A Bayesian ranking statistic to find high-mass black holes in LIGO–Virgo data

(Dated: March 13, 2021)

The detection of intermediate-mass black holes ($10^2 - 10^6$ $M_\odot$) will shed light on the formation of supermassive black holes and thus galaxy formation. Although LIGO is sensitive to the merger of binary black holes with total masses up to 400 $M_\odot$, only 4 of their 50 detections have a total mass > 100 $M_\odot$ with >95% credibility. A possible explanation for the absence of intermediate-mass events may be their misclassification as short-duration instrumental noise transients. Short-duration instrumental transients mimic the short-duration gravitational-wave signals from intermediate-mass binary black hole mergers. [ET: I think the preceding text is misleading. We make it sound like we expect a uniform distribution of total mass, and so it's a surprising that we don't see IMBH. However, in reality, it is theoretically challenging to make IMBH. You need an exceptionally massive progenitor star to avoid pair instability supernova. Or, maybe you can make IMBH from hierarchical mergers or accretion. No one expects them to be common, but they are interesting because of their connection to supermassive black holes. I suggest rewriting the preceding text to explain how IMBH might form and why they are interesting to look for.] Here we demonstrate a ranking statistic utilising Bayesian inference for the detection of high-mass binary black hole mergers (with a total mass > 55 $M_\odot$). [ET: Claims of beating matched filtering not yet borne out.] We apply this technique on the high-mass triggers during LIGO's second observing run to search for previously unresolved gravitational-wave signals from high-mass binary black holes. Our analysis does not discover new gravitational-wave events. However, we find support for some previously identified borderline detections. [ET: Don't oversell Bayesian measure since our method is only semi-Bayesian.]

I. NEW INTRO

In the 1960s, X-ray and optical observations from the Cyg-X1 binary star system led to the discovery of the first stellar-mass black hole []. Shortly afterwards, in 1974, kinematic measurements of stars near the center of our galaxy demonstrated the presence of a supermassive black hole of mass 10^6 $M_\odot$ – Sag A* []. These two detections firmly established the existence of stellar-mass black holes $M_{BH} < 10^2$ $M_\odot$ and supermassive black holes, $M_{BH} > 10^5$ $M_\odot$ in our universe. Since then, astronomers accumulated more definitive evidence for stellar-mass and supermassive black holes. However, a gap in between the mass-ranges of these two categories began to form: there is a deficiency of black holes in the intermediate-mass range $10^2 - 10^5$ $M_\odot$. These intermediate-mass black holes must have existed to form the supermassive black holes detected to exist in the early universe []. However, we only have a handful of candidates and one confirmed intermediate-mass black hole []. The study and discovery of intermediate-mass black holes will help bridge the gap and provide a pathway to forming supermassive black holes.

Intermediate-mass black holes may form various scenarios ranging from hierarchical mergers of stellar-mass black holes to large metal-free stars' direct collapse. Hierarchical mergers in active galactic nuclei migration-traps, globular and stellar clusters may be a feasible method for black holes to accumulate mass and reach the intermediate-mass range. However, when black holes merge, they can receive recoil-kicks up to ∼2,500 km/s. Hence, the merger remnants require their recoil kick to be less than its environment's escape velocity to undergo successive mergers. This tight constraint makes hierarchical mergers a challenging scenario.

Another pathway to form intermediate-mass black involves the direct collapse of massive, metal-free Population III stars. Typically, massive stars > 100 $M_\odot$ can eject more than half of their mass in the form of radiatively driven stellar winds. However, at zero metallicity, mass loss through stellar winds is negligible (Kudritzki 2000). Hence, Population III stars have the opportunity to complete their life-cycles without losing a significant portion of their mass. Moreover, Population III stars with masses above 230 $M_\odot$ will avoid electron-positron pair-instability. Consequently, oxygen and silicon burning will not drive an explosion, allowing the star to undergo a complete collapse to leave behind an intermediate-mass black hole remnant with a mass of $120-280$ $M_\odot$, dependant on the initial star mass (Bond et al. 1984). The intermediate-mass black holes that form as remnants of these large Population III stars may exist today in clusters located at the galaxies' bulges. If these bulges contain intermediate-mass black holes, they could also provide evidence for Population III stars' existence.

As there is inadequate observational evidence for intermediate-mass black holes, it is challenging to state which formation path is favoured. However, the observational work does allude to the existence of intermediate-mass black holes. The search for intermediate-mass black holes is underway using direct kinematic measurements, reverberation mapping, studying luminosity from low-luminous active galactic nuclei and Ultraluminous X-ray Sources.

Dynamic measurements of individual sources orbiting an intermediate-mass black hole is that the sphere of influence too small to resolve motions. Can only be used on close sources. The issue with reverberation mapping and the other EM spectra, cant be sure of anisotropic emission, non-linear emission. large uncertainties. `https://iopscience.iop.org/article/10.3847/2041-8213/ab745b`

The most promising avenues to find $100 - 1000 Msun$

black holes are either to identify stellar binaries, which are expected to be common or to detect them in gravitational radiation. Ground based gravitational-wave limits on the low end of this mass regime are continuing to improve.

LIGO found 190521, can find more if we are more sensitive, and if we can better harness data

II. INTRODUCTION

Since the 1970s, there has been an accumulation of evidence for stellar-mass and supermassive black holes. In April 2019, the Event Horizon Telescope provided the first visual evidence of the supermassive black hole M87 [1]. As of January 2021, the LIGO Scientific Collaboration has confirmed ~ 50 binary black hole systems and listed numerous candidate events [2–4]. Additionally, since the public release of LIGO's first and second observing run's data (O1 and O2), search pipelines such as PyCBC [5] and a research team at the Institute for Advanced Study (IAS) [6–8], have independently identified additional binary black hole gravitational-wave candidates.

[ET: Add a paragraph here discussing basic ideas in IMBH theory. Explain why the are not easy to make. Discuss scenarios by which they can be made. Discuss the role they may play in the formation of supermassive BH. Then, we transition to the next paragraph about observational evidence.]

Until recently, there was no definitive evidence for intermediate-mass black holes with a mass in between the range of stellar-mass and supermassive black holes ($10^2 - 10^6$ M$_\odot$). [ET: There have been numerous claims of IMBH detection. The previous sentence is correct that there are no *definitive* detections before GW190521, but the paragraph is misleading by omission. Insert here an exhaustive review of previous IMBH detection claims, noting who made the claim, the mass, and the method.] This changed with the detection of GW190521, a unique gravitational-wave event that led to the formation of a black hole with a mass 142 M$_\odot$, the first confirmed direct discovery of an intermediate-mass black hole [9]. Although this is the first gravitational-wave observation of an IMBH, ground-based gravitational-wave detectors are sensitive to gravitational-waves from even more massive systems with a total mass of up to 400 M$_\odot$ [2].

[ET: The following paragraph tries to solve a problem that doesn't exist: as noted in comments above, the absence of GWs from IMBH is not surprising. Delete or rewrite if you want to emphasise the difficulty in IMBH detection.] A possible explanation for the absence of intermediate-mass events may be due to their misclassification as short-duration instrumental noise transients known as glitches [10–12]. These glitches can mimic astrophysical signals and hence decrease the significance of true gravitational-wave events.

