [7,41,48], and will benefit from significantly improving the noise model beyond the approximations used in this analysis.

C. Evidence and Bayes factor evaluation

The last quantities we need to evaluate Eqs. (2) and (3) are the signal evidences, and therefore the Bayes factors. For this, we follow the approach, and use the same software, adopted for the analyses of the BBHs reported in GWTC-1 [3].

For each event, we use data from the Gravitational Wave Open Science Center [49,50] and analyze 8s of data around the time of the candidate and within a frequency range of 20–1024 Hz. For simplicity, we restrict our analysis to a 2 detector network, using only the data from Hanford and Livingston, the two most sensitive instruments.[1] The power spectral densities (PSDs) are estimated using the BAYESWAVE algorithm [51,52], which only uses data from the segment that is analyzed, and we marginalize over calibration uncertainty [53–55] using the approximate uncertainty reported in GWTC-1 [3]. Whilst we do not include marginalization over PSD uncertainty when performing parameter estimation [56–59], the BAYESWAVE algorithm does marginalize over uncertainty when estimating the *on-source* PSD. Uncertainty in the PSD will be partially coupled to the uncertainty in the detector calibration, with both terms corresponding to a frequency and detector dependent uncertainty in the amplitude response [53].

An alternative approach, as taken in [46], would be to estimate the PSD using *off-source* methods that are predicated on taking the median- or mean-average of neighboring nonoverlapping segments [60,61]. Such an off-source approach is more susceptible to non-Gaussian transients and nonstationarity in the data [7,25,51,62]. In addition, it is known that Gaussian noise that is whitened using an *off-source* PSD will not follow a unit normal distribution [56,59,63–65]. For a mean-averaged PSD, whitened Gaussian noise can be shown to instead follow a Student's t distribution and the correct likelihood that one should use is given by the Student-Rayleigh distribution [59,63–65]. For a median-averaged off-source PSD, an analogous statement can be shown to hold and one must

---

[1]The BBH candidates for which Virgo was also in science mode are these: GW170729, GW170729A, GW170809, GW170814, GW170801, GW170817A and GW170818. Our analysis is completely general and can be trivially (though rather costly in terms of computer time) extended to a network consisting of an arbitrary number of instruments. We leave this to future work, noting that during the first two observing runs Virgo was sufficiently less sensitive than LIGO that the main conclusions presented here are unaffected.

TABLE II.   The first three columns show the network signal-to-noise ratio and single detector signal-to-noise ratio (the indices "L" and "H" stand for LIGO-Livingston and LIGO-Hanford, respectively) for all the events considered in this analysis, see Table I. The next three columns show the Bayes factors for the signal vs noise hypothesis for the detector network, Eq. (5), and the single detectors, Eq. (4), respectively. The last two columns report the glitch probabilities as estimated from OMICRON triggers in a 24h stretch of data near each trigger. The glitch rates are determined by down-selecting OMICRON triggers based on a single detector signal-to-noise ratio threshold of $\rho > 7$ and by restricting the frequencies to $20 < f_{OM} < 1024$ Hz, see Eq. (15) and discussion in Sec. II B.

Event	ρ_{MF}^{Net}	ρ_{MF}^{H}	ρ_{MF}^{L}	$\log_{10}\mathcal{B}_{s/n}$	$\log_{10}\mathcal{B}_{s/n}^{(H)}$	$\log_{10}\mathcal{B}_{s/n}^{(L)}$	$P(\mathcal{G}_H|\mathcal{N})$	$P(\mathcal{G}_L|\mathcal{N})$
GW150914	25.01	20.30	14.71	121.1	77.77	36.45	5.35×10^{-4}	1.23×10^{-4}
GW151012	9.63	7.23	6.38	8.25	1.97	0.77	8.73×10^{-4}	3.61×10^{-4}
GW151226	12.71	10.30	7.35	18.45	9.26	0.23	9.42×10^{-4}	1.74×10^{-4}
GW170104	14.01	9.75	10.27	30.19	10.96	12.85	6.16×10^{-4}	4.42×10^{-4}
GW170608	15.62	12.68	9.82	34.7	19.45	7.46	5.00×10^{-4}	3.82×10^{-4}
GW170729	10.62	9.15	7.34	15.52	7.4	5.46	3.91×10^{-4}	2.62×10^{-4}
GW170809	12.82	7.77	11.01	23.44	3.47	16.88	3.36×10^{-4}	1.99×10^{-4}
GW170814	16.66	9.27	14.08	46.32	8.96	31.56	2.50×10^{-4}	2.22×10^{-4}
GW170818	11.34	3.70	10.58	15.58	0.05	13.97	2.96×10^{-4}	2.27×10^{-4}
GW170823	12.04	7.19	9.93	21.72	3.97	13.54	2.06×10^{-4}	1.67×10^{-4}
GW151011	7.46	4.05	6.18	4.3	0.28	2.45	8.73×10^{-4}	3.31×10^{-4}
GW151124	8.67	4.99	8.80	5.1	−0.28	7.76	4.40×10^{-4}	6.90×10^{-4}
GW151205	6.81	5.43	4.77	4.67	1.57	1.04	6.81×10^{-4}	7.55×10^{-4}
GW151216	8.16	5.41	6.10	3.84	0.36	0.51	9.24×10^{-4}	1.81×10^{-4}
GW151216A	8.48	6.88	4.52	0.18	−0.08	−0.09	9.24×10^{-4}	1.71×10^{-4}
GW151217	7.99	5.59	3.55	0.15	−0.11	0.04	9.24×10^{-4}	1.71×10^{-4}
GW151222	10.22	10.50	3.50	9.34	10.86	0.13	9.24×10^{-4}	1.71×10^{-4}
GW170104A	7.36	5.76	5.13	1.37	0.05	0.34	6.16×10^{-4}	4.42×10^{-4}
GW170106	9.62	10.05	2.90	8.19	11.15	−0.09	1.58×10^{-3}	1.05×10^{-3}
GW170121	10.55	5.43	9.04	12.26	0.4	8.49	2.78×10^{-4}	3.42×10^{-3}
GW170123	6.20	4.29	5.61	2.34	0.05	0.58	1.85×10^{-4}	1.08×10^{-3}
GW170201	8.25	6.44	6.12	3.64	0.36	0.45	2.64×10^{-4}	3.19×10^{-4}
GW170202	8.50	5.26	7.29	5.33	0.31	1.48	3.82×10^{-4}	3.10×10^{-4}
GW170220	6.59	4.68	4.81	2.18	0.18	0.94	2.57×10^{-4}	2.71×10^{-4}
GW170304	8.50	5.75	7.50	7.53	0.46	5.83	2.20×10^{-4}	1.04×10^{-3}
GW170402	9.00	3.43	9.29	6.16	0.01	6.67	2.92×10^{-4}	7.04×10^{-4}
GW170403	8.20	6.27	6.47	5.8	1.56	1.79	4.00×10^{-4}	9.21×10^{-4}
GW170425	7.99	5.75	5.98	6.74	1.69	2.61	5.99×10^{-4}	4.56×10^{-4}
GW170620	8.36	6.05	6.11	3.38	0.11	0.33	8.77×10^{-4}	3.06×10^{-4}
GW170629	7.56	6.40	3.41	−0.06	−0.12	−0.3	7.62×10^{-4}	3.91×10^{-4}
GW170721	8.55	7.13	5.88	5.66	3.08	0.23	5.51×10^{-4}	2.87×10^{-4}
GW170724	10.76	11.84	6.27	13.46	17.14	−0.2	7.87×10^{-4}	2.69×10^{-4}
GW170727	9.90	4.97	7.66	9.15	0.16	6.27	4.54×10^{-4}	2.52×10^{-4}
GW170729A	11.55	11.17	4.35	10.5	12.73	0.7	3.31×10^{-4}	2.89×10^{-4}
GW170801	8.14	3.14	8.05	2.69	0.05	4.17	2.45×10^{-4}	2.22×10^{-4}
GW170817A	10.65	3.82	10.42	16.62	0.11	15.88	2.55×10^{-4}	1.62×10^{-4}
GW190412	18.47	9.77	15.80	59.27	10.59	42.54	1.00×10^{-4}	1.00×10^{-4}

marginalize over the uncertainty in the estimated PSD in order to obtain the correct likelihood for Gaussian, stationary noise [59]. As a result, when using a median off-source PSD, marginalization over the noise could have a more dramatic effect on the astrophysical signal odds if an incorrect likelihood is used. Based on these considerations, we opt to use on-source BAYESWAVE PSDs throughout this work.

To compute evidences, we perform a coherent Bayesian analysis using the nested sampling algorithm [36,66,67] implemented in LALINFERENCE [61]. We use the precessing waveform model IMRPhenomPv2 [68–71] for both the signal and incoherent glitch model.

In order to analyze each event on an equivalent footing, we adopt the same priors for the BBH parameters for all events. The choice of priors used in our analysis is

templated on the default settings used in GWTC-1 [3]. We use a uniform prior in the component masses $m_{1,2}$ in the range $[5.0, 160.0]\ M_{\odot}$ and an isotropic spin prior with dimensionless spin magnitudes $\chi_{1,2}$ taken to be within $[0, 0.99]$. We further restrict the redshifted chirp mass $(1 + z)\mathcal{M} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$ to lie within $[5.0, 100.0]\ M_{\odot}$ and the mass ratio $q = m_2/m_1$ to lie within $0.05 \leq q \leq 1$. The distance prior is taken to be proportional to the luminosity distance squared, with an upper limit of 5 Gpc.

III. RESULTS

We analyze the event candidates reported in GWTC-1 [3], 2-OGC [4] the PyCBC single detector search [5],[2] and the independent IAS search [6–10]. In addition, we also analyze GW190412 [11] as the first significant BBH reported from the third observing run. We focus only on BBH events, restricting our analysis to event candidates with a redshifted chirp mass $(1 + z)\mathcal{M} > 5\ M_{\odot}$. In order to gauge the robustness of our Bayesian framework, we also analyze a set of known background glitches identified by the PyCBC [25,26] and GstLAL [29,30] search pipelines.

In Table I, we summarize the various Bayesian measures of significance for all event candidates considered in our analysis. We report the astrophysical signal odds $\mathcal{O}_{S/N}$, defined in Eq. (3), and the astrophysical signal probability $P(\mathcal{S}|\mathbf{d})$, Eq. (2). The prior for the signal hypothesis is given by Eq. (13) and the glitch prior per event for each detector is given in Table II of Appendix B. For ease of comparison with the literature, we also show the value of p_{astro} reported by the various search pipelines. In addition we report the Bayes factors under various limiting cases, defined in Eqs. (3), (6), (8) and (11). As a further (obvious) confirmation that p_{astro} is pipeline specific, it is also useful to stress that ranking candidates based on $P(\mathcal{S}|\mathbf{d})$ produces a list which differs from the equivalent lists based on p_{astro}.

A. GWTC-1

For all binaries reported in GWTC-1, bar GW170818 which we discuss below, and GW190412, we find overwhelming evidence in favor of the astrophysical signal hypothesis, $P(\mathcal{S}|\mathbf{d}) \sim 1$. For GW170729, which was assigned a p_{astro} of 0.52 and 0.98 by the two main search pipelines in [3], we find an unambiguous astrophysical signal probability of $P(\mathcal{S}|\mathbf{d}) = 1$.

For GW170818 we find $P(\mathcal{S}|\mathbf{d}) = 0.1$. This signal was initially detected as a triple-coincidence event by GstLAL with a signal-to-noise ratio of ≈ 10 in Livingston, and ≈ 4 in both Hanford and Virgo [3]. We find consistent values of signal-to-noise ratio in our analysis, see Table II, though we

---

[2]Note that we only consider triggers that are found in more than one detector.

do not analyze data from Virgo. The event was initially only detected in the Livingston data by the PyCBC search, though a reanalysis using modified settings around the time of the event did find triggers with a similar signal-to-noise ratio to those reported by GstLAL [3]. This event has also been reported in the dedicated BBH search in [4] and as a significant single-detector trigger in [6,10].

