A Bayesian search to find high-mass black holes in LIGO data

Author list TBD
(Dated: January 5, 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 500 $M_\odot$, only 1 of their 50 detections have a total mass > 100 $M_\odot$. A possible explanation for the absence of intermediate mass events may be due to 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. Here we demonstrate that a search method utilising Bayesian inference could be a more sensitive tool for detecting high-mass binary black hole mergers (systems with a total mass > 55 $M_\odot$) as compared to traditional match-filtering. We have applied this technique on the high-mass triggers during LIGO's second observing run to investigate the possibility of discovering new gravitational-wave signals from high mass black hole binaries, and to re-calculate the significance of high-mass candidate events. Although our search does not discover new candidate events, it does alter the significance of candidates identified by various search pipelines.

I. 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 November 2020, the LIGO Scientific Collaboration has confirmed $\sim$ 50 binary black hole systems and listed numerous candidate events [2–7]. These various discoveries have firmly established the existence of stellar-mass black holes, supermassive black holes and binary black hole systems. Until recently, there was no definitive evidence for intermediate-mass black holes, the black holes that lie in between stellar-mass and supermassive black hole systems with masses between $10^2 - 10^6$ $M_\odot$. This changed with the detection of GW190521, a unique gravitational wave event that lead to the formation of a black hole with a mass 142 $M_\odot$, the first direct discovery of an intermediate mass black hole. Although this is the first gravitational wave that has l;ead to the discovery of a black hole with a mass of > 100 $M_\odot$, ground based gravitational wave detectors are sensitive to gravitational waves from even more massive systems, up to a systems with a total mass of 400 $M_\odot$. Gravitation waves from systems with masses > 100 $M_\odot$ systems should occur at a rate of $0 - 10\mathrm{yr}^{-1}$ [8–11].

However, even after conducting a targeted matched-filter based search for gravitational waves from intermediate-mass black holes the largest total mass detected so far is approximately 80 $M_\odot$ [11–13]. 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 [14–20]. 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 is by utilising an astrophysical Bayesian odds [21–25]. A true Bayesian odds calculated without using bootstrap techniques can provide events with a more accurate significance that is agnostic to a specific search strategy [23–25]. Additionally, a Bayesian odds can include more information than is included in current matched filter searches such as if the gravitational wave event's signal is coherent amongst the network of detectors, if the binary system that created the gravitational waves was precessing, and if the gravitational wave signal contains higher-order modes. It is because Bayesian methods can incorporate all this physical information about a gravitational wave signal that the LIGO Scientific Collaboration uses these methods to determine the source parameters of gravitational wave events [13, 26]. This paper demonstrates that the power of Bayesian methods used in parameter estimation can also successfully be used to discriminate between coherent gravitational-wave signals, incoherent glitches, and Gaussian noise in the form of a Bayesian search.

In this paper we utilise a Bayesian method, called the Bayesian Coherence Ratio ρ_{B} [22], to search for the significant gravitational wave signals from high-mass (systems with total masses in the range of $55-500$ $M_\odot$) compact binary coalescences in the detector data recorded during aLIGO's second observing run (O2). We find that (a) our search does not identify any unreported stellar mass or intermediate mass black holes; (b) high-mass events reported in the GWTC-1, including GW170729 (an event with disputed p_{astro} amongst various search pipelines) have high significance; [AV: and that (c) high-mass events detected from the IAS groups have low significance.]

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

II. METHOD

The gravitational wave community uses Bayesian inference to perform parameter estimation and model selection. In this work, we utilise Bayesian inference to calculate the significance of high-mass candidate events in O2 by using the $\rho_{\rm B}$ as a ranking statistic, taking a step forward to building a unified Bayesian framework to search for candidates and estimate their parameters.