One method to account for glitches while searching for gravitational-waves from coalescing compact binaries (CBCs) is by utilising the astrophysical Bayesian odds [13–17]. A true Bayesian odds, calculated without using bootstrap techniques, can provide more accurate estimate of significance that does note depend on the search pipeline [15–17]. In this paper, we utilise a Bayesian method, called the Bayesian Coherence Ratio $\rho_{\rm BCR}$ [14], to rank the candidate gravitational-wave signals from high-mass compact binary coalescences (systems with total masses in the range of $55 - 500$ M$_\odot$) in the detector data recorded during O2. Although the $\rho_{\rm BCR}$, utilising bootstrap techniques, does not provide the true Bayesian odds, it provides a more astrophysical measure of candidate events' significance than a traditional match-filter significance. [ET: This pitch, that the method is "a more astrophysical measure" is unconvincing. (What does that even mean?) Chat with RS/ET to come up with a better pitch built around the idea that the detection statistic is built on Bayesian inferences results, and that it's a step toward fully Bayesian detection.]

We find that (a) high-mass events reported in the GWTC-1, including GW170729 (the least significant event in GWTC-1) have high significance; (b) high-mass events detected from the IAS group have differing levels of significance; [ET: differing from what? edit this passage to state more clearly that we find statistical support for some of their candidates, but not for others] and that (c) our ranking statistic does not identify any intermediate-mass black holes, but does identify an unreported stellar-mass binary black hole candidate, 170222 [ET: State statistical significance.].

The remainder of this paper is structured as follows. We outline our methods, including details of our ranking statistic and the retrieval of our candidate events in Section III. We present details on the implementation of our analysis in Section IV. Finally, we present our results in Section V, and discuss these results in the context of the significance of gravitational-wave candidates in Section VI.

III. METHOD

The standard framework to identify CBC gravitational-wave signals hidden in data is by quantifying the significance of candidates with null-hypothesis significance testing. In this framework, the candidates' ranking statistic is compared against a background distribution in a frequentist approach. On the other hand, the standard framework for performing parameter estimation and model selection in gravitational-wave astronomy is Bayesian inference. This work utilises Bayesian inference to calculate the Bayesian Coherence odds-ratio [14], $\rho_{\rm BCR}$, of high-mass candidates in LIGO's second observing run. We use the $\rho_{\rm BCR}$ not as an odds-ratio but instead as a ranking statistic, a step toward building a unified Bayesian framework to search for candidates and estimate their parameters.

A. The Bayesian Coherence Ratio

Bayes theorem states that the posterior probability distribution $p(\vec{\theta}|d, \mathcal{H})$ for data d and a vector of parameters $\vec{\theta}$ that describe a model which quantifies a hypothesis $\mathcal{H}$, is given by

$$p(\vec{\theta}|d, \mathcal{H}) = \frac{\mathcal{L}(d|\vec{\theta}, \mathcal{H})\ \pi(\vec{\theta}|\mathcal{H})}{\mathcal{Z}(d|\mathcal{H})}\ , \quad (1)$$

where $\mathcal{L}(d|\vec{\theta}, \mathcal{H})$ is the likelihood of the data given the parameters $\vec{\theta}$ and the hypothesis, $\pi(\vec{\theta}|\mathcal{H})$ is the prior probability of the parameters, and finally,

$$\mathcal{Z}(d|\mathcal{H}) = \int_{\vec{\theta}} \mathcal{L}(d|\vec{\theta}, \mathcal{H})\ \pi(\vec{\theta}|\mathcal{H})\ d\vec{\theta} \quad (2)$$

is the likelihood after marginalising over the parameters $\vec{\theta}$. To compare two hypotheses $\mathcal{H}_A$ and $\mathcal{H}_B$ with the Bayes theorem one can calculate an odds-ratio

$$\mathcal{O}_B^A = \frac{\mathcal{Z}^A\ \pi(\vec{\theta}^A)}{\mathcal{Z}^B\ \pi(\vec{\theta}^B)}\ , \quad (3)$$

where $\mathcal{Z}^A$ and $\mathcal{Z}^B$ are the shorthand for the evidences $\mathcal{Z}(d|\mathcal{H}_A)$ and $\mathcal{Z}(d|\mathcal{H}_B)$. The odds-ratio can tell us which of the two hypotheses is more likely. For example, if $\mathcal{O}_B^A >> 1$, then this odds ratio indicates that the $\mathcal{H}_A$ describes the data much better than $\mathcal{H}_B$.

The $\rho_{\rm BCR}$ is a Bayesian odds-ratio like the above, of a coherent signal hypotheses $\mathcal{H}_S$ and an incoherent instrumental feature hypothesis $\mathcal{H}_I$ (the null-hypotheses) for a network of D detectors. $\mathcal{H}_\mathcal{I}$ states that each detector i has either pure stationary Gaussian noise $\mathcal{H}_N$ or Gaussian noise and an incoherent noise transient (glitch). Taking $\mathcal{H}_G$, Z^S, Z_i^G and Z_i^N as the Bayesian evidences (marginalised likelihoods) for $\mathcal{H}_S$, $\mathcal{H}_N$, and $\mathcal{H}_G$, the $\rho_{\rm BCR}$ is given by

$$\rho_{\rm BCR} = \frac{\alpha Z^S}{\prod_{i=1}^{D} \left[\beta Z_i^G + (1-\beta) Z_i^N\right]}\ , \quad (4)$$

where α and β are the prior-odds of obtaining a signal or a glitch from a stretch of data. The prior-odds can be defined more explicitly as

- $\alpha = P(\mathcal{H}_S)/P(\mathcal{H}_I)$, the prior-odds for obtaining a coherent signal versus an incoherent instrumental feature.
- $\beta = P(\mathcal{H}_G|\mathcal{H}_I)$, the prior-odds for obtaining a glitch assuming there is an incoherent instrumental feature.

When $\mathcal{H}_S$ and $\mathcal{H}_I$ are precisely described and the correct prior-odds are known, the $\rho_{\rm BCR}$ is a Bayesian odds-ratio. As an odds-ratio, the $\rho_{\rm BCR}$ is the optimal discriminator between coherent signals and incoherent instrumental features. However, as the priors-odds are unknown, it is invalid to use the $\rho_{\rm BCR}$ as an odds-ratio to make an informed decision about whether a candidate is from an astrophysical or terrestrial source. Instead of interpreting the $\rho_{\rm BCR}$ as a Bayesian odds-ratio, it can be used as a ranking statistic. Using the $\rho_{\rm BCR}$ as a ranking statistic we can obtain a frequentist significance of a candidate $\rho_{\rm BCR}$-value measured against a background $\rho_{\rm BCR}$ distribution.