It is clear that GW170818 is a relatively weak signal and, due to the markedly different sensitivity of the instruments, is a *de facto* single-detector event. It is therefore unsurprising that the astrophysical signal probability returned by our approach, which critically relies on the notion of coherent signals across a network of detectors, is small and in tension with results reported by the search pipelines. This is expected to be the case whenever the signal in one of the two detectors is too weak to be distinguishable from Gaussian noise ($\rho \sim 5.5$). As further insight into the nature of this candidate, adopting the Bayesian coherence ratio [37] as a measure of significance, we find $\log_{10} \mathcal{B}_{cr} = -0.39$, the only circumstance in which one obtains a negative value for the significant BBH events in GWTC-1.

Finally, we note that for single detector triggers, alternative strategies for determining the significance and astrophysical signal odds should be pursued [3,5,6,72]. In particular see [5], which reports a large positive Bayes factor for a coherent signal versus a null hypothesis consisting of a signal in one detector and noise in all others.

B. Binaries reported by the IAS search

We now focus on the nine event candidates reported by the IAS search pipeline [6,8–10]. GW151216 was first reported as a new, significant trigger in [6] with $p_{\text{astro}} \sim 0.71$ and a high, positive effective spin $\chi_{\text{eff}} \sim 0.8$. This event was subsequently reported in [4], though with a markedly lower $p_{\text{astro}} \sim 0.18$. In our Bayesian framework, we find that the probability that the signal is of astrophysical origin is overwhelmingly disfavored, $P(\mathcal{S}|\mathbf{d}) \sim 4 \times 10^{-4}$, in broad agreement with the results presented in [46].

Of the six binaries reported in [10], we find that *three* of the event candidates, GW170121, GW170425 and GW170727, are of particular interest, with astrophysical probabilities $P(\mathcal{S}|\mathbf{d})$ of 0.53, 0.74 and 0.66, respectively. These three events are also the only candidates with $\log_{10} \mathcal{B}_{cr} > 0$.

GW170121 is the most significant candidate reported in [10], with both the highest SNR and a $p_{\text{astro}} > 0.99$. This trigger was also subsequently found as a significant event candidate in [4], with $p_{\text{astro}} \sim 1$. This event is notable due to its negative effective spin $\chi_{\text{eff}} \approx -0.2$ [10], see also Fig. 8 in Appendix C.

GW170425 and GW170727 are consistent with being generated by heavy BBHs with source frame chirp masses $\mathcal{M} \approx 30\ M_{\odot}$ and effective aligned spins $\chi_{\text{eff}} \sim 0$, in broad agreement with the population of binaries observed in
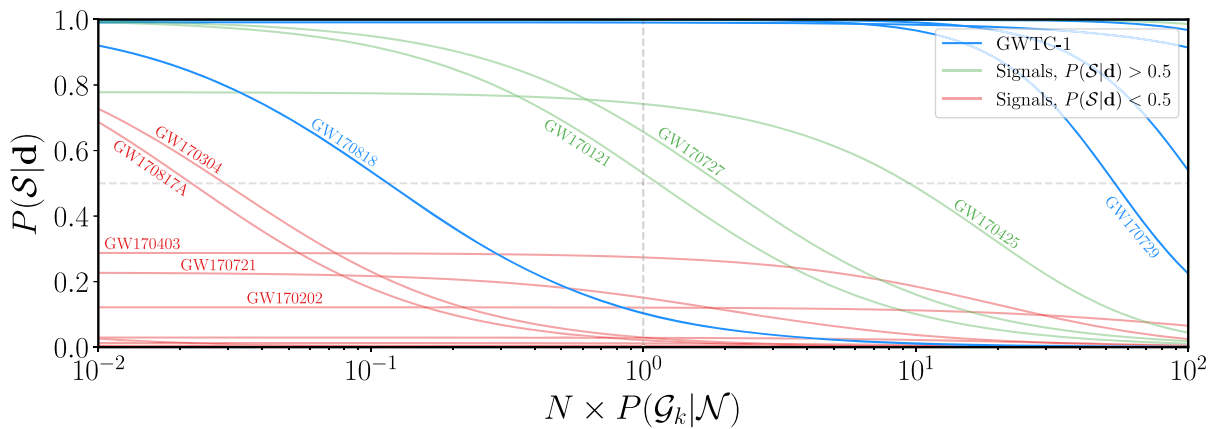
FIG. 2. $P(\mathcal{S}|\mathbf{d})$ as a function of the glitch probability $P(\mathcal{G}_k|\mathcal{N})$. Here we scale the fiducial values $P(\mathcal{G}_k|\mathcal{N})$ by a factor of $N \in [10^{-2}, 10^2]$, keeping the ratio of the glitch probability in each detector fixed, i.e., $P(\mathcal{G}_H|\mathcal{N})/P(\mathcal{G}_L|\mathcal{N}) = \text{const}$. For all GWTC-1 events (blue), bar the single detector event GW170818, we would need to increase the glitch probability by a factor of $\sim \mathcal{O}(10^2)$ in order for the astrophysical probability $P(\mathcal{S}|\mathbf{d})$ to drop below 0.5. Of the interesting event candidates (green), GW170121 is particularly sensitive to the glitch probability, with even a small increase in $P(\mathcal{G}_k|\mathcal{N})$ pushing $P(\mathcal{S}|\mathbf{d})$ below 0.5.

GWTC-1 [35]. We report the measured parameters of these three events in Appendix C and Figs. 8–11.

Of the remaining candidates reported in [10], we find that only GW170403 is of marginal interest, with an astrophysical probability $P(\mathcal{S}|\mathbf{d}) = 0.27$. The other two candidates, GW170202 and GW170304, have a $P(\mathcal{S}|\mathbf{d})$ comparable to GW170818.

The final two event candidates found in the IAS search, GWC170402 and GW170817A, were first reported in [6], a search targeting compact binary mergers that produce a clear signal in one of the detectors and a marginal signal in the other detectors. As per the discussion for GW170818, we do not expect the astrophysical signal probability returned by our framework to be large, finding $P(\mathcal{S}|\mathbf{d}) \sim 0$ and 0.02, respectively. Unlike GW170818, the Bayesian coherence ratio also significantly disfavors these events, with $\log_{10} \mathcal{B}_{\text{cr}}$ being more than an order of magnitude smaller.

C. Binaries reported by the PyCBC search

As per the event candidates reported by the IAS search, we find that only the three events reported in the PyCBC catalog [4], GW170121, GW170425 and GW170727 are of interest. All other event candidates have negligible astrophysical probability. The PyCBC reported p_{astro} is ~ 1 for GW170121, ~ 0.21 for GW170425, and ~ 0.99 for GW170727.

The other significant event reported in [4] is GW170304, which was assigned $p_{\text{astro}} = 0.7$ but for which we find that the astrophysical signal hypothesis is disfavored with $P(\mathcal{S}|\mathbf{d}) = 0.03$.

D. Sensitivity of results to priors and glitches

In Fig. 1 we show how sensitive $P(\mathcal{S}|\mathbf{d})$ is to the choice of the prior probability on the signal, $P(\mathcal{S})$. Even if we

reduce the value of $P(\mathcal{S})$ by two orders of magnitude, we find that $P(\mathcal{S}|\mathbf{d}) \approx 1$ for all GWTC-1 BBHs, with the exception of GW151012, which would be recorded with $P(\mathcal{S}|\mathbf{d}) \approx 0.6$, and GW170818, whose value would be ≈ 0 as per the discussion above.

The astrophysical signal probability of GW170121 and GW170425 is always above 0.5, even when taking the lowest value of $P(\mathcal{S})$ compatible with the 90% probability interval of the local BBH merger rate $\mathcal{R}_0$. For GW170727, the result drops just below our arbitrary threshold of 0.5 to $P(\mathcal{S}|\mathbf{d}) \approx 0.4$. A signal prior $P(\mathcal{S})$ that were at least an order of magnitude higher would yield $P(\mathcal{S}|\mathbf{d}) > 0.5$ for GW170403, GW170721 and GW170202.

In Fig. 2 we explore how sensitive the results are to the prior probability of the glitch hypothesis at a fixed signal probability corresponding to the value in Eq. (2). For all GWTC-1 BBHs, with the exception of GW170818, the results are fairly insensitive to variations in $P(\mathcal{G}_k|\mathcal{N})$ by at least a factor of 100.

For GW170425, one would need to increase the glitch probability by a factor ≈ 10 to reduce $P(\mathcal{S}|\mathbf{d})$ below 0.5. In contrast, both GW170121 and GW170727 are much more sensitive to variations in $P(\mathcal{G}_k|\mathcal{N})$. Increasing the glitch prior $P(\mathcal{G}_k|\mathcal{N})$ by a factor of ≈ 2–3 would be sufficient to reduce $P(\mathcal{S}|\mathbf{d})$ below 0.5, emphasizing the importance of developing a robust model for the glitch hypothesis.

Interestingly, $P(\mathcal{S}|\mathbf{d})$ for GW170403, GW170721 and GW170202 is insensitive to the glitch probability in the region considered here: the signals are weak, and as discussed above $P(\mathcal{S}|\mathbf{d})$ is driven by the prior signal probability.

In Appendix B and Fig. 7, we provide additional diagnostics for a subset of events, including GW170425, GW170121 and GW170227, to show how $P(\mathcal{S}|\mathbf{d})$ varies in the $P(\mathcal{G}_L|\mathcal{N}) - P(\mathcal{G}_H|\mathcal{N})$ plane.
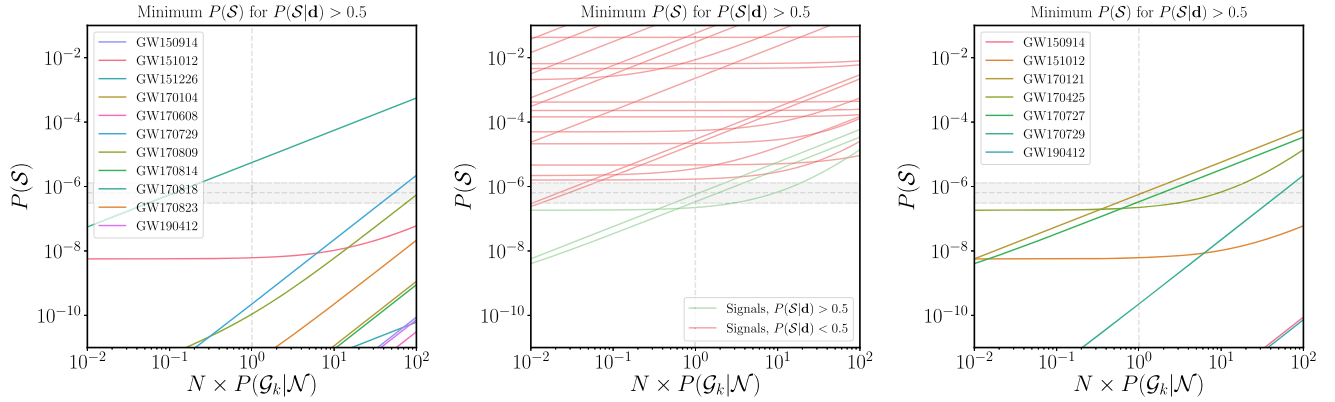
FIG. 3. The minimum $P(\mathcal{S})$, which scales linearly with the total astrophysical merger rate integrated over the entire population, required to obtain a probability that the signal is of astrophysical origin $P(\mathcal{S}|\mathbf{d}) > 0.5$ as a function of the glitch probability $P(\mathcal{G}_k|\mathcal{N})$. Here we scale the fiducial values $P(\mathcal{G}_k|\mathcal{N})$ by a factor $N \in [10^{-2}, 10^2]$, keeping the ratio of the glitch probabilities between the detectors fixed, where the ratio is set by the values reported in Table II. The left panel shows all confident events identified in GWTC-1 and GW190412. The middle panel shows all other triggers listed in Table. I. The right panel shows a subset of interesting events, including the three most interesting triggers identified: GW170121, GW170425 and GW170727. The gray shaded regions denote the range of values $P(\mathcal{S})$ can take by varying the inferred local merger rate $\mathcal{R}_0$ over its 90% credible interval [35]. The vertical dashed lines denote the fiducial values of $P(\mathcal{G}_k|\mathcal{N})$.