Although a dedicated Bayesian search for gravitational waves, as presented by [23], does not require noise estimation using empirical methods, this Bayesian significance ranking technique utilises *time-slides*. To perform time-slides, the data from independent observatories are time-shifted by amounts greater than the light-travel time between the two detectors. Each unique time-slide amount creates an artificial signal-free "background" data set. When gravitational-wave search pipelines scan these background data sets for gravitational wave events, they can find candidate triggers, even though the background data set should not contain coherent astrophysical signals. These candidate triggers in background data are labelled background triggers, while coherent triggers obtained in non-time-slid data are labelled candidate triggers.

Calculating the $\rho_{\rm B}$ ranking statistic for each background trigger builds a background $\rho_{\rm B}$ distribution. With a background distribution, it is possible to assign a statistical significance of how likely a candidate trigger, detected by the search pipelines over non-time slid data, is due to a gravitational wave signal.

This section discusses (a) the method to retrieve triggers, and (b) the $\rho_{\rm B}$ and how it is utilised as a ranking statistic to calculate the significance of candidate triggers.

A. Triggers for Analysis

The LIGO Scientific collaboration operates several search pipelines that scan for gravitational waves from compact binary mergers such as `GstLAL`, `MBTA`, `SPIIR` and `PyCBC` [13].

The output of `PyCBC`'s search is a list of candidate trigger times and their corresponding `PyCBC` ranking statistic $\rho_{\rm PC}$. The $\rho_{\rm PC}$ statistic is akin to the matched-filter signal-to-noise ratio ρ. However, unlike ρ, $\rho_{\rm PC}$ includes some information on the candidate signal's intrinsic and extrinsic properties and other information that feeds into determining if the signal can have astrophysical origins [27]. The additional physical information incorporated in $\rho_{\rm PC}$ makes it a more accurate measure of significance than the standard ρ.

Whenever a local maximum of $\rho_{\rm PC} > \rho_{\rm T}$, where $\rho_{\rm T}$ is some threshold value, the search pipeline produces a single-detector trigger associated with the detector and time where the apparent signal in the data has its merger.

For `PyCBC` to consider a trigger to be a **candidate trigger**, a trigger from astrophysical origins, the trigger must be observed between detectors with the same template and a time of arrival difference less than the gravitational-wave travel time [27]. To test its search, `PyCBC` also conducts searches for **simulated triggers**, artificial triggers manufactured by injecting signals into the detector data. These simulated signal studies provide `PyCBC` with metrics on its search's sensitivity. Finally, to quantify the statistical significance of candidate events, `PyCBC` artificially constructs a **background trigger** set 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 [27]. Note that the time-slides to generate the background triggers are greater than the gravitational-wave travel time between detectors to ensure that the background triggers are not of astrophysical origins.

Our work demonstrates that the $\rho_{\rm B}$ can be used in the same way as $\rho_{\rm PC}$ to measure candidate triggers' statistical significance. The $\rho_{\rm B}$ can be a powerful ranking statistic as the $\rho_{\rm B}$ 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.

Before we discuss how we use the $\rho_{\rm B}$ as a measure of significance, we introduce the method to calculate the $\rho_{\rm B}$ in the following section.

B. 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

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

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,

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

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

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

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 $\mathcal{O}^A_B >> 1$, then this odds ratio indicates that the $\mathcal{H}_A$ describes the data much better than $\mathcal{H}_B$.

The $\rho_{\rm B}$ is a Bayesian odds ratio like the above, of a coherent signal hypotheses $\mathcal{H}_S$ and an incoherent instrumental feature hypothesis $\mathcal{H}_I$ for a network of D detectors. $\mathcal{H}_\mathcal{I}$ states that each detector i has either pure Gaussian noise $\mathcal{H}_N$ or a glitch $\mathcal{H}_G$. Following Isi et al. [22], the $\rho_{\rm B}$ is given by

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

where Z^S, Z_i^G and Z_i^N are the Bayesian evidences (marginalised likelihoods) for $\mathcal{H}_S$, $\mathcal{H}_N$, and $\mathcal{H}_G$. α and β, are the prior odds for obtaining a signal $\alpha = P(\mathcal{H}_S)/P(\mathcal{H}_I)$ and the prior odds for obtaining a glitch $\beta = P(\mathcal{H}_G)/P(\mathcal{H}_I)$. As the rate of signal and glitches are unknown, these priors α and β are tuned to maximise the $\rho_{\rm B}$ distributions for background data (signal-free data) and simulated signals [22].