When using the $\rho_{\rm BCR}$ as a detection statistic, the physical interpretation of the prior-odds is lost. Hence, the prior-odds are empirically tuned to maximise the separation between the $\rho_{\rm BCR}$ distribution of the background (expected to favour the $\mathcal{H}_I$ hypothesis) and the $\rho_{\rm BCR}$ distribution of artificially manufactured simulated signals (expected to favour the $\mathcal{H}_S$ hypothesis). Increasing the separation between the two distributions can improve ability of the $\rho_{\rm BCR}$ to discriminate candidate events as coherent signals or incoherent instrumental features. The tuning process is described in detail in Appendix B.

B. Calculating the Significance of Candidates

Candidate $\rho_{\rm BCR}$-values are either statistically insignificant compared to the background $\rho_{\rm BCR}$ distribution, implying the candidate is more probable to be an incoherent instrumental feature (the $\mathcal{H}_I$ null-hypothesis), or statistically significant to the background distribution, indicating the possible presence of an astrophysical signal (the $\mathcal{H}_S$ hypothesis). A false alarm probability with trial factors, FAP, for the candidate $\rho_{\rm BCR}$-values can quantify the significance. The FAP is the probability that a candidate originating from a non-astrophysical source can be is falsely identified as a signal.

To calculate the FAP, each candidate $\rho_{\rm BCR}$ is considered a single statistical trail that can occur at a fixed false alarm probability f, where f is the probability of observing a background $\rho_{\rm BCR}$ greater than or equal to the candidate $\rho_{\rm BCR}$,

$$f = \frac{\text{Count of } {\rho_{\rm BCR}}' \leq \rho_{\rm BCR}}{\text{Count of } {\rho_{\rm BCR}}'}\ . \quad (5)$$

The false alarm probability with trails FAP that the $\rho_{\rm BCR}$ measurement occurs at least once for N trials ($N > 0$), where N is the number of candidate triggers is

$$\text{FAP} = 1 - (1-f)^N\ . \quad (6)$$

Finally, the FAP can be used to construct a $p_{\rm astro}$, the probability that a signal is of astrophysical origin [18–20]

$$p_{\rm astro} = 1 - \text{FAP}\ . \quad (7)$$

C. Data for Analysis

[AV: Probably need a better section label...] The LIGO Scientific collaboration operates several search pipelines that scan for gravitational-waves from compact binary mergers such as `GstLAL`, `MBTA`, `SPIIR` and `PyCBC` [2]. The output of `PyCBC`'s search is a list of times and their corresponding `PyCBC` ranking statistic $\rho_{\rm PC}$ values. The $\rho_{\rm PC}$ ranking-statistic is akin to the matched-filter signal-to-noise ratio ρ. However, unlike ρ, $\rho_{\rm PC}$ includes candidate signal's intrinsic and extrinsic properties and other information that feeds into determining if the signal can have astrophysical origins [21]. Whenever a local maximum of $\rho_{\rm PC} > \rho_{\rm T}$, where $\rho_{\rm T}$ is some predetermined threshold value, the `PyCBC` search pipeline produces a single-detector *trigger* associated with the detector and time where the apparent signal in the data has its merger [21].

When `PyCBC` observes a trigger between detectors with coincident parameters and a time of arrival difference less than the gravitational-wave travel time between detectors, the trigger is labelled a *candidate event trigger*, a trigger that may be from astrophysical origins [22]. To test the pipeline's sensitivity `PyCBC` also conducts searches for *simulated triggers*, artificial triggers manufactured by injecting signals into the detector data. Finally, to quantify the statistical significance of candidate triggers, `PyCBC` artificially constructs *background triggers* to compare against the candidate events. These background triggers are coherent signal-free events, constructed by applying relative offsets, or time-slides, between the data of different detectors [21]. The background trigger's $\rho_{\rm PC}$ distribution is used to calculate the candidate trigger's significance, using null-hypothesis significance testing, under the assumption that all candidate event triggers are due to noise.

Our work demonstrates that the $\rho_{\rm BCR}$ can be used in the same way as $\rho_{\rm PC}$ to measure candidate triggers' statistical significance. The $\rho_{\rm BCR}$ can be a powerful ranking statistic as it incorporates information of not only all possible binary black hole systems that might have merged to produce the trigger but also the various incoherent glitches that might cause a false-detection.

IV. ANALYSIS

A. Acquisition of triggers

Advanced LIGO's second observing run O2 lasted 38 weeks [23]. The software package, `PyCBC` [5], was used by LIGO to process the O2 data in 22 time-frames (approximately 2 weeks for one time-frame) and found several gravitational-wave events and numerous gravitational-wave candidates [21, 22, 24–28]. Some candidate events were vetoed to be glitches, while others were rejected due to their low significance. The data is divided into these time-frames because the detector's sensitivity does not stay constant throughout the eight-month-long observing period.

In addition to finding candidate events, `PyCBC` also identified several million background triggers for each time-frame, by searching background data manufactured by time-sliding data within that time-frame. The background triggers help quantify the candidate events' significance for the respective time-frames. Finally, to test the search's sensitivity, `PyCBC` produced and searched for thousands of simulated signals.

TABLE I. High-mass parameter space (parameters corresponding to signals with durations < 454 ms).

	Minimum	Maximum
Component Mass 1 [$M_\odot$]	31.54	491.68
Component Mass 2 [$M_\odot$]	1.32	121.01
Total Mass [$M_\odot$]	56.93	496.72
Chirp Mass [$M_\odot$]	8.00	174.56
Mass Ratio	0.01	0.98

For our study, we filter the background, simulated and candidate events to include only high-mass events with masses in the ranges of the parameters presented in Table I. A plot of the `PyCBC` triggers from one time-frame, during April 23 - May 8, 2017, is presented in Figure 1. This figure also depicts the gravitational-wave templates used during the search through this time-frame of data.

B. Calculating the BCR for triggers

To evaluate Z^S, Z_i^G and Z_i^N and calculate the $\rho_{\rm BCR}$ Eq. 4 for triggers, we carry out Bayesian inference with Bilby [29, 30], employing dynesty [31] as our nested sampler. Nested sampling, an algorithm introduced by Skilling [32, 33], provides an estimate of the true Bayesian evidence and is often utilised for parameter estimation within the LIGO collaboration [29, 34, 35].

The most computationally intensive step during Bayesian inference is evaluating the likelihood $\mathcal{L}(d_i|\mu(\vec{\theta}))$. To accelerate our analysis, we use a likelihood that explicitly marginalises over coalescence time, phase at coalescence, and luminosity distance (Eq. 80 from Thrane and Talbot [36]). While this marginalised likelihood reduces the run time without introducing errors to our evidence evaluation, it does not generates samples for the marginalised parameters. However, these parameter samples can be calculated as a post-processing step [36].

We set the priors $\pi(\vec{\theta}|\mathcal{H}_S)$ and $\pi(\vec{\theta}|\mathcal{H}_G)$ to be identical. These priors restrict signals with mass parameters in the ranges presented in Table I. The spins are aligned over a uniform range for the dimensionless spin magnitude from $[0, 1]$. The luminosity distance prior assigns probability uniformly in comoving volume, with an upper cutoff of 5 Gpc. The full list of the priors, along with their shapes, limits and boundary conditions are documented in Table II.