A complementary way of exploring how the astrophysical signal odds and associated signal posterior probability depend on the signal and glitch prior probabilities is displayed in Fig. 3. We show the minimum $P(\mathcal{S})$ required to obtain an astrophysical probability of $P(\mathcal{S}|\mathbf{d}) > 0.5$ as a function of the glitch probability $P(\mathcal{G}_k|\mathcal{N})$, where the ratio of $P(\mathcal{G}_k|\mathcal{N})$ between the detectors is kept constant, and set by the values reported in Table II. The left panel of Fig. 3 shows all confident GWTC-1 events plus GW190412. The middle panel shows all remaining triggers, with the three most compelling events GW170121, GW170425 and GW170727 highlighted in green. The right panel of Fig. 3 focuses on a subset of interesting events. We show GW150914 and

GW190412 as clear, unambiguous detections such that increasing the glitch rate by a factor $\sim \mathcal{O}(10^2)$ would still be insufficient to disfavor the astrophysical hypothesis. GW151012 was initially identified with a high FAR [32] but improvements to the analysis pipelines used for the GWTC-1 reanalysis of this event substantially reduced the FAR, leading to a $p_{astro} = 0.96$ [3]. In our Bayesian framework, GW151012 is unambiguously identified as an astrophysical signal with $P(\mathcal{S}|\mathbf{d}) \sim 0.99$, in agreement with [37,46] and showcasing the utility of our unified framework for calculating the astrophysical signal probability. Finally, we highlight the three most compelling events from our analysis presented: GW170121, GW170425 and GW170727.
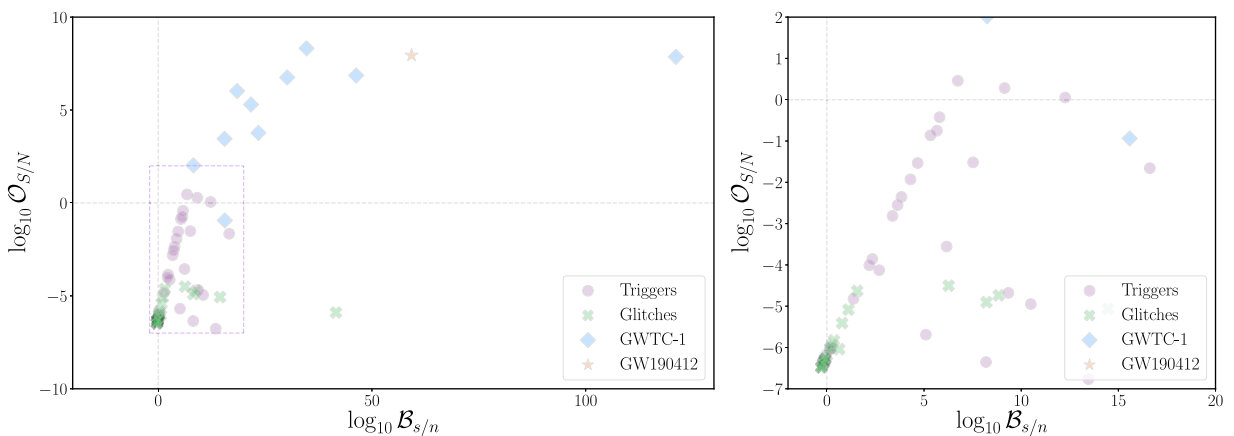
FIG. 4. Bayes factor for the signal to noise hypothesis $\mathcal{B}_{s/n}$ against the astrophysical signal odds $\mathcal{O}_{S/N}$ for all events in GWTC-1 (blue), GW190412 (orange), the candidate events listed in Table I (red) and a set of known glitches (green). Despite relatively high signal-to-noise Bayes factors, the astrophysical signal odds for the population of known glitches is efficiently suppressed within this framework. In the left panel we show all triggers and in the right panel we focus on the boxed region demarcated by the purple dashed lines.

The horizontal shaded area in Fig. 3 denotes the allowed range of the prior probability on the signal as inferred from Eq. (13) obtained by varying the local astrophysical merger rate $\mathcal{R}_0$ over its 90% confidence interval. For all events with significant support for the astrophysical signal hypothesis, the minimum $P(\mathcal{S})$ required to meet the threshold of 0.5 is far below the astrophysical rate estimated in Appendix A for typical glitch rates estimated from the distribution of OMICRON triggers, with an upper limit on the glitch prior of $P(\mathcal{G}_k|\mathcal{N}) \sim$ few $\times 10^{-3}$. As an interesting case study, for GW170729 we would need to increase the glitch rate by a factor of $\sim 10^2$ before $P(\mathcal{S}|\mathbf{d})$ starts to drop below 0.5, emphasizing the robustness of this result.

It would be reasonable to wonder if our approach is robust against noise transients of instrumental origin that pollute the data. In Fig. 4, we show the Bayes factor for the signal to noise hypothesis $\mathcal{B}_{s/n}$ against the astrophysical signal odds $\mathcal{O}_{S/N}$ for all triggers listed in Table I. In addition, we plot the same quantities evaluated for a population of known glitches identified by the PyCBC search. Figure 4 serves to demonstrate the efficacy with which the inferred $P(\mathcal{S}|\mathbf{d})$ for glitches is suppressed and as a caution that the standard Bayes factor for signal vs noise $\mathcal{B}_{s/n}$ is a poor discriminator between astrophysical signals and instrumental noise transients.

IV. IMPLICATIONS FOR POPULATION INFERENCE

In our Bayesian analysis, we have identified three additional event candidates with a $P(\mathcal{S}|\mathbf{d}) > 0.5$ (see Table I) which would be sufficient for inclusion in typical population studies [3,35]. In this section, we assess the impact of including GW170121, GW170425 and GW170727 on the inferred BBH population properties [35].

Following the methodology described in [35], we perform hierarchical Bayesian inference, incorporating measurement uncertainty and selection effects into the analysis [34,73–75]. The rate of events is modeled as a Poisson process whose mean is dependent on the parameter distribution of the population of binary black holes

$$\mathcal{L}(\{d_n\}|\Lambda) \propto e^{-\mu(\Lambda)} \prod_{n=1}^{N_{\mathrm{obs}}} \mathcal{L}(d_n|\theta,z) \frac{dN}{d\theta dz}(\Lambda) d\theta dz, \quad (16)$$

where θ are the intrinsic parameters of the binary, z the redshift, N the total number of mergers that occur within the detection horizon and $\mu(\Lambda)$ the rate constant describing the mean number of events as a function of the population hyperparameters Λ. Here N_{obs} is the number of detections and the individual-event likelihood for the nth detection is $\mathcal{L}(d_n|\theta,z)$. The expected number of events is dependent on the selection effects for the detectors. This can be characterized by the sensitive spacetime volume $VT(\theta)$ for a detector network in terms of the detection probability for a

binary with parameters θ and the total observation time, as discussed in Appendix A.

For the population inference, we use Model C of [35] as a representative model. This model is based on a power law mass distribution with an additional Gaussian component at high masses [76]. The low-mass cutoff m_{min} is tapered by a smoothing scale δm to account for environmental effects, such as metallicity, that can blur the edge of the lower mass gap [35,76] and the maximum allowed BH mass is given by m_{max}. At low masses, we recover a standard mass power law governed by a power-law index α on the primary BH and a mass ratio power-law index β_q. The Gaussian component models the possible build up of high-mass BBHs from pulsational pair-instability supernovae [77–81] and is parametrized by the mean μ_m and standard deviation σ_m and well as the fraction of primary BHs in the Gaussian component λ_m. The distribution of spin magnitudes χ_i is taken to be drawn from a beta distribution [73] and can be parametrized by either the moments of the distribution α_χ and β_χ or, equivalently, by the mean μ_χ and variance σ_χ^2. The distribution of spin orientations is modeled as a mixture of two distributions corresponding to an isotropic and a preferentially aligned component [82]. The fraction of binaries preferentially aligned with the orbital angular momentum is denoted ξ and the degree of spin misalignment is denoted by σ_{χ_i}. The redshift evolution of the model follows the prescription in [75], where, for simplicity, we assume a merger rate density that is uniform in comoving volume and source-frame time.

The likelihood $\mathcal{L}(d_n|\theta,z)$ is recycled from the posterior samples calculated using LALINFERENCE [61]. We assume uniform priors on the population parameters, as detailed in Table 2 of [35]. For the event rate, we take a log-uniform distribution bounded between $[10^{-1}, 10^3]$. The hierarchical population inference is carried out using the GWPOPULATION package [83] and BILBY [38].

As a proxy for incorporating full information regarding the astrophysical signal probability, we use the posterior probability of the signal hypothesis $P(\mathcal{S}|\mathbf{d})$ as a criterion for including events in our population inference. Following [3,35], the threshold is taken to be $P(\mathcal{S}|\mathbf{d}) > 0.5$. Based on the analysis discussed in Sec. III, the events GW170121, GW170425 and GW170727 satisfy this selection criterion for our default priors and we can incorporate them into the hierarchical inference.

Various approaches for incorporating p_{astro} in hierarchical population inference have recently been presented [15,84,85]. In particular, these methods allow one to simultaneously infer both the foreground and background rates along with concomitant astrophysical distributions. In [84], the authors applying such a method to all GWTC-1 candidates, as well as the eight binaries reported in [8–10], finding that all binaries are up weighted to a $p_{\mathrm{astro}} \sim 1$ once including for the population model. This places the results of [84] in some tension with the results presented both here

and in [46]. Whilst a detailed reanalysis of the impact of p_{astro} on population inference is beyond the scope of this paper, discarding information regarding the coherence of the signal could be partially culpable for the discrepancies

observed and could potentially lead to nontrivial biases in the inferred astrophysical merger rate.