C. Bayesian Evidence Evaluation

1. Noise Model

We assume that each detector's noise is Gaussian and stationary over the period being analysed [28]. 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

$$Z_i^N = \mathcal{N}(d_i) = \frac{1}{2\pi P(f)_i} \exp\left(-\frac{1}{2}\frac{d_i}{P(f)_i}\right) \,, \tag{5}$$

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 s data segment d_i for a network of D detectors, Hence, the coherent signal evidence is calculated as

$$Z^S = \int_{\vec{\theta}} \prod_{i=1}^{D} \left[\mathcal{L}(d_i|\mu(\vec{\theta}))\right] \pi(\vec{\theta}|\mathcal{H}_S)\, {\rm d}\vec{\theta} \,, \tag{6}$$

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 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 [21]. Modelling glitches with $\mu(\vec{\theta})$ captures the worst case scenario: when glitches appear similar to gravitational wave signals. Thus, we can write Z_i^G as

$$Z_i^G = \int_{\vec{\theta}} \mathcal{L}(d_i|\mu(\vec{\theta}))) \; \pi(\vec{\theta}|\mathcal{H}_G) \; {\rm d}\vec{\theta} \,, \tag{7}$$

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

D. Tuning the BCR

After calculating the $\rho_{\rm B}$ 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, $p_b(\rho_{\rm B})$ and $p_s(\rho_{\rm B})$. We expect the background trigger and simulated signal $\rho_{\rm B}$ values to favour the incoherent glitch and the coherent signal hypothesis, respectively. Ideally these distributions that represent two unique populations should be distinctly separate and have no overlap in their $\rho_{\rm B}$ values. The prior odds parameters α and β from Eq. 4 help separate the two distributions. Altering α translates the $\rho_{\rm B}$ 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 $\rho_{\rm B}$ distribution. In this study, we tune α and β to maximally separate the $\rho_{\rm B}$ distributions for the background and simulated triggers.

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

$$D_{KL}(p_b|p_s) = \sum_{i=1}^{N} p_b({\rho_{\rm B}}^i) \cdot (\log p_b({\rho_{\rm B}}^i) - \log p_s({\rho_{\rm B}}^i) \,. \tag{8}$$

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 $[E-10, E0]$ 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 $\rho_{\rm B}$ 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.

E. Calculating the Significance of Candidate events

With the tuned values of α and β we can calculate the $\rho_{\rm B}$ for candidate events. As mentioned previously, irrespective of the $\rho_{\rm B}$'s Bayesian interpretation, we treat the $\rho_{\rm B}$ as a traditional detection statistic to obtain a frequentist estimate of the significance of candidate event measured against the background of signal-free data.

We expect the background trigger $\rho_{\rm B}$ values to favour the incoherent glitch hypothesis (the null hypothesis). Candidate event $\rho_{\rm B}$ values will either be statistically insignificant compared to the background triggers, implying the candidate is a glitch, or statistically significant to the background distribution, indicating the possible presence of an astrophysical signal. To quantify the level of significance we can calculate a p-value for the candidate events. Here, the p-value tells us how probable it is for the candidate event to be a glitch. Hence, we can use the p-value to calculate

$$p_{\rm astro} = 1 - \text{p-value} \ , \tag{9}$$

where $p_{\rm astro}$ is the probability that a signal is of astrophysical origin [29–31].

III. ANALYSIS

A. Acquisition of triggers

Advanced LIGO's second observing run O2 lasted 38 weeks [32]. The software package, PyCBC [33], processed the O2 data in 22 time-frames (approximately 2 weeks for one time-frames) and found several gravitational wave events and numerous gravitational wave candidates [34–39]. Some candidate events were vetoed to be glitches, while others due to their low significance. The data is divided into these time-frames because the detector's sensitivity does not stay constant throughout 8 month long observing period.