The waveform template we utilise is IMRPhenomPv2, a phenomenological waveform template constructed in the frequency domain that models the inspiral, merger, and ringdown (IMR) of a compact binary coalescence [38]. Although there exist gravitational-wave templates such as SEOBNRv4PHM [39] which incorpo-

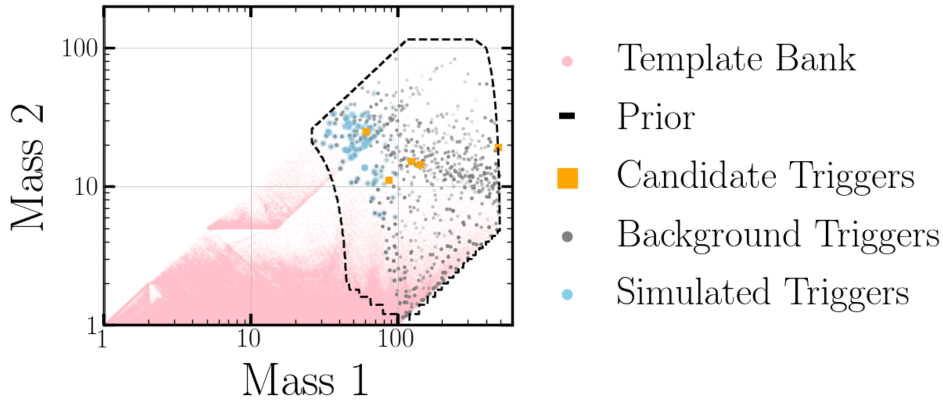

FIG. 1. The template bank (pink) used by PyCBC to search a section of O2 data from April 23 - May 8, 2017. Our search is constrained to the high-mass parameter space enclosed by the dashed line. The candidate, background and simulated triggers detected in this region of the parameter space during this period are plotted in orange, grey and blue respectively. [AV: Should I remove the 'template bank' and only display the various triggers?]

TABLE II. Prior settings for the parameters used during our parameter estimation. The definitions of the parameters are documented in Romero-Shaw *et al.* [37] Table E1.

Parameter	Shape	Limits	Boundary
$\mathcal{M}/\mathrm{M}_\odot$	Uniform	7–180	Reflective
q	Uniform	0.1–1	Reflective
$M/\mathrm{M}_\odot$	Constraint	50–500	–
$d_\mathrm{L}/\mathrm{Mpc}$	Comoving	100–5000	Reflective
a_1, a_2	Uniform	0–1	Reflective
θ_{JN}	Sinusoidal	0–π	Reflective
ψ	Uniform	0–π	Periodic
ϕ	Uniform	0–2π	Periodic
ra	Uniform	0–2π	Periodic
dec	Cosine	0–2π	Reflective

rate more physics, such as information on higher-order modes, we use IMRPhenomPv2 as it is computationally inexpensive compared to others.

We take 31 neighbouring, off-source, non-overlapping, 4-second segments of time-series data before the analysis data segment d_i to generate the PSD. We use off-source to avoid the inclusion of a signal in the PSD calculation. A Tukey window with 0.2-second roll-offs is applied to each data segment to suppress spectral leakage after which the segments are fast-Fourier transformed and median-averaged to create a PSD [40]. Like other PSD estimation methods, this method adds statistical uncertainties to the PSD [41, 42]. To marginalise over the statistical uncertainty, we use the median-likelihood presented by Talbot and Thrane [41] as a post-processing step. [AV: Greg would like me to discuss the effect of the PSD marginalization in more detail]

Finally, we neglect detector calibration uncertainty and acquire data from from the gravitational-wave Open Science Center [23]. The data we use is the publicly accessible O2 strain data from the Hanford and Livingston detectors, recorded while the detectors are in "Science Mode". We obtain the data using GWpy [43].

C. Assigning p_astro to candidate events

After the calculating the ρ_BCR for the entire set of high-mass background and simulated triggers, we calculate probability distributions $p_b(\rho_\mathrm{BCR})$ and $p_s(\rho_\mathrm{BCR})$ for each 2-week time-frame of O2 data. These distributions are used to 'tune' prior-odd α and β values.

Using the tuned prior odds the ρ_BCR for the candidate events can be calculated. Figure 2 shows the ρ_BCR distributions for the background triggers, simulated triggers and candidate events. The bulk of the background and simulated trigger distributions are separate but slightly overlap due to some of the simulated signal's being very faint. The separation suggests that the ρ_BCR can successfully distinguish signals from noise or glitches. The vertical lines in Figure 2 displays the ρ_BCR for gravitational-wave candidate events. On comparing the candidate event ρ_BCR values with the background distribution, we can estimate p_astro values for the candidate events.

V. RESULTS

We analyse the 60 996 background, 5 146 simulated, and 25 candidate triggers reported by PyCBC's search on the data from LIGO's second observing run, restricting our analysis to the triggers that fall within our mass-space as described in Section III. In addition to these triggers, we also analyse events and candidate events reported by GWTC-1 and the IAS group (note that some of these were identified as candidates by the PyCBC search). In Table III, we summarise the $p_\mathrm{astro}^\mathrm{BCR}$, along with the

TABLE III. The $p_{\rm astro}$ of gravitational wave events from various detection pipelines, along with the event candidates with $p_{\rm astro}^{\rm BCR} > 0.3$. Only the candidates and events within our prior space are displayed. The various pipeline $p_{\rm astro}$ represented in this table, $p_{\rm astro}^{\rm ext}$, are from the following pipelines: GstLAL $^{\heartsuit}$ [2], PyCBC $^{\spadesuit}$ [2], PyCBC OGC-2 $^{\clubsuit}$ [44], PyCBC 'single-search' $^{\diamondsuit}$ [45], IAS $^{\bigstar}$ [7, 8], and Pratten and Vecchio [17]'s significances $^{\blacktriangle}$. The catalogues labelled IAS-1 and IAS-2 correspond to the candidates published in Venumadhav *et al.* [7] and Zackay *et al.* [8].

Event	Catalogue	$p_{\rm astro}^{\rm BCR}$	$p_{\rm astro}^{\rm ext}$	GPS
GW170104	GWTC-1	0.94	$1.00^{\heartsuit}; 1.00^{\spadesuit}; 1.0^{\blacktriangle}$	1167559934.60
GW170121	IAS-1	0.76	$1.00^{\clubsuit}; 1.00^{\bigstar}; 0.53^{\blacktriangle}$	1169069152.57
170222	-	0.49	-	1171814474.97
170302	IAS-1	0.63	$0.45^{\bigstar}; 0.0^{\blacktriangle}$	1172487815.48
GW170304	IAS-1	0.83	$0.70^{\clubsuit}; 0.99^{\bigstar}; 0.03^{\blacktriangle}$	1172680689.36
GWC170402	IAS-2	0.38	$0.68^{\bigstar}; 0.03^{\diamondsuit}; 0.0^{\blacktriangle}$	1175205126.57
GW170403	IAS-1	0.33	$0.03^{\clubsuit}; 0.56^{\bigstar}; 0.27^{\blacktriangle}$	1175295987.22
GW170425	IAS-1	0.10	$0.21^{\clubsuit}; 0.77^{\bigstar}; 0.74^{\blacktriangle}$	1177134830.18
GW170608	GWTC-1	0.95	$0.92^{\heartsuit}; 1.00^{\spadesuit}; 1.0^{\blacktriangle}$	1180922492.50
GW170727	IAS-1	0.92	$0.99^{\clubsuit}; 0.98^{\bigstar}; 0.66^{\blacktriangle}$	1185152686.02
GW170729	GWTC-1	0.96	$0.98^{\heartsuit}; 0.52^{\spadesuit}; 1.0^{\blacktriangle}$	1185389805.30
GW170809	GWTC-1	0.98	$0.99^{\heartsuit}; 1.00^{\spadesuit}; 1.0^{\blacktriangle}$	1186302517.75
GW170814	GWTC-1	1.00	$1.00^{\heartsuit}; 1.00^{\spadesuit}; 1.0^{\blacktriangle}$	1186741859.53
GW170817A	IAS-2	0.83	$0.86^{\bigstar}; 0.36^{\diamondsuit}; 0.02^{\blacktriangle}$	1186974182.72