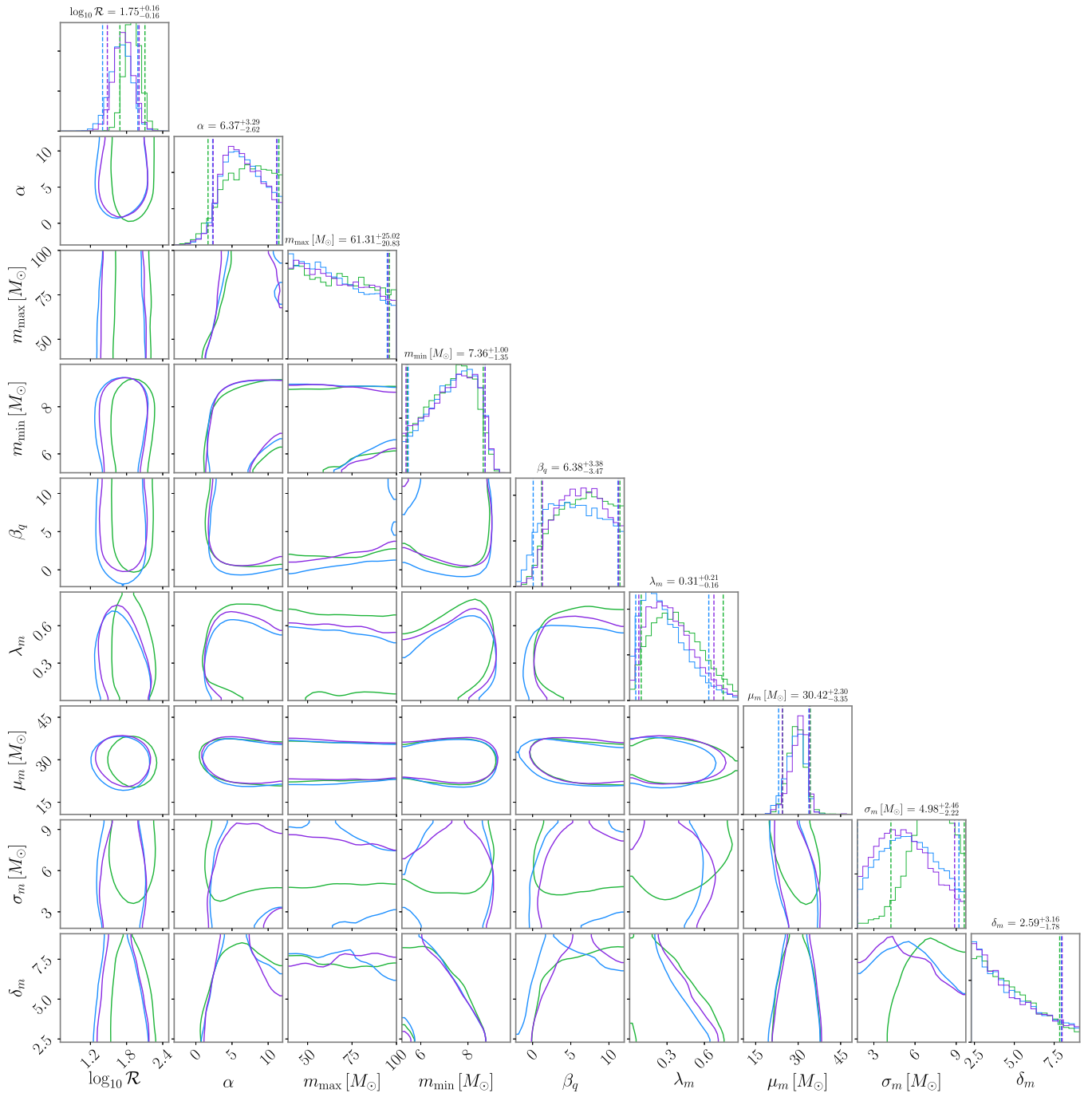In Figs. 5 and 6 we compare the posterior distributions for the mass and spin population hyperparameters,

FIG. 5. The 1D and 2D posterior distributions for the hyperparameters describing the BBH population model for the masses and rates. Here the blue posteriors denote the results of the analysis considering all GWTC-1 events in the population analysis. The purple posteriors show the results when including the three events with $P(\mathcal{S}|\mathbf{d}) > 0.5$, GW170121, GW170425 and GW170727 in addition to all GWTC-1 events. The green posteriors refer to the inferred population based on all GWTC-1 events plus all triggers reported in [6–9] irrespective of $P(\mathcal{S}|\mathbf{d})$. We also report the 1D median values and 90% probability interval inferred using all GWTC-1 binaries plus the three events with $P(\mathcal{S}|\mathbf{d}) > 0.5$.
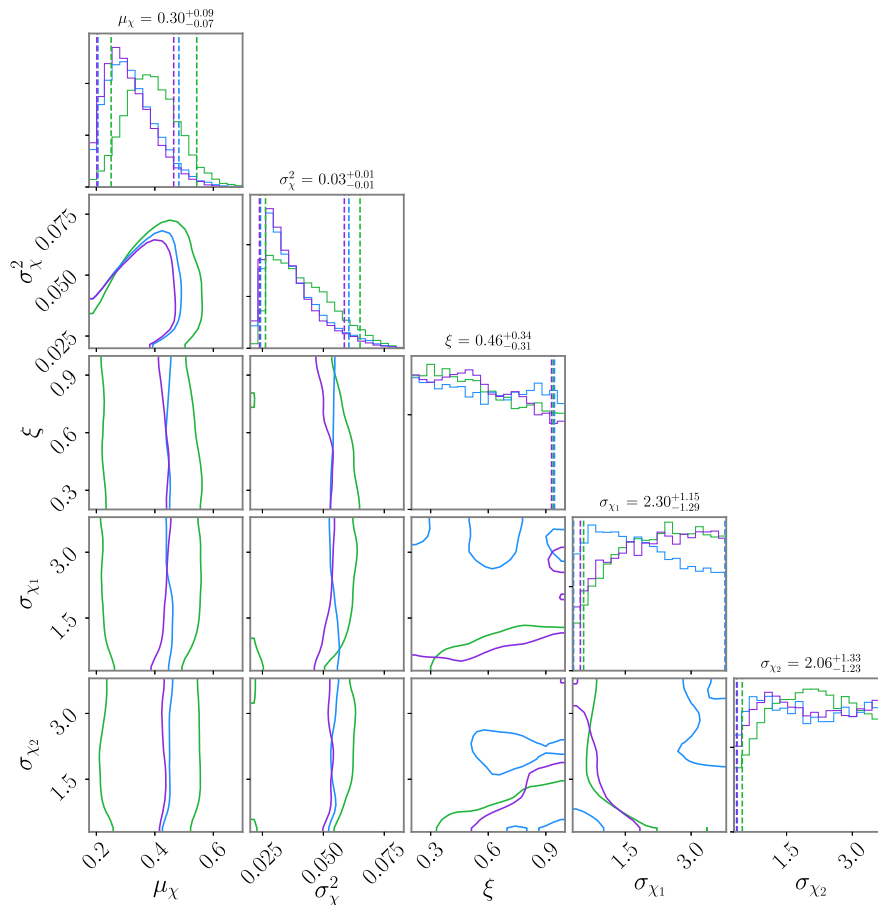
FIG. 6. The 1D and 2D posterior distributions for the hyperparameters describing the spin population model. As per Fig. 5, the blue posteriors denotes all GWTC-1 events. The purple posteriors show all GWTC-1 events plus the three events with $P(\mathcal{S}|\mathbf{d}) > 0.5$, GW170121, GW170425 and GW170727. The green posteriors show all GWTC-1 events plus all triggers reported in [6–9]. We report the 1D median values and 90% probability interval inferred using all GWTC-1 binaries plus the three events with $P(\mathcal{S}|\mathbf{d}) > 0.5$.

respectively, by including in the population study: (i) GWTC-1 binaries alone (blue curves), (ii) GWTC-1 binaries and the three new event candidates with significant astrophysical probability as determined by our analysis (purple curves) and (iii) GWTC-1 binaries and *all* binaries reported in the IAS search, excluding GW170402, irrespective of $P(\mathcal{S}|\mathbf{d})$ as in [84]. The population hyperparameters inferred from GWTC-1 plus the three new events are in broad agreement with the results inferred using the GWTC-1 binaries alone. We see a slight shift to higher values of the astrophysical merger rate $\mathcal{R}_0$ and some of the other hyperparameters, notably the degree of spin misalignment for the primary BH σ_{χ_1}.

Incorporating the eight IAS event candidates, we find that the population hyperparameters are increasingly discrepant with respect the results from using GWTC-1 alone. In particular, we observe that the astrophysical merger rate $\mathcal{R}_0$, the width of the Gaussian mass peak σ_m and the mean of the spin magnitude distribution μ_χ shift quite significantly. We observe the same shift in σ_{χ_1} as when including

the three events with significant astrophysical probability returned by our analysis.

Incorporating information regarding the coherence of the signal will likely be an important consideration when performing population inference on an ever growing catalog of observed compact binaries. This highlights the necessity of a robust, unified framework for calculating the astrophysical signal odds agnostic to any specific search pipeline of strategy, such as the framework presented here.

V. DISCUSSION

We have revisited the problem of determining whether an observed GW transient is due to an astrophysical source. We employ a Bayesian framework to derive the posterior odds ratio, $\mathcal{O}_{S/N}$, allowing us to determine the probability that the signal is of astrophysical origin, $P(\mathcal{S}|\mathbf{d})$, independent of any search pipeline. The astrophysical posterior odds is determined through a coherent analysis of the data, depending only on the underlying model hypotheses and the choice of prior probabilities. In order to set the

astrophysical signal prior $P(\mathcal{S})$, we leveraged prior knowledge on the population of astrophysical BBHs. The glitch probability for a given detector $P(\mathcal{G}_k|\mathbf{d})$ was determined using the density of OMICRON triggers in a 24h stretch of data around each event, allowing us to account for asymmetric noise profiles between the detectors and for changes in the behavior of the detectors over time. This work builds on recent studies in the literature [28,36–38].

We analyze all confident BBH events reported in GWTC-1 [3] as well as event candidates recently reported by independent search pipelines [4–6,8,10]. Using the astrophysical posterior odds $\mathcal{O}_{S/N}$, we find that all GWTC-1 binaries, except for the single detector event GW170818, have overwhelming odds in favor of the astrophysical signal hypothesis. Of all other event candidates analyzed, only GW170121, GW170425 and GW170727 have odds in favor of the astrophysical signal hypothesis, $P(\mathcal{S}|\mathbf{d}) > 0.5$, meeting the threshold for inclusion in typical population studies [35]. We characterized the robustness of these results to changes in the signal $P(\mathcal{S})$ and glitch $P(\mathcal{G}_k|\mathcal{N})$ priors. Due to the significant computational burden of our analysis, in this paper we have considered only BBH candidates; our method however is fully general and can be applied to all classes of binary systems.

The methodology presented in this paper can be further improved in a number of ways. Due to the limited coincident data available, and the nontrivial computational processing load we restricted our analysis to a 2-detector network. As the duty cycle of the instruments improves, the analysis should incorporate data from all available detectors in a given network at any given time. The fourth observing run (O4) is scheduled to last for one year, with the LIGO detectors nearing design sensitivity and Phase 1 of the Advanced VIRGO + upgrade nearing completion [86]. In addition, KAGRA [87,88] will be operational with a nominal BNS range of 25–130 Mpc [86], providing the potential opportunity for analyzing coherent data from four detectors. The discriminating power of our approach between astrophysical signals and instrumental transient will correspondingly increase as data from more instruments are included.

With detector sensitivities ever increasing, the approximation that signals and glitches do not overlap can break down, as was the case for GW170817 [89]. Whilst glitch subtraction can be an effective tool to mitigate such scenarios [52], the signal model used here can be generalized to include a more robust glitch model which may improve the discriminating power of this framework. Similarly, we can extend the framework to account for the possibility of overlapping astrophysical signals [90–93].

For computational efficiency, we used the precessing waveform model IMRPhenomPv2 [68–71]. In future analyses, we plan to utilize improved and more accurate waveform models [94] that incorporate both higher

modes [95–97] and precession [98,99]. As a first step, we have reanalyzed GW170121, GW170425 and GW170727 with the precessing higher-mode waveform model PhenomXPHM [98], while keeping all the other parameters of our analysis unchanged. We find that $P(\mathcal{S}|\mathbf{d})$ increases from 0.53 to 0.56 for GW170121, from 0.74 to 0.84 for GW170425, and from 0.66 to 0.67 for GW170727.

To process large catalogs of candidate events or inferring $P(\mathcal{S}|\mathbf{d})$ in low latency—the latter can be done with rather minor changes to the present implemented software infrastructure and relatively small additional computational cost—it will be important to adopt methods that further reduce waveform generation costs [100,101] and mitigate the computational cost of the Bayesian inference [102,103].