In addition to finding candidate events, PyCBC also manufactured several million background triggers for each time-frames, to quantify the significance of the candidate events when compared to the background for the respective time-frames. Finally, to test the search's sensitivity, PyCBC simulated and searched for thousands of artificial signals.

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.

TABLE I: Template Banks's parameters for templates with duration < 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

This figure also depicts the gravitational wave templates used during the search through this time-frame of data.

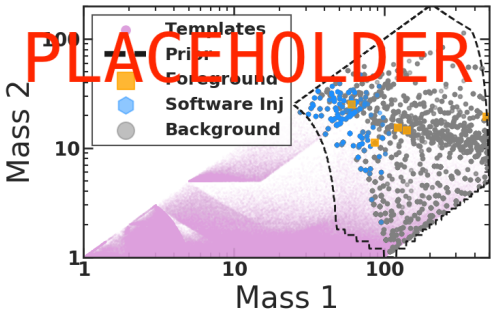

FIG. 1: The template bank used by PyCBC to search a section of O2 data from April 23 - May 8, 2017. Our search is constrained to the high-mass parameter space enclosed by the dashed line.

B. Calculating the BCR for triggers

To evaluate Z^S, Z_i^G and Z_i^N as shown in Eqs. 5-7 and calculate the $\rho_{\rm B}$ Eq. 4 for these triggers, we carry out Bayesian inference with Bilby [40], employing DYNESTY [41] as our nested sampler. Nested sampling, an algorithm introduced by [**?**], provides an estimate of the true Bayesian evidence and is often utilised for parameter estimation within the LIGO collaboration [40].

The most computationally intensive step during Bayesian inference is the evaluation of 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 [42]). 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 [42].

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 remaining priors are the same as the those used for in GWTC-1.

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 [43]. Although gravitational wave templates such as SEOBNRv4PHM [**?**] which incorporate more physics, such as information on higher-order modes, we still use IMRPhenomPv2 as it is inexpensive compared to waveforms fitted against numerical-relativity simulations.

To generate the PSD, we take 31 neighbouring, off-source, non-overlapping, 4-second segments of time-series data prior to the data segment d_i being analysed. A Tukey window with a roll off of 0.2-seconds 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 [28]. This method, like other PSD estimation methods, adds statistical uncertainties to the PSD [44]. To marginalise over the statistical uncertainty we use the median-likelihood presented by Talbot and Thrane [44] as a post-processing step and shift our Bayesian Evidence estimations closer to their true astrophysical values.

Finally, we neglect detector calibration uncertainty and acquire data from from the Gravitational Wave Open Science Center [32]. The data we use is the publicly accessible O2 strain data from the Hanford and Livingston detectors. To ensure the data is usable we verify that the analysis and PSD data are obtained when the detectors are in "Science Mode". The data requisition and quality checks are conducted using `GWPY` [**?**].

The run-time to calculate a single Bayesian evidence after using DYNESTY with $1,000$ live points and 100 walkers is usually between $0.5-12$ hours (where the run time often depends on the SNR of the data segment).

C. Assigning $p_{\rm astro}$ to candidate events

After the calculating the $\rho_{\rm B}$ for the entire set of high-mass background and simulated triggers, we calculate probability distributions $p_b(\rho_{\rm B})$ and $p_s(\rho_{\rm B})$ for each 2-week time-frame of O2 data. These distributions are used to obtain the "tuned" prior-odd α and β values that maximise $D_{KL}(p_b|p_s)$ for each time-frame of data.

Finally, using the tuned prior odds the $\rho_{\rm B}$ for the candidate events can be calculated. Figure 2 shows the $\rho_{\rm B}$ distributions for the background triggers, simulated triggers and candidate events. The bulk of the background and simulated trigger distributions are separate but have a slight overlap due to some of the simulated signal's being very faint. This suggests that the $\rho_{\rm B}$ can successfully distinguish signals from noise or glitches. The vertical lines in Figure 2 displays the $\rho_{\rm B}$ for gravitational wave candidate events. On comparing the candidate event $\rho_{\rm B}$ values with the background distribution, we can estimate $p_{\rm astro}$ values for the candidate events.