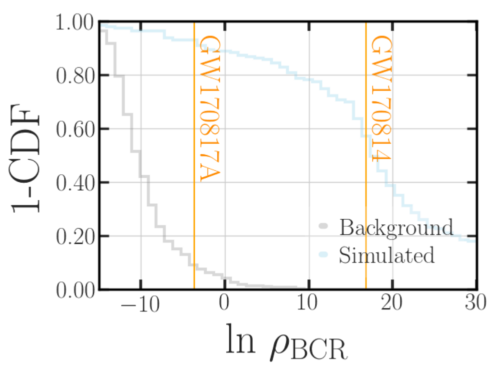

FIG. 2. Histograms represent the survival function (1-CDF) from our high-mass selection of background triggers (grey) and simulated signals (blue) triggers obtained from `PyCBC`'s search of data from August 13 - 21, 2017. Vertical lines mark the ln $\rho_{\rm BCR}$ of IAS's GW170817 and GWTC-1's GW170814.

$p_{\rm astro}$ of other pipelines for comparison. Although the various pipeline $p_{\rm astro}$ are not mathematically equivalent, by comparing pipeline $p_{\rm astro}$ values for a given candidate, we can compare how significant each pipeline deems various candidates. The α and β values utilised for each time-frame are reported in Appendix C.

A. GWTC-1 Events

All the confirmed gravitational-wave events from binary black hole mergers reported in GWTC-1 and within our prior space, (specifically GW170104, GW170608, GW170729, GW170809 and GW170814), have $p_{\rm astro}^{\rm BCR}$ greater than 0.9, indicating a high probability of the presence of an astrophysical signal.

In addition to the above confirmed gravitational-wave events from GWTC-1, we have also analysed several candidate events from GWTC-1, most of which have low $p_{\rm astro}^{\rm BCR}$. For example, consider the candidate event 170412, assigned a $p_{\rm astro}$ of 0.06 by GstLAL and has a $p_{\rm astro}^{\rm BCR}$ of 0.01. This candidate was reported to be excess power caused due noise appearing non-stationary between 60-200 Hz [2]. This candidate acts as an example of how $p_{\rm astro}^{\rm BCR}$ may be utilised to eliminate candidates originating from terrestrial noise sources.

B. IAS Events

Our analysis of the high-mass IAS events and candidates in O2 has resulted in three events with disfavourable $p_{\rm astro}^{\rm BCR} < 0.5$ (GWC170402, GW170403, GW170425), and four events and one candidate with $p_{\rm astro}^{\rm BCR} \geq 0.5$ (GW170121, 170302, GW170304, GW170727, GW170817A).

GWC170402, detected by Zackay *et al.* [8], is reported to have a signal that is not described well by waveforms for circular binaries with aligned spins. Hence, we might have received a low $p_{\rm astro}^{\rm BCR}$ due to our usage of IMRPhenomPv2, a waveform that does not account for eccentricity. Additionally, the search conducted by Zackay *et al.* [8] was a single-detector search. Our ranking statistic relies on the signal to appear coherent, even if just faintly coherent, amongst the various detectors to have a high $p_{\rm astro}^{\rm BCR}$. The lack of coherence and the non-eccentric waveform may be the leading factors for a low $p_{\rm astro}$. GW170403 and GW170425 which have $p_{\rm astro}^{\rm BCR} < 0.35$ also have low $p_{\rm astro}$ reported by Nitz *et al.* [44], suggesting that these events may have been false alarms.

From the candidates with $p_{\rm astro}^{\rm BCR} > 0.5$, GW170727 and 170302 are of particular interest, with $p_{\rm astro}^{\rm BCR}$ of 0.92 and 0.63. GW170727 was emitted from a black hole binary system with a source frame total mass $\approx 70\ {\rm M}_{\odot}$. In addition to the high $p_{\rm astro}^{\rm BCR}$ reported by our study, Venumadhav *et al.* [7] and Nitz *et al.* [44] have also reported high $p_{\rm astro}$ values of 0.98 and 0.99, making it a viable gravitational-wave event candidate. Similarly, the sub-marginal-candidate 170302 reported by [7] with a $p_{\rm astro}$ of 0.45 appears to have a higher significance from our analysis, resulting in a $p_{\rm astro}^{\rm BCR}$ of 0.63.

C. New Candidate Events

Although no clear detections are made with the $\rho_{\rm BCR}$, a marginal-candidate 170222 has been discovered with a $p_{\rm astro}^{\rm BCR} \sim 0.5$. This candidate has its similar masses when compared to those of GWTC-1. The remaining coherent trigger candidates all have $p_{\rm astro}^{\rm BCR} \ll 0.5$ making them unlikely to originate from astrophysical sources. [AV: wjat more can I say about this?]

VI. CONCLUSION

In this paper, we demonstrate that the Bayesian Coherence Ratio [14] can be used as a ranking statistic to provide a better measure of significance for gravitational-wave candidates by re-analysing the significance of high-mass binary black hole triggers from O2. This method takes a step towards building a unified Bayesian framework that provides a search-pipeline agnostic measure of significance, utilising the same level of physical information incorporated during parameter estimation.

We focused our analysis on the high-mass regime as this region of the parameter space is plagued with a high number of short duration terrestrial artefacts that can mimic signals. In addition to the high-mass triggers, we also analyse the high-mass binary black hole events in O2 reported by LIGO [2] and IAS [7, 8]. Using $p_{\rm astro}^{\rm BCR}$, we find that the analysed GWTC-1 events have high probabilities of originating from an astrophysical source. We also find that some of the GWTC-1 marginal triggers that have corroborated terrestrial sources (for example candidate 170412) have low $p_{\rm astro}^{\rm BCR}$, indicating this method's ability to discriminate between terrestrial artefacts and astrophysical signals. Our analysis on the IAS events has demonstrated that GW17072 is highly likely to originate from an astrophysical source, while GW17040 is not. Finally, we did not identify any new gravitational-wave events, but we did find some a marginal candidate 170222.