Due to the lack of a robust glitch model, we made the simplifying approximation that $\mathcal{G}_k \approx \mathcal{S}_k$. Whilst this is the most conservative choice that could be made [36], our understanding of the instruments' behavior is continuously improving, as is our ability to classify and describe glitches. In the future, a reliable glitch model could be available and allow us to more adequately distinguish between astrophysical signals and glitches, thereby increasing the significance of the inferred $P(\mathcal{S}|\mathbf{d})$.

The astrophysical signal prior $P(\mathcal{S})$ used in this analysis was determined by calculating the expected number of observed BBHs within the relevant time span from a population of binaries taking into account selection effects, see Eq. (A4). A more detailed characterization of the detector sensitivity $\langle VT \rangle_\Lambda$, e.g., through Monte Carlo integration [35,104] or novel applications of machine learning [105], could allow for a more accurate determination of $P(\mathcal{S})$ as a function of the population hyperparameters Λ. In addition our analysis can be improved by dividing the whole observing period into shorter segments, and updating $P(\mathcal{S})$ as the analysis progresses. Finally, the default priors used in calculating the Bayes factors are uniform in the component masses, uniform in spin magnitudes and isotropic in spin orientations. The use of such agnostic priors is of particular importance in the absence of any *a priori* knowledge about the expected shape of the distributions. However, as the number of observations increases, it will be important to consider population informed priors, e.g., by constructing the posterior predictive distributions [84,106–108].

ACKNOWLEDGMENTS

APPENDIX A: CALCULATION OF THE SIGNAL PRIOR PROBABILITY

Here we provide details of the calculation of the prior signal probability, $P(\mathcal{S})$, given by Eq. (13).

In order to calculate the expected number of detections $\mu(\Lambda)$, we need to incorporate selection effects from a network of detectors. Here, we calculate the sensitive spacetime volume $\langle VT \rangle$ following a semianalytic prescription [32,35,42]. For a network of gravitational wave detectors, the sensitive spacetime volume is

$$VT(\theta) = T_{\text{obs}} \int_0^{z_{\text{max}}} f(z|\theta) \frac{dV_c}{dz} \frac{1}{1+z} dz, \quad (A1)$$

where we assume a constant sensitivity over an observing run T_{obs} and $f(z|\theta)$ is the detection probability for a binary with parameters θ at a redshift z [110] averaged over extrinsic parameters [42]. For each binary, we calculate the optimal SNR ρ_{opt} using the IMRPhenomXAS phenomenological waveform model [94], corresponding to the SNR that would be observed for a face-on source located directly overhead a detector. For an isotropic distribution of sources, with arbitrary sky locations and inclinations, the distribution of SNRs can be captured by introducing a coefficient Θ that parametrizes the angular dependence, $0 \leq \Theta = \rho/\rho_{\text{opt}}$, and where Θ has a known distribution [42]. Here $\Theta = 1$ corresponds to an optimally oriented source and $\Theta = 0$ corresponds to a binary that is in a blind spot of the detector. The probability of detecting a source at a redshift z can therefore be written as

$$f(z|\theta) \equiv p_{\text{det}}(z|\theta) = P(\rho \geq \rho_{\text{th}}), \quad (A2)$$

where ρ_{th} is the threshold for observing a binary. Here we assume that sources will only be detected if they have a single detector SNR above a threshold $\rho_{\text{ifo}} > 7$.

We incorporate both mass and spin dependence in the calculation of $VT(\theta)$, though spins have a subdominant effect on the ρ calculation for the population of binaries considered here [35]. For a given population with hyperparameters Λ, the total sensitive spacetime volume is [35]

$$\langle VT \rangle_\Lambda = \int_\theta p(\theta|\Lambda) VT(\theta) d\theta. \quad (A3)$$

If the merger rate evolves with redshift, then the expected number of detections will be given by

$$\mu(\Lambda) = T_{\text{obs}} \int_\theta \int_0^\infty p(\theta|\Lambda) f(z|\theta) \mathcal{R}(z) \frac{dV_c}{dz} \frac{1}{1+z} dz d\theta. \quad (A4)$$

Following [75], we can parametrize the evolving merger rate density $\mathcal{R}(z)$ in the comoving frame by

$$\mathcal{R}(z|\lambda) = \mathcal{R}_0 (1+z)^\lambda, \quad (A5)$$

where $\mathcal{R}_0$ is the local merger rate density at $z = 0$. Here, $\lambda = 0$ corresponds to a uniform in comoving volume merger rate and $\lambda \sim 3$ is a merger rate that approximately follows the star formation rate [111], at redshifts below $z \approx 2$, which corresponds to the peak of the star formation rate. For the population of binaries here, we take $\lambda = 2$, consistent with [11,35]. In the simpler scenario in which the merger rate is constant with redshift, the expected number of observations reduces to

$$\mu(\Lambda) = \mathcal{R}_0 \langle VT \rangle_\Lambda. \quad (A6)$$

The expectation value of $VT(\theta)$ can be calculated using standard Monte Carlo methods

$$\langle VT \rangle_\Lambda = \int p(\theta|\Lambda) VT(\theta) d\theta \approx \frac{1}{N} \sum_k^N VT(\theta_k). \quad (A7)$$

We use PSDs that are representative of the detector performance in O2 [3,50] and assume a total coincident observing time of $T_{\text{obs}} = 0.46$ y [3]. For the population of binaries, we use a parametrized mass model

$$p(m_1, m_2 | m_{\text{min}}, m_{\text{max}}, \alpha, \beta_q) \propto C(m_1) m_1^{-\alpha} q^{\beta_q}, \quad (A8)$$

where the minimum and maximum black hole masses are taken to be $m_{\text{min}} = 5\ M_\odot$ and $m_{\text{max}} = 50\ M_\odot$, as motivated by the apparent lower limit for x-ray binary observations [112] and the maximum mass above which a pair-instability supernova is thought to completely disrupt the star [77]. The slope of the mass power law is taken to be $\alpha = 1.8$ and the exponent for the mass ratio to be $\beta_q = 2.0$, which is in broad agreement with [11,35]. The distribution of spin magnitudes χ_i are taken to follow a beta distribution [73]

$$p(\chi_i | \alpha_\chi, \beta_\chi) = \frac{\chi_i^{\alpha_\chi - 1} (1 - \chi_i)^{\beta_\chi - 1}}{B(\alpha_\chi, \beta_\chi)}, \quad (A9)$$

with $\alpha_\chi = 1.5$ and $\beta_\chi = 3.5$. The spin orientations are taken to follow a mixture model consisting of an isotropic component and a preferentially aligned component [82]

$$p(\cos t_1, \cos t_2 | \sigma_1, \sigma_2, \zeta)$$
$$= \frac{(1-\zeta)}{4} + \frac{2\zeta}{\pi} \prod_{i \in \{1,2\}} \frac{\exp(-(1-\cos t_i)^2/(2\sigma_i^2))}{\sigma_i \mathrm{erf}(\sqrt{2}/\sigma_i)}, \quad \text{(A10)}$$

where we take $\zeta = 0.5$ and $\sigma_i = 1$. In practice, we only use the spins projected along the orbital angular momentum $\vec{L}$ when evaluating the aligned-spin waveform model to calculate the SNR.

Using the population of binaries detailed above, the sensitive spacetime volume is estimated to be

$$\langle VT \rangle_\Lambda \approx 0.877 \text{ Gpc}^3 \text{ y}. \quad \text{(A11)}$$

Using a fiducial local astrophysical merger rate of $\mathcal{R}_0 = 53.2 \text{ Gpc}^{-3} \text{ y}^{-1}$ [35], the probability of a signal within a coalescence interval of $0.2s$ yields Eq. (13).

APPENDIX B: IMPACT OF GLITCH PRIOR PROBABILITY ON THE POSTERIOR SIGNAL PROBABILITY OF A SIGNAL

In Table II we report the glitch probability for each of the two instruments that we have used in the analysis, see Table I, and we also report the SNRs (signal detector and across the network) for each of the candidate events.

In Fig. 7 we explore the dependence of $P(\mathcal{S}|\mathbf{d})$ on the glitch prior for each of the detectors for selected events. The fiducial values of $P(\mathcal{G}_k|\mathbf{d})$ estimated from the OMICRON triggers are denoted with a gray star.

APPENDIX C: POSTERIORS FOR GW170121, GW170425 AND GW170727

In Figs. 8–10 we show the posterior distributions for a subset of key intrinsic parameters for the three candidate events with $P(\mathcal{S}|\mathbf{d}) > 0.5$: GW170121, GW170425 and GW170727. The posterior distributions show that the instrinsic parameters of the BBHs that generated these systems are broadly compatible with those that characterize the population of BBHs observed in O1 and O2 as reported in GWTC-1 [3,35], though we note that GW170121 has a negative effective aligned spin parameter $\chi_{\mathrm{eff}} \simeq -0.19^{+0.16}_{-0.17}$ [10,113].

In Fig. 11, we perform a sanity check on the posterior distributions for the three events with $P(\mathcal{S}|\mathbf{d}) > 0.5$, GW170729—which was reported in GWTC-1 with different values of p_{astro} by different pipelines—and GW150914, an unambiguous BBH. For a coherent astrophysical signal, the posterior distributions for parameters that can be well constrained from the data should be broadly consistent between the values inferred from each detector individually as well as the coherent network analysis. GW150914 and GW170728 both show significant information gain from Hanford and Livingston, respectively. The posteriors for the new events of interest show that information is gained from each detector individually.

