

PAPER • OPEN ACCESS

## Establishing significance of gravitational-wave signals from a single observatory in the PyCBC offline search

To cite this article: Gareth S Cabourn Davies and Ian W Harry 2022 *Class. Quantum Grav.* **39** 215012

View the [article online](#) for updates and enhancements.

### You may also like

- [2-OGC: Open Gravitational-wave Catalog of Binary Mergers from Analysis of Public Advanced LIGO and Virgo Data](#)  
Alexander H. Nitz, Thomas Dent, Gareth S. Davies et al.
- [Real-time Search for Compact Binary Mergers in Advanced LIGO and Virgo's Third Observing Run Using PyCBC Live](#)  
Tito Dal Canton, Alexander H. Nitz, Bhooshan Gadre et al.
- [Potential Gravitational-wave and Gamma-ray Multi-messenger Candidate from 2015 October 30](#)  
Alexander H. Nitz, Alex B. Nielsen and Collin D. Capano



**IOP | ebooks™**

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

# Establishing significance of gravitational-wave signals from a single observatory in the PyCBC offline search

Gareth S Cabourn Davies\*  and Ian W Harry 

University of Portsmouth, Portsmouth PO1 3FX, United Kingdom

E-mail: [gareth.davies@port.ac.uk](mailto:gareth.davies@port.ac.uk)

Received 17 March 2022, revised 27 July 2022

Accepted for publication 9 August 2022

Published 29 September 2022



## Abstract

Gravitational-wave observations of compact binary coalescences are allowing us to see black holes and neutron stars further into the Universe and recent results represent the most sensitive searches for compact objects ever undertaken. Most searches for gravitational waves from compact binary coalescence currently rely on detecting coincident triggers from multiple detectors. In this paper, we describe a new method for extrapolating significance of single-detector signals beyond the live-time of the analysis. Using this method, we can recover loud signals which only triggered in a single detector. We demonstrate this method in a search of O3 data, and recover seven single-detector events with a false alarm rate less than two per year. These were the same events as discovered in the GWTC-2.1 and GWTC-3 searches in a single detector, and all but one event from 3-OGC and 4-OGC. Through a campaign of injected signals, we estimate that the total time–volume sensitivity increases by a factor of up to  $1.20 \pm 0.02$  at a false alarm rate of one per two years compared to completely ignoring single-detector events.

Keywords: gravitational-wave astrophysics, gravitational waves, binary compact objects

(Some figures may appear in colour only in the online journal)

\* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

## 1. Introduction

Gravitational-wave searches and observations of compact binary coalescences are helping us to understand more and more about the Universe. Signals detected in the Advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) [1] and Advanced Virgo [2] have allowed us to see more black holes and neutron stars, in more configurations and at further distances than previously observed.

In order to detect and understand these objects, we must first find them within the data. This is done in two ways; in low-latency (e.g. [3]), which aims to rapidly detect signals which can be disseminated to the wider astronomical community to enable follow-up observations, as seen for GW170817 [4]. Additional analyses are performed offline (e.g. [5–8]), which aim to more accurately assess the significance of the candidate events, in order to gain a more confident list of events for use in further study such as parameter estimation [9–11], population analyses [12–14] and tests of general relativity [15].

Multiple search analyses are used for gravitational wave (GW) searches, using different search methods and configurations in order to ensure that we remain able to detect any possible compact binary coalescences (CBCs) in the data. The most sensitive searches for GWs from CBCs, both in low-latency and in offline searches [16–25], are based on comparing the data to a bank of waveform templates [26, 27]. Additionally, unmodelled searches for coherent excess power between detectors [28, 29] can find many of the CBC signals, as well as loud transient signals not from CBCs.