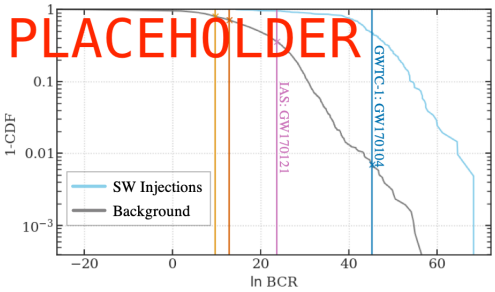

FIG. 2: Histograms represent the survival function (1-CDF) from our selection of $\sim$ 3,000 background triggers (grey) and 648 simulated signals (blue) triggers obtained from `PyCBC`'s search of data from April 23 - May 8, 2017. Vertical lines mark the ln BCRs of two glitches (orange and yellow), IAS's GW170121 (pink), and GWTC-1's GW170104 (dark blue).

IV. RESULTS

[AV: I should use Pratten et al's table + values here] In aLIGO's first two observing runs, eleven gravitational wave events were found in the data by the LIGO-Virgo scientific collaboration [13]. Since the public release of LIGO's first and second observing run's data, several groups have searched the data for gravitational waves independently of LIGO. One particular research group of interest is a research team at the Institute for Advanced Study (IAS). The group constructed searches to look for the LIGO-confirmed the gravitational wave events detected by the LIGO-Virgo collaboration, and in the process of doing so, claim to have discovered several others events [4–6]. Some of these events have total masses > 85 $M_\odot$, which is larger than the average total mass of the LIGO detections. Some of these IAS and LIGO events are displayed in Table II with their $p_{\rm astro}$ reported by various LIGO and IAS search pipelines.

From Table II, it is evident there is some uncertainty if these events can be considered real gravitational-wave events – are these events significantly different from the background or not? The various pipelines have different answers. Our $\rho_{\rm B}$ $p_{\rm astro}$ shows support that the LIGO-VIRGO events come from an astrophysical source, while the candidate events from the IAS group appear to have a lower probability of originating from an astrophysical source.

V. CONCLUSION

[AV: Although a true Bayesian search for gravitational waves, as presented by [23], does not require noise estimation using empirical methods, this Bayesian framework utilises 'time-slides']

TABLE II: $p_{\rm astro}$ from several detection pipelines for a subset of the O2 foreground triggers.

Event	Catalogue	PyCBC	GstLAL	cWB	IAS	BCR
GW170104	GWTC-1	1	1	1	0.99	0.99
GW170121	IAS-1	NA	NA	NA	0.99	0.65
GWC170402	IAS-2	NA	0.086	NA	0.68	0.33
GW170403	IAS-1	NA	NA	NA	0.56	0.33
IMBHC170423	IMBH-marginal	NA	0.95	1	NA	0.01
GW170425	IAS-1	NA	NA	NA	0.77	0.36
GW170502	Udall et al.	NA	NA	NA	NA	0.40
GW170729	GWTC-1	0.52	0.98	0.94	NA	0.98

[RS: lead with something specific to your work; what did you find/learn? how can it be extended? what does it imply for high mass BBH? How do the reported BCRs on interesting events compare to other bayesian measures of significance/odds? etc...It might be useful to take a look at some of the pycbc search papers for how to summarize results from a search]The detection of high mass black holes > 100 $M_\odot$ will shed light on the formation of globular clusters, supermassive black holes and thus galaxy formation [45, 46]. LIGO is theoretically sensitive to the merger of binary black holes with total masses up to 500 $M_\odot$ which are expected to occur at a rate of 0-10 yr^{-1} [8, 9]. However, even after Abbott et al. [11]'s targeted match-filter based search for gravitational waves from high-mass black holes the largest total mass detected so far is approximately 80 $M_\odot$ [13]. A possible explanation for the absence of high mass events may be due to their misclassification as short-duration instrumental noise transients [**?**]. High-mass mergers have very few in-band wave cycles, and hence can easily be mistaken for short-duration instrumental transients.