APPENDIX D: POSTERIORS FOR GW170402, GW170817A AND GW170818

In Fig. 12, we show the posterior distributions for the two single detector IAS candidates GW170402 and GW170817A as well as GW170818 for comparison. For the two IAS events, we report $\mathcal{M}_c \simeq 24.0^{+15.7}_{-2.0}$ and $\mathcal{M}_c \simeq 65.5^{+4.9}_{-6.8}$ for GW170402 and GW170817A, respectively. Similarly, we find $\chi_{\mathrm{eff}} \simeq 0.5^{+0.2}_{-0.38}$ and $\chi_{\mathrm{eff}} \simeq 0.31^{+0.13}_{-0.16}$.
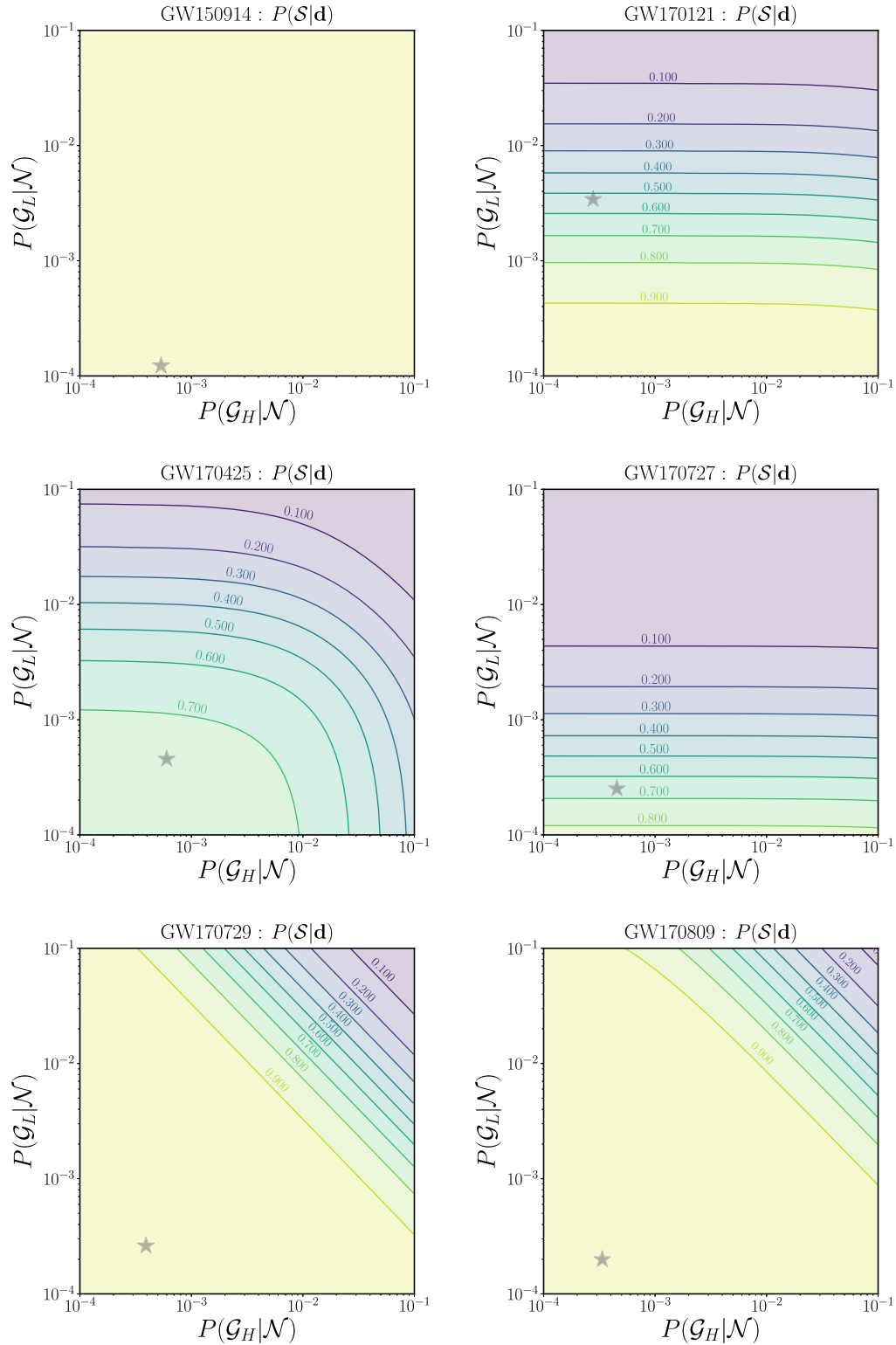
FIG. 7. Posterior probability that the signal is of astrophysical origin $P(\mathcal{S}|\mathbf{d})$ for select events as a function of the detector glitch probabilities $P(\mathcal{G}_H|\mathcal{N})$ and $P(\mathcal{G}_L|\mathcal{N})$ for the LIGO-Hanford and LIGO-Livingston instruments, respectively. The gray star denotes the glitch probability estimated from OMICRON triggers, see Table II. GW150914 is unambiguous, having an astrophysical probability ≈ 1 irrespective of the glitch probability. Similarly, we find that GW170729 has $P(\mathcal{S}|\mathbf{d}) \approx 1$ for the glitch probabilities estimated from a 24h window around the event. For the candidate events GW170121, GW170425 and GW170727, the probability that the signal is of astrophysical origin is more sensitive to prior probability on the glitch and signal hypotheses, respectively.
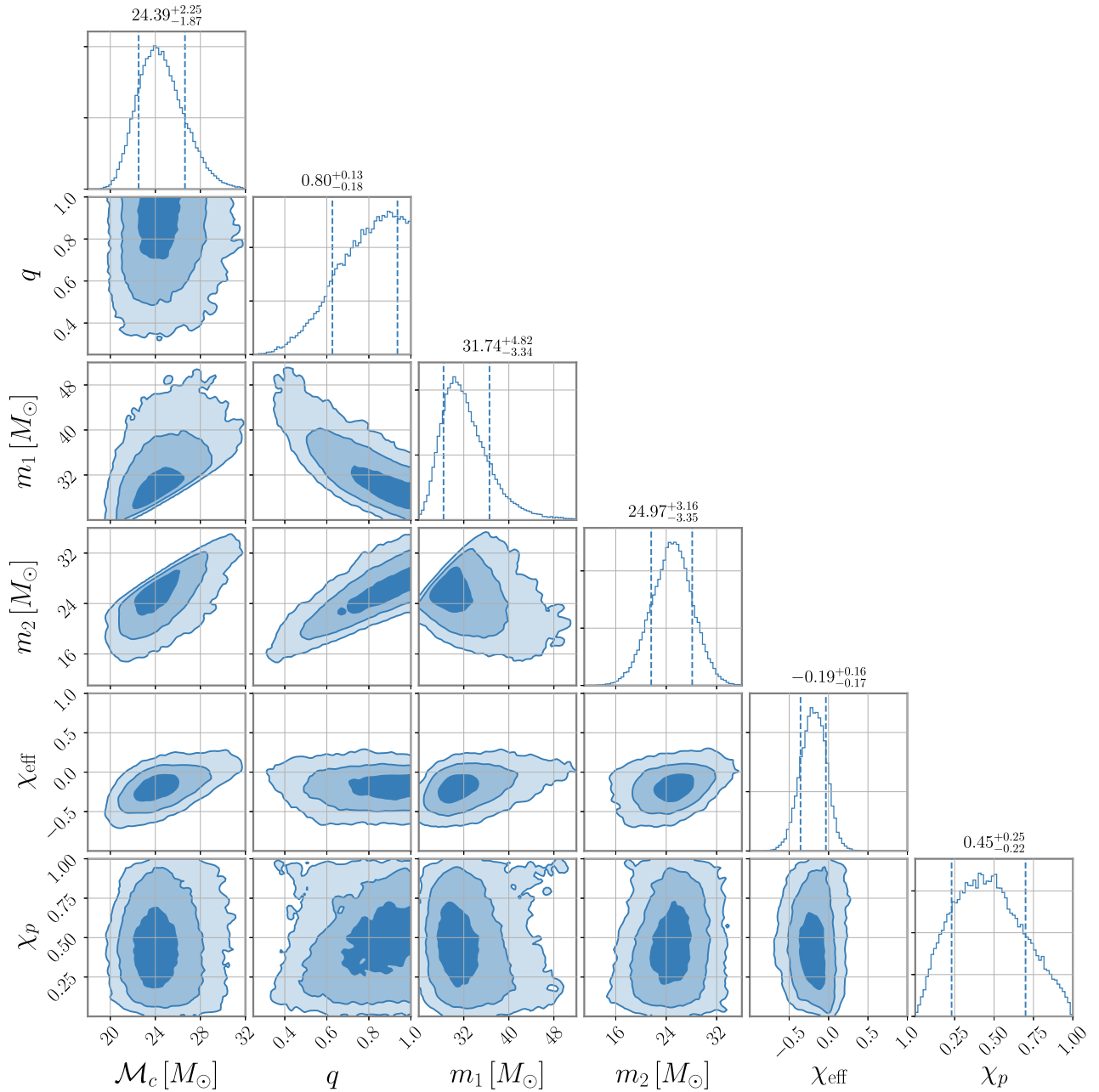
FIG. 8. Posterior distributions for GW170121. We show source frame parameters: chirp mass $\mathcal{M}_c \equiv (m_1 m_2)^{3/5}/(m_1+m_2)^{1/5}$, mass ratio $q \equiv m_2/m_1 \leq 1$, component masses m_1 and m_2, effective aligned spin χ_{eff} and effective precessing spin χ_p [69], see e.g., Eq. (6) and (9), respectively, in [114]. Note that whilst GW170121 has negative effective spin $\chi_{\text{eff}} \simeq -0.19$ [10,113], the intrinsic parameters are broadly compatible with the population of BBHs observed in O1 and O2 [3,35].
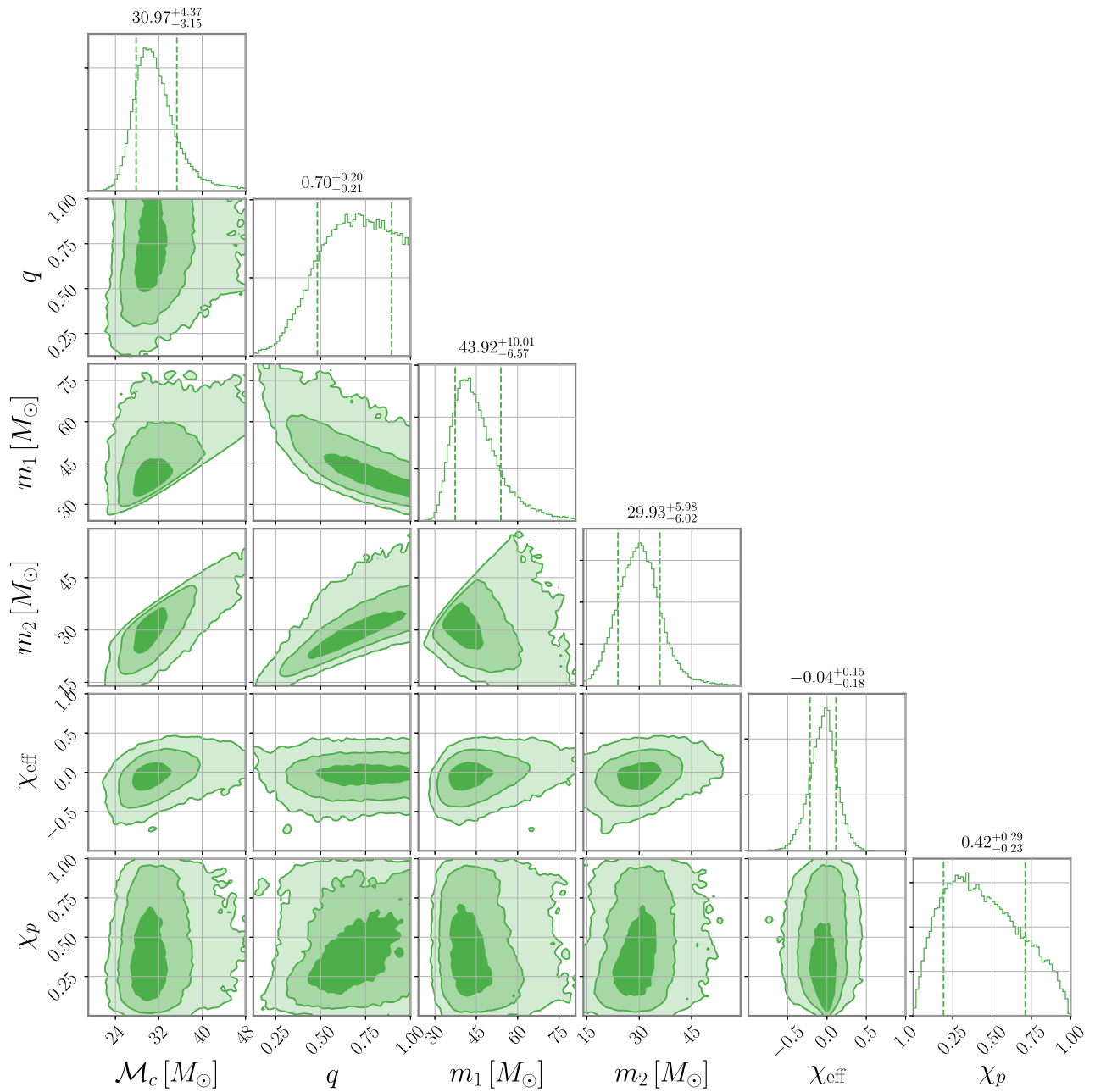
FIG. 9.   Posterior distributions for GW170425 as per Fig. 8.
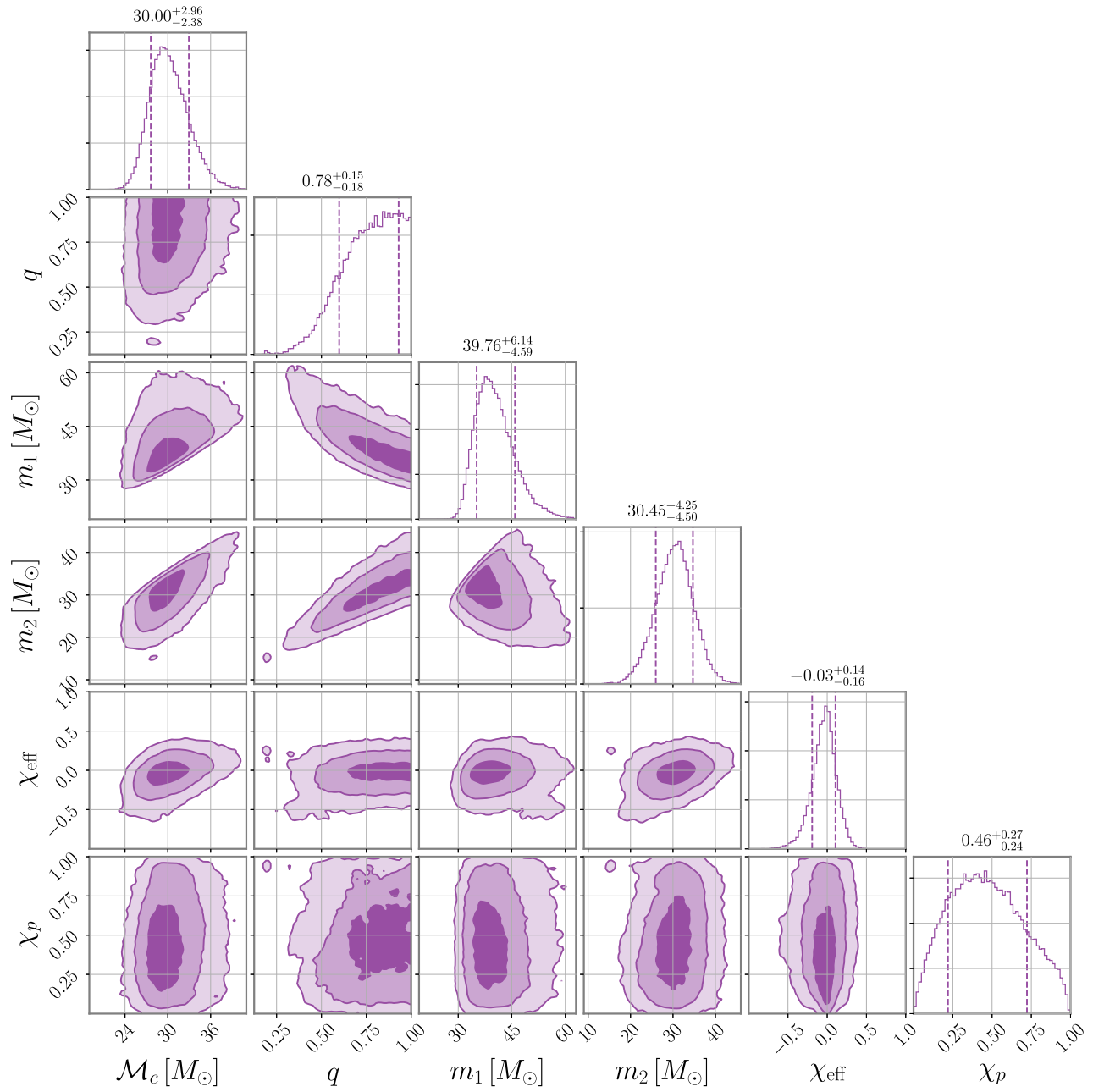
FIG. 10. Posterior distributions for GW170727 as per Fig. 8.
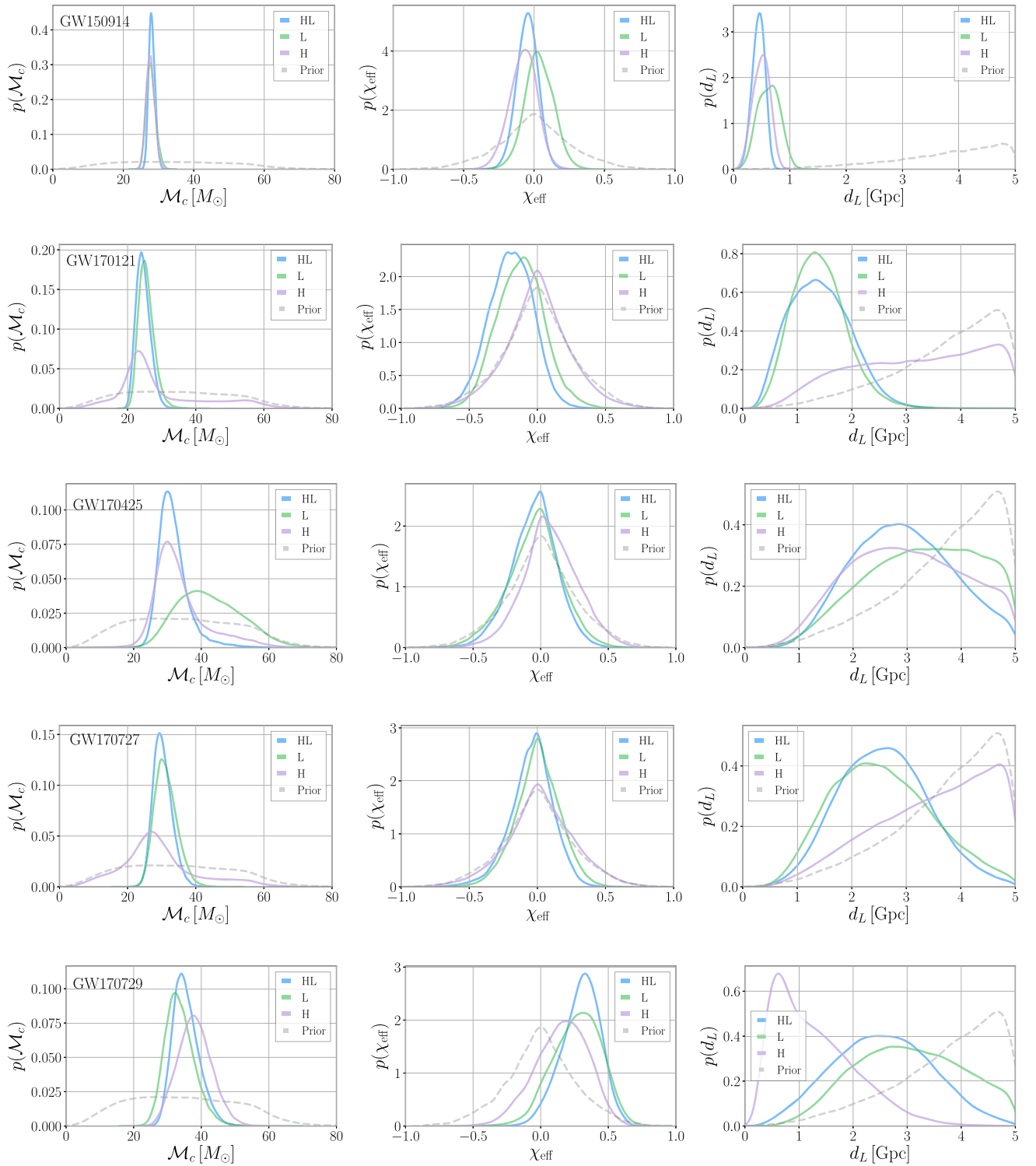
FIG. 11.   For coherent astrophysical signals, the posterior distributions for parameters that can be best measured from the data should be broadly consistent between the values inferred individually from each of the detectors and from the coherent network analysis. Here we show the posteriors for the chirp mass $\mathcal{M}_c$ and effective spin χ_{eff} for the three candidate events GW170121, GW170425 and GW170727 as well as the GWTC-1 events GW150914 and GW170729. All events show evidence for posterior information being driven by both detectors.
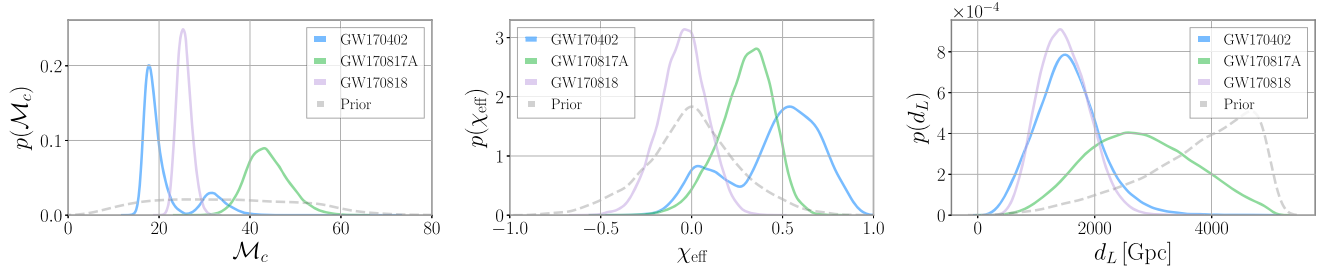
FIG. 12. Posterior distributions for the two IAS single detector triggers, GW170402 and GW170818, as well as GW170818 for comparison. We only show the source frame chirp mass $\mathcal{M}_\downarrow$, the effective aligned spin $\chi_{\rm eff}$, and the luminosity distance d_L. We do not expect our framework to yield a large astrophysical probability for these events as they are in effect single detector triggers.