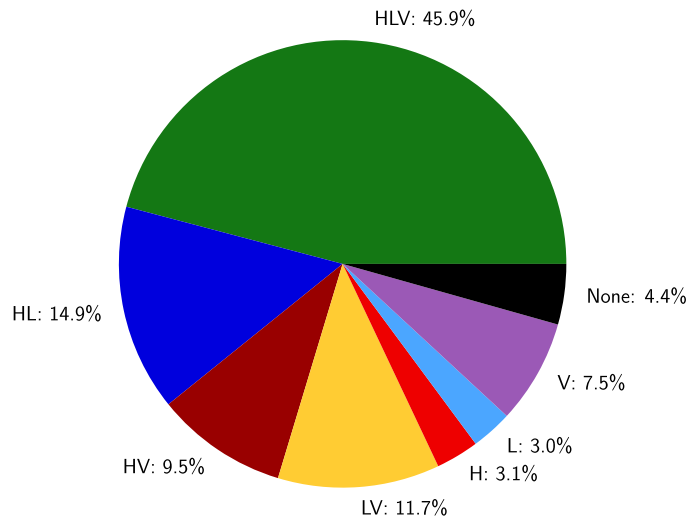
PyCBC is a suite of mainly python-based software for use in the analysis of gravitational-wave data, consisting of a highly modular and configurable set of libraries for searches and parameter estimation of CBCs [30]. PyCBC workflows take advantage of diverse computational resources including local clusters, XSEDE, and the Open Science Grid [31] using the Pegasus workflow management system [32]. PyCBC software includes low-latency and offline CBC searches [3, 16–18], and parameter estimation [9]. This paper utilises the PyCBC offline search for GWs from CBCs.

Gravitational-wave detectors do not have a 100 percent duty cycle, and so we must consider detection of events which occur outside of times when all detectors are operating as designed. Figure 1 shows the fraction of time that different combinations of detectors at LIGO-Hanford, LIGO-Livingston and Virgo were active during the third observing run (O3).

Modelled searches generally use peaks in signal-to-noise ratio (SNR) time-series, *triggers*, from each detector and coincident sets of these triggers are matched with one another to form *events*, which are candidate signals. Standard offline PyCBC searches for GWs, for example, generate coincidences from multiple detectors and then the false alarm rate (FAR) is estimated by counting the number of higher-ranked events in a background manufactured from time shifts [16–18]. There are however a few notable exceptions to this requirement for coincidence in GW searches.

An alternative search pipeline, GstLAL, also has offline and low latency searches. The GstLAL searches assign significance to single-detector events based on likelihood estimates calculated through distributions of triggers in the individual detectors [21]. A ‘singles penalty’ is applied to this likelihood in order to penalise a signal that is not seen in multiple detectors, [33].

Low latency PyCBC searches have also recently introduced a method to estimate FARs [3], based on extrapolating the background using a fit to an exponential decrease in the number of



**Figure 1.** The fraction of time during O3 for which each combination of detectors was observing, with each detector combination signified by its initials, LIGO-Hanford (H), LIGO-Livingston (L) and Virgo (V). Times do not include the month-long commissioning break in October 2019. We see that a significant fraction of the observing run time is when a single observatory is operating (13.6%), or where one of the LIGO observatories is coincident with Virgo only (21.2%).

triggers at a given single-detector ranking statistic. This method is designed to provide alerts in low latency for single-detector events that could have multimessenger counterparts, and so insists on strict event selection cuts which exclude most binary black hole (BBH) mergers.

Recent work means that the PyCBC offline search is able to provide estimates of probability of astrophysical origin  $p_{\text{astro}}$  for single-detector events [34]. However  $p_{\text{astro}}$  should be considered as part of the wider context of all figures of merit of the significance of the candidate event. For example,  $p_{\text{astro}}$  can have significant uncertainty when the underlying signal rate is unknown [35], and particularly for previously undetected populations of CBC (e.g. events with extreme mass ratio). As a result, we wish to also obtain an estimate of the FAR in order to provide further information on the event.

In this paper, we will discuss events in different combinations of detectors in the gravitational wave detector network. Detectors will be referred to by indicative letters, LIGO-Hanford (H), LIGO-Livingston (L) and Virgo (V). The network of detectors will be referred to by the combined initials, Hanford–Livingston–Virgo (HLV), Hanford–Livingston (HL), Hanford–Virgo (HV) and Livingston–Virgo (LV).