We have developing a targeted search for gravitational waves from high-mass black hole systems. This targeted search utilises Bayesian inference and provides a ranking statistic that contains a lot of physical information about the high-mass systems. We have applied this technique on all the high-mass triggers identified by `PyCBC` during LIGO's second observing run to investigate the possibility of discovering new gravitational-wave signals from high mass black hole binaries. Although we were unable to uncover new gravitational waves events, we were able to report high p_{astro} for events already detected by LIGO, and low p_{astro} for some events identified by external pipelines.

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.

[1] Event Horizon Telescope Collaboration, K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, J. Barrett, et al., ApJ **875**, L4 (2019), 1906.11241.

[2] B. P. Abbott, R. Abbott, T. D. Abbott, et al. (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. X **9**, 031040 (2019), URL `https://link.aps.org/doi/10.1103/PhysRevX.9.031040`.

[3] B. P. Abbott, R. Abbott, T. D. Abbott, et al., arXiv e-prints arXiv:2010.14527 (2020), 2010.14527.

[4] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, Physical Review D **100**, 023011 (2019).

[5] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, arXiv e-prints arXiv:1904.07214 (2019), 1904.07214.

[6] B. Zackay, L. Dai, T. Venumadhav, J. Roulet, and M. Zaldarriaga, arXiv e-prints arXiv:1910.09528 (2019), 1910.09528.

[7] A. H. Nitz, T. Dent, G. S. Davies, S. Kumar, C. D. Capano, I. Harry, S. Mozzon, L. Nuttall, A. Lundgren, and M. Tápai, ApJ **891**, 123 (2020), 1910.05331.

[8] J. M. Fregeau, S. L. Larson, M. C. Miller, R. O'Shaughnessy, and F. A. Rasio, The Astrophysical Journal Letters **646**, L135 (2006), URL `https://iopscience.iop.org/article/10.1086/507106/pdf`.

[9] I. Mandel, D. A. Brown, J. R. Gair, and M. C. Miller, The Astrophysical Journal **681**, 1431 (2008), URL `http://www.chgk.info/~ilyamandel/papers/mbgm2008.pdf`.

[10] C. L. Rodriguez, M. Morscher, B. Pattabiraman, S. Chatterjee, C.-J. Haster, and F. A. Rasio, Physical Review Letters **115**, 051101 (2015), URL `https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.115.051101`.

[11] B. P. Abbott, R. Abbott, T. D. Abbott, et al. (LIGO Scientific Collaboration and Virgo Collaboration), arXiv e-prints arXiv:1906.08000 (2019), 1906.08000.

[12] J. Aasi, B. Abbott, R. Abbott, T. Abbott, M. Abernathy, T. Accadia, F. Acernese, K. Ackley, C. Adams, T. Adams, et al., Physical Review D **89**, 122003 (2014), URL `https:`