Although our analysis targets high-mass triggers, this method can be extended to include the entire body range of LIGO-detectable gravitational-wave sources. Additionally, to further improve the method's infrastructure, we can use more robust gravitational-wave templates (such as templates that incorporate higher-order modes), and sophisticated glitch models. Future analysis can also incorporate data from all available detectors in a network to increase the sensitivity of $p_{\rm astro}^{\rm BCR}$. The BCR can discern better whether a candidate is a coherent astrophysical candidate or an incoherent glitch with data from more detectors.

[AV: as the core of this method is PE, improvements in PE can be adapted into this method]

ACKNOWLEDGMENTS

This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (https://www.gw-openscience.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by Polish and Hungarian institutes.

Thank Stuart Anderson, CIT

We greatly appreciate the contributions of all these computing allocations. All pe performed for this study including test runs and failed simulations used about 2.5M core-hours which under the assumption of 30W per core-hour and a CO2 intensity of electricity of 600 kg CO2 per MWh amounts to a carbon footprint of 45t of CO2.

Appendix A: Bayesian Evidence Evaluation

1. Noise Model

[RS: I suggest moving this subsection to an appendix] We assume that each detector's noise is Gaussian and stationary over the period being analysed [40]. In practice, we assume that the noise has a mean of zero that the noise variance σ^2 is proportional to the noise power spectral density (PSD) $P(f)$ of the data. Using the $P(f)$, for each data segment d_i in each of the i detectors in a network of D detectors, we can write

$$Z_i^N = \mathcal{N}(d_i) = \frac{1}{2\pi P(f)_i} \exp\left(-\frac{1}{2}\frac{d_i}{P(f)_i}\right) , \quad \text{(A1)}$$

where $\mathcal{N}(d_i)$ is a normal distribution with $\mu = 0$ and $\sigma^2 \sim P(f)$.

2. Coherent Signal Model

We model coherent signal using a binary black hole waveform template $\mu(\vec{\theta})$, where the vector $\vec{\theta}$ contains a point in the 15 dimensional space describing precessing binary-black hole mergers. For the signal to be coherent, $\vec{\theta}$ must be consistent in each 4-second data segment d_i for a network of D detectors, Hence, the coherent signal evidence is calculated as

$$Z^S = \int_{\vec{\theta}} \prod_{i=1}^{D} \left[\mathcal{L}(d_i|\mu(\vec{\theta}))\right] \pi(\vec{\theta}|\mathcal{H}_S)\, \mathrm{d}\vec{\theta} , \quad \text{(A2)}$$

where $\pi(\vec{\theta}|\mathcal{H}_S)$ is the prior for the parameters in the coherent signal hypothesis, and $\mathcal{L}(d_i|\mu(\vec{\theta})))$ is the likelihood for the coherent signal hypothesis that depends on the gravitational-wave template $\mu(\vec{\theta})$ and its parameters $\vec{\theta}$.

3. Incoherent Glitch Model

Finally, as glitches are challenging to model and poorly understood, we follow Veitch and Vecchio [13] and utilise a surrogate model for glitches: the glitches are modelled using gravitational-wave templates $\mu(\vec{\theta})$ with uncorrelated parameters amongst the different detectors such that $\vec{\theta}_i \neq \vec{\theta}_j$ for two detectors i and j [13]. Modelling glitches with $\mu(\vec{\theta})$ captures the worst case scenario: when glitches are identical to gravitational-wave signals (excluding coherent signals). Thus, we can write Z_i^G as

$$Z_i^G = \int_{\vec{\theta}} \mathcal{L}(d_i|\mu(\vec{\theta}))\ \pi(\vec{\theta}|\mathcal{H}_G)\ \mathrm{d}\vec{\theta} , \quad \text{(A3)}$$

where $\pi(\theta|\mathcal{H}_G)$ is the prior for the parameters in the incoherent glitch hypothesis.

Appendix B: Tuning the prior-odds

After calculating the ρ_{BCR} for a set of background triggers and simulated triggers from as stretch of detector-data (a data chunk), we can compute probability distributions for the background and simulated triggers, $p_b(\rho_{\mathrm{BCR}})$ and $p_s(\rho_{\mathrm{BCR}})$. We expect the background trigger and simulated signal ρ_{BCR} values to favour the incoherent glitch and the coherent signal hypothesis, respectively. Ideally, these distributions representing two unique populations should be distinctly separate and have no overlap in their ρ_{BCR} values. The prior odds parameters α and β from Eq. 4 help separate the two distributions. Altering α translates the ρ_{BCR} probability distributions while adjusting β spreads the distributions. Although Bayesian hyper-parameter estimation can determine the optimal values for α and β, an easier approach is to adjust the parameters for each data chunk's ρ_{BCR} distribution. In this study, we tune α and β to maximally separate the ρ_{BCR} distributions for the background and simulated triggers.

To calculate the separation between $p_b(\rho_{\mathrm{BCR}})$ and $p_s(\rho_{\mathrm{BCR}})$, we use the Kullback–Leibler divergence (KL divergence) D_{KL}, given by

$$D_{KL}(p_b|p_s) = \sum_{x\in\rho_{\mathrm{BCR}}} p_b(x) \log\left(\frac{p_b(x)}{p_a(x)}\right) . \quad \text{(B1)}$$

The $D_{KL} = 0$ when the distributions are identical and increases as the asymmetry between the distributions increases.

We limit our search for the maximum KL-divergence in the α and β ranges of $[10^{-10}, 10^{0}]$ as values outside this range are nonphysical. We set our values for α and β to those which provide the highest KL-divergence and calculate the ρ_{BCR} for candidate events present in this data chunk. Note that we conduct the analysis in data chunks of a few days rather than an entire data set of a few months as the background may be different at different points of the entire data set.

Appendix C: Tuned prior odds

O2 lasted several months over which the detector's sensitivity varied. Hence, a part of our analysis entailed tuning the prior odds for obtaining a signal and a glitch, α and β, as described in Section III. Table IV presents the signal and glitch prior odds utilised for each time-frame of O2 data. [AV: note: havent tuned all chunks yet – and tbh the 'tuning' could be better, atm brute force searching a 2d log-spce grid]

Tuning the prior odds can dramatically affect the $p_{\mathrm{astro}}^{\mathrm{BCR}}$. For example, consider Table V, which reports tuned $p_{\mathrm{astro}}^{\mathrm{BCR}}$ and un-tuned $p_{\mathrm{astro}}^{\mathrm{BCR}\prime}$ (where $\alpha = 1$ and $\beta = 1$) for various high-mass events and candidates. By tuning the prior odds, the $p_{\mathrm{astro}}^{\mathrm{BCR}}$ for some IAS events (for example, GW170403 and GW170817A) can change by more than 0.5, resulting in the promotion/demotion of a candidate's significance.

TABLE IV. The prior odds used for each time-frame of data from O2. Each time frame commences at the start date and concludes at the following time-frame's start date.