The discussion requires two notions of the detector network, the first of which is the network of detectors which is active at the time of the event; we will refer to this as an event being in the time of the detectors, e.g. *HL time* refers to a time when LIGO-Hanford and LIGO-Livingston are operating, but not Virgo. When discussing the time in which exactly one detector is operating, this is *single-detector time*, similarly *double-* and *triple times* refer to when exactly two or three detectors are operating. Times where one-or-more detectors are operating is *any-detector time*, and where two or more detectors is operating is *coincident time*.

The other network we will discuss is the network of detectors which generated triggers that contributed to the event; this is a subset of the active network, and contains only the detectors which contributed to the significance of the event. We will refer to events of this nature as being,

e.g. *HV events* from LIGO-Hanford and Virgo, where these detectors' triggers contribute to the event. Again, these events can be referred to as *single-detector events*, *double* or *triple* events, and *coincident* events, with equivalent definitions to the same terms for the active network.

By adding the ability to detect single-detector events, we gain sensitivity in a number of situations: the first is in the single-detector time, which was the case for 45.3 days, or 13.6% of O3. By allowing the use of single-detector time, in O3 we increase the time available to the search from 273.5 to 318.8 days, an increase of 16.6%. This will not directly correspond to an increase of the same amount in the sensitive volume–time of the search, as the additional time would need to be weighted by the sensitive distance of the available network.

Secondly, we gain sensitivity where a signal is only observed in a single detector but other detectors are operating and did not see the signal. For example, the detection of GW200105\_162426 during the second part of O3 (O3b), where the signal was not seen in Virgo but was in LV time. This was because the signal was not strong enough to be seen in that detector. In O3, we were therefore more likely to have single-detector detections during two-detector time when one of the operating detectors was Virgo, as Virgo was less sensitive than the LIGO detectors. As the two LIGO detectors had similar sensitivities to one another, it is unlikely that a signal would be loud enough to be seen in one detector but not the other. In figure 1 we see that the detector network had one of the LIGO detectors coincident with only Virgo for 70.8 days, or 22.2% of the any-detector time.

Where we have significantly mismatched sensitivities of detectors in a coincident search, we are presented with an additional problem. When calculating significance, we remove confident detections from the estimated background, however we cannot do so for these single-detector events using current methods. If the signal cannot be removed from the background, it can cause contamination in the background used for significance estimates, and cause false alarm rates to be overestimated.

We present here a method for estimating false alarm rates for use in the PyCBC offline search, explaining how we extrapolate beyond the noise background limit in section 2, and how this significance is used within a wider search analysis in section 3. In section 4 we present results of a search on O3 data using this method, comparing to a coincident-only search similar to the PyCBC gravitational-wave transient catalogue (GWTC) analyses GWTC-2.1 for the first part of O3 (O3a) and GWTC-3 for O3b. Section 5 then describes the results of associated injection campaigns for estimates of the increase in sensitivity given by the search.

The results presented here are for the PyCBC-broad and PyCBC-BBH searches as presented in GWTC-2.1 and GWTC-3; these are searches for a broad range of compact binary coalescence parameters, and for systems which are similar to the majority of previously detected BBH systems respectively, with details given in [5, 6].

## 2. Going beyond the edge of the noise distribution

In order to assign a false alarm rate to a candidate event, we compare a ranking statistic—a measure of how signal-like we think the event is—to a background, counting the rate of background events ranked higher than our candidate event. PyCBC coincident searches use time shifts, where the triggers from each detector are shifted relative to the others in order to build up a background which is entirely made up of events which cannot possibly be real, as the time shifts are far greater than the time taken for the GW to travel between detectors. However when considering the background of single-detector events, we cannot build up this background through time shifts, and so we need an alternative method to approximate the background.

Without any form of extrapolation, the highest-ranked single-detector event in a search would have a FAR of one per the live time of the search under the assumption that it, and

all lower-ranked events, are noise. The GstLAL-based CBC search uses extrapolation based on a KDE approach [20]. Other methods use signal and noise populations to estimate a  $p_{\text{astro}}$  based on assumptions of the noise distribution beyond the loudest-ranked event. One option for this is to use the expected signal distribution, normalised to assume a certain number of signals above the second-loudest noise candidate [36]. Alternatively, one can form the noise distribution by assuming it is proportional to the signal density function at high statistic, normalised so that one noise event is expected between the event and next-loudest event [34].