[1] J. Aasi *et al.* (LIGO Scientific Collaboration), Classical Quantum Gravity **32**, 074001 (2015).

[2] F. Acernese *et al.* (VIRGO Collaboration), Classical Quantum Gravity **32**, 024001 (2015).

[3] B. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. X **9**, 031040 (2019).

[4] A. H. Nitz, T. Dent, G. S. Davies, S. Kumar, C. D. Capano, I. Harry, S. Mozzon, L. Nuttall, A. Lundgren, and M. Tápai, Astrophys. J. **891**, 123 (2019).

[5] A. H. Nitz, T. Dent, G. S. Davies, and I. Harry, Astrophys. J. **897**, 169 (2020).

[6] B. Zackay, L. Dai, T. Venumadhav, J. Roulet, and M. Zaldarriaga, Phys. Rev. D **104**, 063030 (2021).

[7] B. Zackay, T. Venumadhav, J. Roulet, L. Dai, and M. Zaldarriaga, Phys. Rev. D **104**, 063034 (2021).

[8] B. Zackay, T. Venumadhav, L. Dai, J. Roulet, and M. Zaldarriaga, Phys. Rev. D **100**, 023007 (2019).

[9] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, Phys. Rev. D **100**, 023011 (2019).

[10] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, Phys. Rev. D **101**, 083030 (2020).

[11] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. D **102**, 043015 (2020).

[12] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Astrophys. J. **896**, L44 (2020).

[13] W. M. Farr, J. R. Gair, I. Mandel, and C. Cutler, Phys. Rev. D **91**, 023005 (2015).

[14] S. J. Kapadia *et al.*, Classical Quantum Gravity **37**, 045007 (2020).

[15] S. M. Gaebel, J. Veitch, T. Dent, and W. M. Farr, Mon. Not. R. Astron. Soc. **484**, 4008 (2019).

[16] B. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. Lett. **116**, 061102 (2016).

[17] B. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Astrophys. J. Lett. **833**, L1 (2016).

[18] B. F. Schutz, Classical Quantum Gravity **28**, 125023 (2011).

[19] H.-Y. Chen and D. E. Holz, arXiv:1409.0522.