Start Date	α	β
2016-11-15	-	-
2016-11-30	-	-
2016-12-23	1.00E+00	6.25E-01
2017-01-22	-	-
2017-02-03	1.00E-10	2.44E-01
2017-02-12	1.76E-08	5.96E-02
2017-02-20	6.55E-10	2.22E-03
2017-02-28	1.00E-10	5.96E-02
2017-03-10	2.56E-10	3.91E-01
2017-03-18	1.60E-10	1.00E+00
2017-03-27	1.10E-08	5.96E-02
2017-04-04	3.73E-02	2.33E-02
2017-04-14	1.05E-09	2.44E-01
2017-04-23	2.68E-09	1.46E-02
2017-05-08	1.00E+00	2.44E-01
2017-06-18	6.55E-10	3.39E-04
2017-06-30	2.02E-05	5.69E-03
2017-07-15	1.05E-09	9.54E-02
2017-07-27	-	-
2017-08-05	2.12E-04	3.73E-02
2017-08-13	2.68E-09	8.69E-04
2017-08-21	-	-

[1] Event Horizon Telescope Collaboration, K. Akiyama, A. Alberdi, W. Alef, *et al.*, First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole, ApJ **875**, L4 (2019), arXiv:1906.11241 [astro-ph.GA].

[2] B. P. Abbott, R. Abbott, T. D. Abbott, *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs, Phys. Rev. X **9**, 031040 (2019).

[3] B. P. Abbott, R. Abbott, T. D. Abbott, *et al.*, GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run, arXiv e-prints , arXiv:2010.14527 (2020), arXiv:2010.14527 [gr-qc].

[4] GraceDB, Gravitational-Wave Candidate Event Database (2020).

[5] A. Nitz, I. Harry, D. Brown, C. M. Biwer, J. Willis, T. D. Canton, C. Capano, L. Pekowsky, T. Dent, A. R. Williamson, G. S. Davies, S. De, M. Cabero, B. Machenschalk, P. Kumar, S. Reyes, D. Macleod, F. Pannarale, dfinstad, T. Massinger, M. Tápai, L. Singer, S. Khan, S. Fairhurst, S. Kumar, A. Nielsen, shasvath, I. Dorrington, A. Lenon, and H. Gabbard, gwastro/pycbc: PyCBC Release 1.16.4 (2020).

[6] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, New search pipeline for compact binary mergers: Results for binary black holes in the first observing run of Advanced LIGO, Physical Review D **100**, 023011 (2019).

[7] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, New Binary Black Hole Mergers in the Second Observing Run of Advanced LIGO and Advanced Virgo, arXiv e-prints , arXiv:1904.07214 (2019), arXiv:1904.07214 [astro-ph.HE].

[8] B. Zackay, L. Dai, T. Venumadhav, J. Roulet, and M. Zaldarriaga, Detecting Gravitational Waves With Disparate Detector Responses: Two New Binary Black Hole Mergers, arXiv e-prints , arXiv:1910.09528 (2019),

TABLE V. The BCR p-astro after tuning the prior odds, $p_{\rm astro}^{\rm BCR}$, and without tuning the prior odds, $p_{\rm astro}^{\rm BCR\prime}$.

Event	Catalogue	$p_{\rm astro}^{\rm BCR}$	$p_{\rm astro}^{\rm BCR\prime}$
161202	-	0.09	0.41
GW170104	GWTC-1	0.94	0.93
GW170121	IAS-1	0.76	0.72
170206	-	0.11	0.51
170222	-	0.49	0.48
170302	IAS-1	0.63	0.54
GW170304	IAS-1	0.83	0.81
GWC170402	IAS-2	0.38	0.01
GW170403	IAS-1	0.33	0.89
GW170425	IAS-1	0.10	0.22
GW170608	GWTC-1	0.95	0.95
GW170727	IAS-1	0.92	0.96
GW170729	GWTC-1	0.96	0.94
GW170809	GWTC-1	0.98	0.99
GW170814	GWTC-1	1.00	1.00
GW170817A	IAS-2	0.83	0.36

arXiv:1910.09528 [astro-ph.HE].

[9] B. P. Abbott, R. Abbott, T. D. Abbott, *et al.*, GW190521: A Binary Black Hole Merger with a Total Mass of 150 M_{odot}, Phys. Rev. Lett. **125**, 101102 (2020), arXiv:2009.01075 [gr-qc].

[10] A. H. Nitz, Distinguishing short duration noise transients in LIGO data to improve the PyCBC search for gravitational waves from high mass binary black hole mergers, Classical and Quantum Gravity **35**, 035016 (2018), arXiv:1709.08974 [gr-qc].

[11] J. Powell, Parameter estimation and model selection of gravitational wave signals contaminated by transient detector noise glitches, Classical and Quantum Gravity **35**, 155017 (2018), arXiv:1803.11346 [astro-ph.IM].

[12] M. Cabero, A. Lundgren, A. H. Nitz, T. Dent, D. Barker, E. Goetz, J. S. Kissel, L. K. Nuttall, P. Schale, R. Schofield, and D. Davis, Blip glitches in Advanced LIGO data, Classical and Quantum Gravity **36**, 155010 (2019), arXiv:1901.05093 [physics.ins-det].

[13] J. Veitch and A. Vecchio, Bayesian coherent analysis of in-spiral gravitational wave signals with a detector network, Phys. Rev. D **81**, 062003 (2010), arXiv:0911.3820 [astro-ph.CO].

[14] M. Isi, R. Smith, S. Vitale, T. J. Massinger, J. Kanner, and A. Vajpeyi, Enhancing confidence in the detection of gravitational waves from compact binaries using signal coherence, Phys. Rev. D **98**, 042007 (2018), arXiv:1803.09783 [gr-qc].

[15] G. Ashton, E. Thrane, and R. J. E. Smith, Gravitational wave detection without boot straps: A Bayesian approach, Phys. Rev. D **100**, 123018 (2019), arXiv:1909.11872 [gr-qc].

[16] G. Ashton and E. Thrane, The astrophysical odds of GW151216, MNRAS 10.1093/mnras/staa2332 (2020), arXiv:2006.05039 [astro-ph.HE].

[17] G. Pratten and A. Vecchio, Assessing gravitational-wave binary black hole candidates with Bayesian odds, arXiv e-prints , arXiv:2008.00509 (2020), arXiv:2008.00509 [gr-qc].

[18] W. M. Farr, J. R. Gair, I. Mandel, and C. Cutler, Counting and confusion: Bayesian rate estimation with multiple populations, Phys. Rev. D **91**, 023005 (2015), arXiv:1302.5341 [astro-ph.IM].

[19] S. J. Kapadia, S. Caudill, J. D. E. Creighton, W. M. Farr, G. Mendell, A. Weinstein, K. Cannon, H. Fong, P. Godwin, R. K. L. Lo, R. Magee, D. Meacher, C. Messick, S. R. Mohite, D. Mukherjee, and S. Sachdev, A self-consistent method to estimate the rate of compact binary coalescences with a Poisson mixture model, Classical and Quantum Gravity **37**, 045007 (2020), arXiv:1903.06881 [astro-ph.HE].

[20] S. M. Gaebel, J. Veitch, T. Dent, and W. M. Farr, Digging the population of compact binary mergers out of the noise, MNRAS **484**, 4008 (2019), arXiv:1809.03815 [astro-ph.IM].