Here we outline a method to estimate single-detector event significance beyond this limit without assuming any signal distribution through simple extrapolation of the background. As we do not assume any signal distribution, we are able to estimate the FAR based solely on the noise distribution, which is more appropriate where the signal distribution is unknown or poorly-known. A single-detector event cannot *a priori* be considered as signal or noise, and so we estimate the bulk distribution of events and assume that it is dominated by noise events; this is a safe assumption in current detectors.

Searches will assign a low FAR to an event which is rare compared to other events in the analysis. The most common events in the data come from Gaussian noise triggers, which have well-described statistical properties. Signals in the data occur rarely in the analysis, and have different properties to the Gaussian-produced events, meaning that they stand out from the background. However, non-Gaussian transient glitches in the data can also produce triggers which stand out from the background, some of which mimic certain types of CBC signal [37].

Most glitches are removed by coincidence requirements; as the glitch times are uncorrelated between detectors they will be removed from consideration as possible events. Without this coincidence requirement, however, non-Gaussian transients could be found to be very significant.

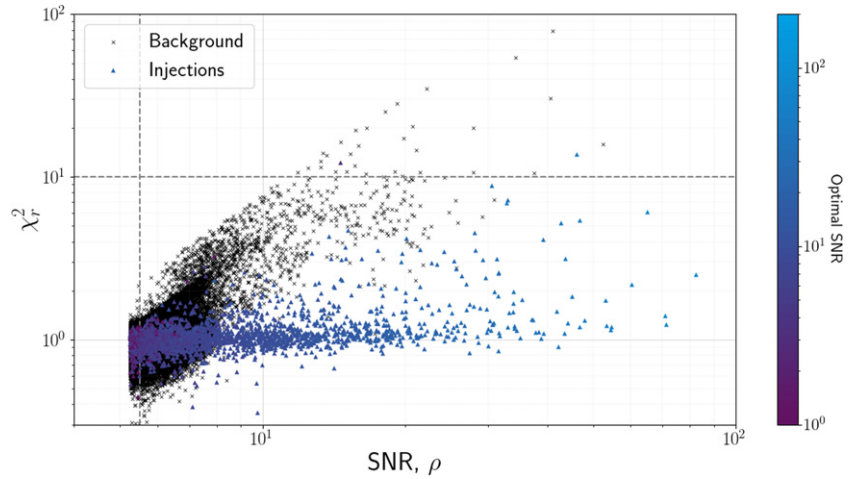
### 2.1. Removal of confident noise events

As a result of these additional difficulties in extrapolating single-detector significance, we will place more stringent limits on certain properties of the candidate events and their associated triggers.

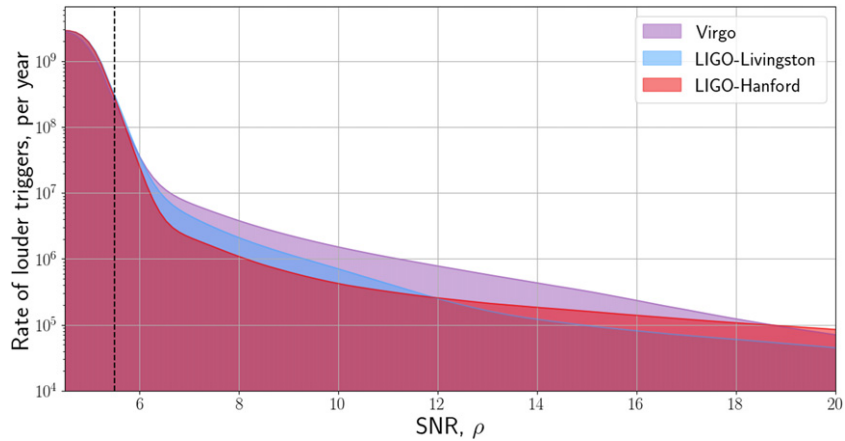
Firstly, in order to prevent unnecessary computational effort, we make a cut on triggers which we can be confident are a part of the noise.