//arxiv.org/abs/1404.2199.
[13] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, et al., Physical Review X **9**, 031040 (2019), URL https://arxiv.org/pdf/1811.12907.pdf.
[14] L. Blackburn, L. Cadonati, S. Caride, S. Caudill, S. Chatterji, N. Christensen, J. Dalrymple, S. Desai, A. Di Credico, G. Ely, et al., Classical and Quantum Gravity **25**, 184004 (2008), 0804.0800.
[15] N. J. Cornish and T. B. Littenberg, Classical and Quantum Gravity **32**, 135012 (2015), 1410.3835.
[16] L. K. Nuttall, T. J. Massinger, J. Areeda, J. Betzwieser, S. Dwyer, A. Effler, R. P. Fisher, P. Fritschel, J. S. Kissel, A. P. Lundgren, et al., Classical and Quantum Gravity **32**, 245005 (2015), 1508.07316.
[17] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, M. Adamo, C. Adams, T. Adams, P. Addesso, et al., Classical and Quantum Gravity **33**, 134001 (2016), 1602.03844.
[18] A. H. Nitz, Classical and Quantum Gravity **35**, 035016 (2018), 1709.08974.
[19] J. Powell, Classical and Quantum Gravity **35**, 155017 (2018), 1803.11346.
[20] M. Cabero, A. Lundgren, A. H. Nitz, T. Dent, D. Barker, E. Goetz, J. S. Kissel, L. K. Nuttall, P. Schale, R. Schofield, et al., Classical and Quantum Gravity **36**, 155010 (2019), 1901.05093.
[21] J. Veitch and A. Vecchio, Phys. Rev. D **81**, 062003 (2010), 0911.3820.
[22] M. Isi, R. Smith, S. Vitale, T. J. Massinger, J. Kanner, and A. Vajpeyi, Phys. Rev. D **98**, 042007 (2018), 1803.09783.
[23] G. Ashton, E. Thrane, and R. J. E. Smith, Phys. Rev. D **100**, 123018 (2019), 1909.11872.
[24] G. Ashton and E. Thrane, MNRAS (2020), 2006.05039.
[25] G. Pratten and A. Vecchio, arXiv e-prints arXiv:2008.00509 (2020), 2008.00509.
[26] B. Abbott, S. Jawahar, N. Lockerbie, and K. Tokmakov, PHYSICAL REVIEW D Phys Rev D **93**, 122003 (2016), URL https://arxiv.org/pdf/1602.03839.pdf.
[27] G. S. Davies, T. Dent, M. Tápai, I. Harry, C. McIsaac, and A. H. Nitz, Phys. Rev. D **102**, 022004 (2020), 2002.08291.
[28] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, V. B. Adya, et al., arXiv e-prints arXiv:1908.11170 (2019), 1908.11170.
[29] W. M. Farr, J. R. Gair, I. Mandel, and C. Cutler, Phys. Rev. D **91**, 023005 (2015), 1302.5341.
[30] 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, et al., Classical and Quantum Gravity **37**, 045007 (2020), 1903.06881.
[31] S. M. Gaebel, J. Veitch, T. Dent, and W. M. Farr, MNRAS **484**, 4008 (2019), 1809.03815.
[32] 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, et al., arXiv e-prints arXiv:1912.11716 (2019), 1912.11716.
[33] A. Nitz, I. Harry, D. Brown, C. M. Biwer, J. Willis, T. D. Canton, C. Capano, L. Pekowsky, T. Dent, A. R. Williamson, et al., *gwastro/pycbc: Pycbc release 1.16.4* (2020), URL https://doi.org/10.5281/zenodo.3904502.
[34] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, Phys. Rev. D **85**, 122006 (2012), gr-qc/0509116.
[35] B. Allen, Phys. Rev. D **71**, 062001 (2005), gr-qc/0405045.
[36] A. H. Nitz, T. Dent, T. Dal Canton, S. Fairhurst, and D. A. Brown, ApJ **849**, 118 (2017), 1705.01513.
[37] 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, et al., Phys. Rev. D **90**, 082004 (2014), 1405.6731.
[38] 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, et al., Classical and Quantum Gravity **33**, 215004 (2016), 1508.02357.
[39] A. H. Nitz, T. Dal Canton, D. Davis, and S. Reyes, Phys. Rev. D **98**, 024050 (2018), 1805.11174.
[40] G. Ashton, M. Hübner, P. Lasky, and C. Talbot, *Bilby: A User-Friendly Bayesian Inference Library* (2019).
[41] J. S. Speagle, MNRAS **493**, 3132 (2020), 1904.02180.
[42] E. Thrane and C. Talbot, PASA **36**, e010 (2019), 1809.02293.
[43] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Physical Review D **93**, 044007 (2016), URL https://arxiv.org/abs/1508.07253.
[44] C. Talbot and E. Thrane, arXiv e-prints arXiv:2006.05292 (2020), 2006.05292.
[45] G. Lodato and P. Natarajan, Monthly Notices of the Royal Astronomical Society **371**, 1813 (2006).
[46] F. Koliopanos, arXiv preprint arXiv:1801.01095 (2018).