[21] G. S. Davies, T. Dent, M. Tápai, I. Harry, C. McIsaac, and A. H. Nitz, Extending the PyCBC search for gravitational waves from compact binary mergers to a global network, Phys. Rev. D **102**, 022004 (2020), arXiv:2002.08291 [astro-ph.HE].

[22] B. Allen, χ^2 time-frequency discriminator for gravitational wave detection, Phys. Rev. D **71**, 062001 (2005), arXiv:gr-qc/0405045 [gr-qc].

[23] The LIGO Scientific Collaboration, the Virgo Collaboration, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, and et al., Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo, arXiv e-prints , arXiv:1912.11716 (2019), arXiv:1912.11716 [gr-qc].

[24] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, FINDCHIRP: An algorithm for detection of gravitational waves from inspiraling compact binaries, Phys. Rev. D **85**, 122006 (2012), arXiv:gr-qc/0509116 [gr-qc].

[25] A. H. Nitz, T. Dent, T. Dal Canton, S. Fairhurst, and D. A. Brown, Detecting Binary Compact-object Mergers with Gravitational Waves: Understanding and Improving the Sensitivity of the PyCBC Search, ApJ **849**, 118 (2017), arXiv:1705.01513 [gr-qc].

[26] T. Dal Canton, A. H. Nitz, A. P. Lundgren, A. B. Nielsen, D. A. Brown, T. Dent, I. W. Harry, B. Krishnan, A. J. Miller, K. Wette, K. Wiesner, and J. L. Willis, Implementing a search for aligned-spin neutron star-black hole systems with advanced ground based gravitational wave detectors, Phys. Rev. D **90**, 082004 (2014), arXiv:1405.6731 [gr-qc].

[27] S. A. Usman, A. H. Nitz, I. W. Harry, C. M. Biwer, D. A. Brown, M. Cabero, C. D. Capano, T. Dal Canton, T. Dent, S. Fairhurst, M. S. Kehl, D. Keppel, B. Krishnan, A. Lenon, A. Lundgren, A. B. Nielsen, L. P. Pekowsky, H. P. Pfeiffer, P. R. Saulson, M. West, and J. L. Willis, The PyCBC search for gravitational waves from compact binary coalescence, Classical and Quantum Gravity **33**, 215004 (2016), arXiv:1508.02357 [gr-qc].

[28] A. H. Nitz, T. Dal Canton, D. Davis, and S. Reyes, Rapid detection of gravitational waves from compact binary mergers with PyCBC Live, Phys. Rev. D **98**, 024050

(2018), arXiv:1805.11174 [gr-qc].

[29] G. Ashton, M. Hübner, P. Lasky, and C. Talbot, Bilby: A User-Friendly Bayesian Inference Library (2019).

[30] G. Ashton, I. Romero-Shaw, C. Talbot, C. Hoy, and S. Galaudage, bilby pipe: 1.0.1 (2020).

[31] J. S. Speagle, DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences, MNRAS **493**, 3132 (2020), arXiv:1904.02180 [astro-ph.IM].

[32] J. Skilling, Nested Sampling, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, American Institute of Physics Conference Series, Vol. 735, edited by R. Fischer, R. Preuss, and U. V. Toussaint (2004) pp. 395–405.

[33] J. Skilling, Nested sampling for general Bayesian computation, Bayesian Analysis **1**, 833 (2006).

[34] G. Ashton, M. Hübner, P. D. Lasky, C. Talbot, K. Ackley, S. Biscoveanu, Q. Chu, A. Divakarla, P. J. Easter, B. Goncharov, F. Hernandez Vivanco, J. Harms, M. E. Lower, G. D. Meadors, D. Melchor, E. Payne, M. D. Pitkin, J. Powell, N. Sarin, R. J. E. Smith, and E. Thrane, BILBY: A User-friendly Bayesian Inference Library for Gravitational-wave Astronomy, ApJS **241**, 27 (2019), arXiv:1811.02042 [astro-ph.IM].

[35] R. J. E. Smith, G. Ashton, A. Vajpeyi, and C. Talbot, Massively parallel Bayesian inference for transient gravitational-wave astronomy, MNRAS **498**, 4492 (2020), arXiv:1909.11873 [gr-qc].

[36] E. Thrane and C. Talbot, An introduction to Bayesian inference in gravitational-wave astronomy: Parameter estimation, model selection, and hierarchical models, PASA **36**, e010 (2019), arXiv:1809.02293 [astro-ph.IM].

[37] I. M. Romero-Shaw, C. Talbot, S. Biscoveanu, V. D'Emilio, G. Ashton, *et al.*, Bayesian inference for compact binary coalescences with BILBY: validation and application to the first LIGO-Virgo gravitational-wave transient catalogue, MNRAS **499**, 3295 (2020), arXiv:2006.00714 [astro-ph.IM].

[38] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era, Physical Review D **93**, 044007 (2016).

[39] S. Ossokine, A. Buonanno, S. Marsat, R. Cotesta, S. Babak, T. Dietrich, R. Haas, I. Hinder, H. P. Pfeiffer, M. Pürrer, C. J. Woodford, M. Boyle, L. E. Kidder, M. A. Scheel, and B. Szilágyi, Multipolar effective-one-body waveforms for precessing binary black holes: Construction and validation, Phys. Rev. D **102**, 044055 (2020), arXiv:2004.09442 [gr-qc].

[40] B. P. Abbott, R. Abbott, T. D. Abbott, *et al.*, A guide to LIGO-Virgo detector noise and extraction of transient gravitational-wave signals, arXiv e-prints , arXiv:1908.11170 (2019), arXiv:1908.11170 [gr-qc].

[41] C. Talbot and E. Thrane, Gravitational-wave astronomy with an uncertain noise power spectral density, arXiv e-prints , arXiv:2006.05292 (2020), arXiv:2006.05292 [astro-ph.IM].

[42] K. Chatziioannou, C.-J. Haster, T. B. Littenberg, W. M. Farr, S. Ghonge, M. Millhouse, J. A. Clark, and N. Cornish, Noise spectral estimation methods and their impact on gravitational wave measurement of compact binary mergers, Phys. Rev. D **100**, 104004 (2019).

[43] D. Macleod, A. L. Urban, S. Coughlin, T. Massinger, M. Pitkin, paulaltin, J. Areeda, E. Quintero, T. G. Badger, L. Singer, and K. Leinweber, gwpy/gwpy: 1.0.1 (2020).

[44] A. H. Nitz, T. Dent, G. S. Davies, S. Kumar, C. D. Capano, I. Harry, S. Mozzon, L. Nuttall, A. Lundgren, and M. Tápai, 2-OGC: Open Gravitational-wave Catalog of Binary Mergers from Analysis of Public Advanced LIGO and Virgo Data, ApJ **891**, 123 (2020), arXiv:1910.05331 [astro-ph.HE].

[45] A. H. Nitz, T. Dent, G. S. Davies, and I. Harry, A Search for Gravitational Waves from Binary Mergers with a Single Observatory, ApJ **897**, 169 (2020), arXiv:2004.10015 [astro-ph.HE].