The SNR,  $\rho$ , of the trigger is generally reduced according to certain properties to produce the *reweighted SNR*,  $\hat{\rho}$  [19]. One such property is the reduced  $\chi^2$  discriminator of [38],  $\chi_r^2$ , which checks that the SNR is accumulated in a way which is consistent with the frequency evolution of the template. Instead of only re-weighting the SNR, we decide to impose a stricter criterion, which is a hard cut on the value of  $\chi_r^2$ . By insisting that  $\chi_r^2 < 10$  we remove the loudest and most obvious glitches. We see in figure 2 that background triggers can have much higher  $\chi_r^2$  values than those of (injected) signals, and that nearly all injected signals in this analysis were found using triggers with  $\chi_r^2 < 10$ .

The next cut is on the SNR,  $\rho$ , of the signal, which in this work we enforce must be above 5.5. Figure 3 shows the distribution of trigger SNR in each detector. We see that the chosen SNR criterion of  $\rho > 5.5$  is low enough to be well within the noise distribution, but high enough to still remove the majority of triggers. The triggers are clustered, which means that the trigger with the maximum reweighted SNR within a certain window is kept, and so we need to choose a SNR cut high enough to not be within the region affected by this clustering. Though the



**Figure 2.** Plot of SNR and  $\chi_r^2$  for background triggers and triggers associated with recovered injections for an analysis of seven days during O3b. Each point of the scatter plot represents a single-detector trigger, either as found in a coincident injection (coloured triangles), or in the exclusive background of the coincident search (black crosses). The colour of the injection triggers is based on the optimal SNR; the expected SNR of the injection if it were to be recovered by an exact match to the waveform it was injected with. For this figure, the optimal SNR can be considered as a measure of the loudness of the injection. The dashed lines indicate the cuts on single-detector triggers used before ranking statistic computation, where triggers with SNR  $\rho < 5.5$  or with  $\chi_r^2 > 10$  are removed from consideration as single-detector triggers. The two particularly high  $\chi_r^2$  injections, with  $\chi_r^2$  above ten, are signals injected within seconds of rapid bursts of loud glitches.



**Figure 3.** Count of triggers above given values of SNR for an analysis of seven days during O3b. We see that the SNR requirement,  $\rho > 5.5$ , shown by the dashed line, removes most of the quietest triggers, but retains many triggers for use in understanding the background distribution.

triggers just above the SNR cut-off are still very unlikely to come from signals, they are still informative for the distribution of the single-detector ranking statistic, and so are useful for later stages of the extrapolation of background.

## 2.2. Ranking for false alarm rate calculation

The ranking statistic we use to compare events is based on the ratio of the expected rate density of signals to an empirically measured noise rate density [16]. For coincident events, this takes the form (equation (16) of [16]);

$$R = -\log A_{N\{d\}} - \sum_d \log r_{di}(\hat{\rho}) + \log p(\vec{\Omega}|S) - \log p(\vec{\Omega}|N) + R_{\sigma,i}. \quad (1)$$

The constituent parts of the ranking statistic are:

- $A_{N\{d\}}$  is the allowed ( $(N_{\text{detectors}} - 1)$ -dimensional) time window for coincidences in each detector combination.
- $r_{di}(\hat{\rho})$  is the measured rate density of triggers in template  $i$  and detector  $d$  at reweighted SNR  $\hat{\rho}$  [17]. The measured rate density is that of all triggers, rather than of noise triggers only, and so could include a slight bias from the inclusion of signal triggers, this bias will be naturally very small as the vast majority of triggers will come from noise. To mitigate this bias, we remove triggers within a 0.1 s window of the highest  $\hat{\rho}$  triggers before calculation of the trigger rate density.
- $p(\vec{\Omega}|S)$  is the probability of a signal having the extrinsic parameters,  $\vec{\Omega}$ , of the event (time difference, phase difference, amplitude ratio) given by prior histograms obtained through a Monte Carlo calculation.
- $p(\vec{\Omega}|N)$  is the same probability given a noise distribution, which is assumed to be constant.
- $R_{\sigma,i}$  is (the log of) network sensitive volume for a given template and triggered detector network, which is proportional to the expected rate of signals, normalised compared to a reference network.  $R_{\sigma,i} \equiv 3(\log \sigma_{\min,i} - \log \bar{\sigma}_{HL,i})$  where  $\sigma_i$  is the expected SNR of a signal with the same parameters as template  $i$  directly overhead at a distance of 1 Mpc [39].  $\sigma_{\min,i}$  is where we take the minimum over detectors in the triggered network, and  $\bar{\sigma}_{HL,i}$  is the reference sensitivity, for which we use the minimum over detectors of the median  $\sigma$  for triggers from template  $i$  in the HL detector network. The reference sensitivity is kept the same for all detector combinations in order to ensure that the ranking statistic is comparable over different triggered network configurations.