[20] L. Nuttall *et al.*, Classical Quantum Gravity **32**, 245005 (2015).

[21] B. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Classical Quantum Gravity **33**, 134001 (2016).

[22] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Classical Quantum Gravity **35**, 065010 (2018).

[23] M. Cabero *et al.*, Classical Quantum Gravity **36**, 155010 (2019).

[24] D. Davis, L. V. White, and P. R. Saulson, Classical Quantum Gravity **37**, 145001 (2020).

[25] S. A. Usman *et al.*, Classical Quantum Gravity **33**, 215004 (2016).

[26] A. H. Nitz, T. Dent, T. Dal Canton, S. Fairhurst, and D. A. Brown, Astrophys. J. **849**, 118 (2017).

[27] R. Smith and E. Thrane, Phys. Rev. X **8**, 021019 (2018).

[28] G. Ashton, E. Thrane, and R. J. E. Smith, Phys. Rev. D **100**, 123018 (2019).

[29] C. Messick *et al.*, Phys. Rev. D **95**, 042001 (2017).

[30] S. Sachdev *et al.*, arXiv:1901.08580.

[31] B. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. Lett. **116**, 221101 (2016);

[32] B. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. X **6**, 041015 (2016); **8**, 039903(E) (2018).

[33] B. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. D **100**, 104036 (2019).

[34] B. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Astrophys. J. Lett. **818**, L22 (2016).

[35] B. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Astrophys. J. Lett. **882**, L24 (2019).

[36] J. Veitch and A. Vecchio, Phys. Rev. D **81**, 062003 (2010).

[37] M. Isi, R. Smith, S. Vitale, T. J. Massinger, J. Kanner, and A. Vajpeyi, Phys. Rev. D **98**, 042007 (2018).

[38] G. Ashton *et al.*, Astrophys. J. Suppl. Ser. **241**, 27 (2019).

[39] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Classical Quantum Gravity **37**, 055002 (2020).

[40] S. Babak *et al.*, Phys. Rev. D **87**, 024033 (2013).

[41] R. Essick, P. Godwin, C. Hanna, L. Blackburn, and E. Katsavounidis, arXiv:2005.12761.

[42] L. S. Finn and D. F. Chernoff, Phys. Rev. D **47,** 2198 (1993).

[43] S. Chatterji, L. Blackburn, G. Martin, and E. Katsavounidis, Classical Quantum Gravity **21,** S1809 (2004).

[44] F. Robinet, See https://tds.ego-gw.it/ql (2015).

[45] F. Robinet, N. Arnaud, N. Leroy, A. Lundgren, D. Macleod, and J. McIver, SoftwareX **12,** 100620 (2020).

[46] G. Ashton and E. Thrane, Mon. Not. R. Astron. Soc. **498,** 1905 (2020).

[47] A. H. Nitz, Classical Quantum Gravity **35,** 035016 (2018).

[48] S. Mozzon, L. K. Nuttall, A. Lundgren, T. Dent, S. Kumar, and A. H. Nitz, Classical Quantum Gravity **37,** 215014 (2020).

[49] R. Abbott et al. (LIGO Scientific, Virgo Collaborations), SoftwareX **13,** 100658 (2021).

[50] LIGO Scientific Collaboration, Virgo Collaboration, Gravitational Wave Open Science Center: GWTC-1, https://www.gw-openscience.org/GWTC-1/ (2019).

[51] T. B. Littenberg and N. J. Cornish, Phys. Rev. D **91,** 084034 (2015).

[52] N. J. Cornish and T. B. Littenberg, Classical Quantum Gravity **32,** 135012 (2015).

[53] S. Vitale, W. Del Pozzo, T. G. Li, C. Van Den Broeck, I. Mandel, B. Aylott, and J. Veitch, Phys. Rev. D **85,** 064034 (2012).

[54] C. Cahillane et al. (LIGO Scientific Collaboration), Phys. Rev. D **96,** 102001 (2017).

[55] L. Sun et al., Classical Quantum Gravity **37,** 225008 (2020).

[56] K. Chatziioannou, C.-J. Haster, T. B. Littenberg, W. M. Farr, S. Ghonge, M. Millhouse, J. A. Clark, and N. Cornish, Phys. Rev. D **100,** 104004 (2019).

[57] S. Banagiri, M. W. Coughlin, J. Clark, P. D. Lasky, M. Bizouard, C. Talbot, E. Thrane, and V. Mandic, Mon. Not. R. Astron. Soc. **492,** 4945 (2020).

[58] S. Biscoveanu, C.-J. Haster, S. Vitale, and J. Davies, Phys. Rev. D **102,** 023008 (2020).

[59] C. Talbot and E. Thrane, Phys. Rev. Research **2,** 043298 (2020).

[60] P. Welch, IEEE Trans. Audio Electroacoust. **15,** 70 (1967).

[61] J. Veitch et al., Phys. Rev. D **91,** 042003 (2015).

[62] G. Vajente, Y. Huang, M. Isi, J. C. Driggers, J. S. Kissel, M. J. Szczepanczyk, and S. Vitale, Phys. Rev. D **101,** 042003 (2020).

[63] C. Rover, R. Meyer, and N. Christensen, Classical Quantum Gravity **28,** 015010 (2011).

[64] C. Rover, Phys. Rev. D **84,** 122004 (2011).

[65] T. Yamamoto, K. Hayama, S. Mano, Y. Itoh, and N. Kanda, Phys. Rev. D **93,** 082005 (2016).

[66] J. Skilling, Bayesian Anal. **1,** 833 (2006).

[67] J. Veitch and A. Vecchio, Phys. Rev. D **78,** 022001 (2008).

[68] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, Phys. Rev. Lett. **113,** 151101 (2014).

[69] P. Schmidt, F. Ohme, and M. Hannam, Phys. Rev. D **91,** 024043 (2015).

[70] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza, and A. Bohé, Phys. Rev. D **93,** 044006 (2016).

[71] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Phys. Rev. D **93,** 044007 (2016).

[72] T. Callister, J. Kanner, T. Massinger, S. Dhurandhar, and A. Weinstein, Classical Quantum Gravity **34,** 155007 (2017).

[73] D. Wysocki, J. Lange, and R. O'Shaughnessy, Phys. Rev. D **100,** 043012 (2019).

[74] I. Mandel, W. M. Farr, and J. R. Gair, Mon. Not. R. Astron. Soc. **486,** 1086 (2019).

[75] M. Fishbach, D. E. Holz, and W. M. Farr, Astrophys. J. Lett. **863,** L41 (2018).

[76] C. Talbot and E. Thrane, Astrophys. J. **856,** 173 (2018).

[77] Z. Barkat, G. Rakavy, and N. Sack, Phys. Rev. Lett. **18,** 379 (1967).

[78] S. E. Woosley and T. A. Weaver, Radiation Hydrodynamics in Stars and Compact Objects, Proceedings of IAU Colloq. 89, held at Copenhagen University, June 11-20, 1985, Lecture Notes in Physics, (Springer-Verlag Publishers, 1986).

[79] A. Heger and S. Woosley, Astrophys. J. **567,** 532 (2002).

[80] E. Chatzopoulos and J. C. Wheeler, Astrophys. J. **748,** 42 (2012).

[81] S. Woosley, Astrophys. J. **836,** 244 (2017).

[82] C. Talbot and E. Thrane, Phys. Rev. D **96,** 023012 (2017).

[83] C. Talbot, R. Smith, E. Thrane, and G. B. Poole, Phys. Rev. D **100,** 043030 (2019).

[84] S. Galaudage, C. Talbot, and E. Thrane, Phys. Rev. D **102,** 083026 (2020).

[85] J. Roulet, T. Venumadhav, B. Zackay, L. Dai, and M. Zaldarriaga, Phys. Rev. D **102,** 123022 (2020).

[86] B. Abbott et al. (KAGRA, LIGO Scientific, VIRGO Collaborations), Living Rev. Relativity **21,** 3 (2018).

[87] Y. Aso, Y. Michimura, K. Somiya, M. Ando, O. Miyakawa, T. Sekiguchi, D. Tatsumi, and H. Yamamoto (KAGRA Collaboration), Phys. Rev. D **88,** 043007 (2013).

[88] T. Akutsu et al. (KAGRA Collaboration), Nat. Astron. **3,** 35 (2019).

[89] B. Abbott et al. (LIGO Scientific, Virgo Collaborations), Phys. Rev. Lett. **119,** 161101 (2017).

[90] A. Samajdar, J. Janquart, C. Van Den Broeck, and T. Dietrich, Phys. Rev. D **104,** 044003 (2021).

[91] E. Pizzati, S. Sachdev, A. Gupta, and B. Sathyaprakash, arXiv:2102.07692.

[92] Y. Himemoto, A. Nishizawa, and A. Taruya, Phys. Rev. D **104,** 044010 (2021).

[93] P. Relton and V. Raymond, Phys. Rev. D **104,** 084039 (2021).

[94] G. Pratten, S. Husa, C. Garcia-Quiros, M. Colleoni, A. Ramos-Buades, H. Estelles, and R. Jaume, Phys. Rev. D **102,** 064001 (2020).

[95] A. Nagar, G. Pratten, G. Riemenschneider, and R. Gamba, Phys. Rev. D **101,** 024041 (2020).

[96] A. Nagar, G. Riemenschneider, G. Pratten, P. Rettegno, and F. Messina, Phys. Rev. D **102,** 024077 (2020).

[97] C. García-Quirós, M. Colleoni, S. Husa, H. Estellés, G. Pratten, A. Ramos-Buades, M. Mateu-Lucena, and R. Jaume, Phys. Rev. D **102,** 064002 (2020).

[98] G. Pratten et al., Phys. Rev. D **103,** 104056 (2021).

[99] S. Ossokine et al., Phys. Rev. D **102,** 044055 (2020).

[100] S. Vinciguerra, J. Veitch, and I. Mandel, Classical Quantum Gravity **34,** 115006 (2017).

[101] C. García-Quirós, S. Husa, M. Mateu-Lucena, and A. Borchers, Classical Quantum Gravity **38,** 015006 (2021).

[102] P. Canizares, S. E. Field, J. Gair, V. Raymond, R. Smith, and M. Tiglio, Phys. Rev. Lett. **114**, 071104 (2015).

[103] R. Smith, S. E. Field, K. Blackburn, C.-J. Haster, M. Pürrer, V. Raymond, and P. Schmidt, Phys. Rev. D **94**, 044031 (2016).

[104] V. Tiwari, Classical Quantum Gravity **35**, 145009 (2018).

[105] D. Gerosa, G. Pratten, and A. Vecchio, Phys. Rev. D **102**, 103020 (2020).

[106] M. Fishbach, W. M. Farr, and D. E. Holz, Astrophys. J. Lett. **891**, L31 (2020).

[107] S. Miller, T. A. Callister, and W. Farr, Astrophys. J. **895**, 128 (2020).

[108] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Astrophys. J. Lett. **913**, L7 (2021).

[109] https://www.gw-openscience.org.

[110] R. O'Shaughnessy, V. Kalogera, and K. Belczynski, Astrophys. J. **716**, 615 (2010).

[111] P. Madau and M. Dickinson, Annu. Rev. Astron. Astrophys. **52**, 415 (2014).

[112] W. M. Farr, N. Sravan, A. Cantrell, L. Kreidberg, C. D. Bailyn, I. Mandel, and V. Kalogera, Astrophys. J. **741**, 103 (2011).

[113] Y. Huang, C.-J. Haster, S. Vitale, A. Zimmerman, J. Roulet, T. Venumadhav, B. Zackay, L. Dai, and M. Zaldarriaga, Phys. Rev. D **102**, 103024 (2020).

[114] B. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. Lett. **116**, 241102 (2016).