For the single-detector event ranking statistic, we remove all terms which come from the coincident nature of the event. We therefore remove  $A_{N\{d\}}$ ,  $p(\vec{\Omega}|S)$ , and  $p(\vec{\Omega}|N)$  and the ranking statistic becomes

$$R = -\log r_{di}(\hat{\rho}) + R_{\sigma,i}. \quad (2)$$

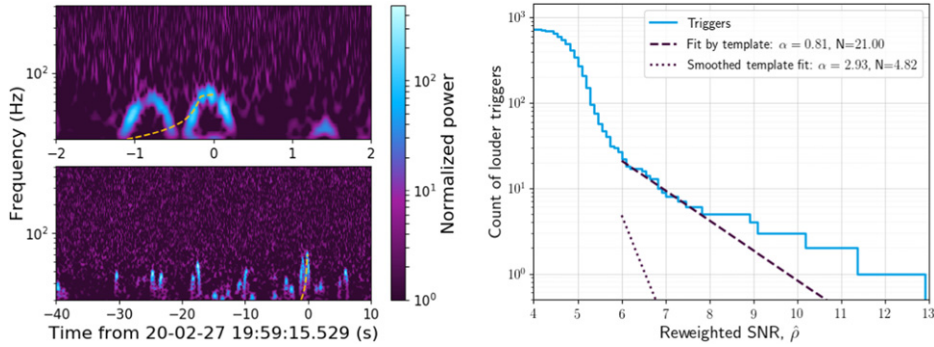
Outside of the differences in sensitive distance, the ranking statistic is therefore entirely dependent on the measured rate density of triggers in each template and detector.

The measured trigger rate for each template and detector,  $r_{di}(\hat{\rho})$ , is modelled as an exponential decay of the reweighted SNR  $\hat{\rho}$  of triggers [17];

$$r_{di}(\hat{\rho}) = N_{di} \alpha_{di} \exp[-\alpha_{di}(\hat{\rho} - \hat{\rho}_{\text{thresh}})], \quad (3)$$

where  $N_{di}$  is the number of triggers in the template with  $\hat{\rho} > \hat{\rho}_{\text{thresh}}$ ,  $\hat{\rho}_{\text{thresh}}$  is a threshold value of reweighted SNR, and  $\alpha_{di}$  is the exponential decay constant.





**Figure 4.** Example of problematic smoothing of trigger distribution fits (right), and a loud trigger from this template (left). We see that the triggers in the template are fit with a falling exponential with fit coefficient  $\alpha = 0.81$ , and there are 21 triggers with reweighted SNR above threshold. However once this is smoothed over local templates, the exponential fit steepens to  $\alpha = 2.93$  and there are 4.82 triggers above threshold per template. As a result, the rate of triggers at the SNR of the loudest events falls by many orders of magnitude, boosting the ranking of these events. The example trigger has reweighted SNR 22.5, which would have had a trigger count density of around  $2.7 \times 10^{-5}$ , which changes to a count density of  $1.4 \times 10^{-20}$  with the smoothed fit parameters. N.B. The example trigger is removed from the triggers used for fitting by a process which removes the loudest triggers in the search. This example trigger comes from scattered light in the detector.

We use a maximum-likelihood procedure to fit the  $\alpha_{di}$  and  $N_{di}$  parameters, which are then smoothed over nearby templates. The smoothed  $\alpha_{di}$  and  $N_{di}$  are then used to measure the rate of triggers at that value of  $\hat{\rho}$ .

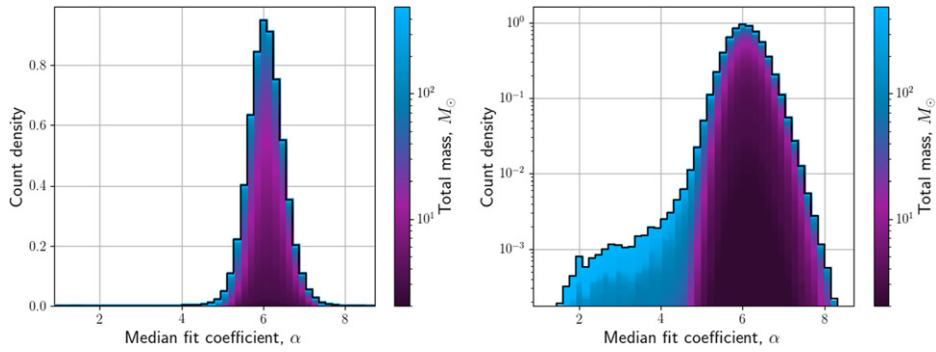
The smoothing is useful for templates where we have very few triggers with  $\hat{\rho} > \hat{\rho}_{\text{thresh}}$ , and we gain more triggers to avoid problems with small sample-size statistics. However we can face occasional problems with this procedure.

The most problematic situation for this procedure would be where a set of detector artefacts trigger many times in a specific template, but not in nearby templates, for example this can be the case for scattered-light artefacts [40, 41].

The template fit in figure 4 shows an example of this, with a slow exponential decay in this template. However when it is smoothed by surrounding templates, the expected exponential decay is much steeper, and the calculated rate of triggers at this reweighted SNR would be much lower than it should be. Given Gaussian noise, we would expect the decay in each template to follow an exponential decay distribution as a function of reweighted SNR with decay constant of around 6<sup>1</sup>, as seen in figure 5.

In order to remove the effect of over-smoothing where it results in overly optimistic rate estimates, we make our final cut on triggers in templates with an exponential decay constant  $\alpha$  below 2.5 before the smoothing is applied. This has the effect of removing a relatively small fraction of the bank in most cases, as seen in figure 5 and table 1. Table 1 shows the number of removed templates for each search, with summary values over different chunks of analysis.

<sup>1</sup> As the signal consistency tests we perform are one-sided, and only penalise rather than boost the SNR in reweighted SNR calculations, the exponential decay constant ends up being just below 6.



**Figure 5.** A histogram of fit coefficients from LIGO-Hanford during the O3 PyCBC-broad analysis, left, and with logarithmic count density plotted, right. We use the median of the fit coefficient in each template from different analyses over the course of the observing run in order to remove significant outliers and cases where templates had no triggers with reweighted SNR above threshold. We see that the peak of the distribution is around six. The colour of the histogram indicates the total mass of the template which contributed to the bin of the histogram, and we see that many of the higher-mass templates have low fit coefficients; this comes from the relatively short high-mass templates matching well to glitches and therefore having increased trigger rates at higher SNR. The lower-mass templates generally have a well-constrained distribution of fit coefficients around the mode.

**Table 1.** Number and fraction of templates which had their triggers removed from single-detector event analysis in each detector for the chunks of data analysed by the PyCBC-broad and PyCBC-BBH searches. We present the maximum number and the median, showing worst-case and usual scenarios. We expect a small number of templates to be removed due to Gaussian noise fluctuations, dependent on the threshold SNR value, which we use  $\hat{\rho}_{\text{thresh}} = 6$ , and proportional to the number of templates and the amount of time analysed. The PyCBC-broad search used 428 725 templates in one week chunks, and the PyCBC-BBH search used 17 094 templates in chunks of around a month, explaining the many fewer affected templates.

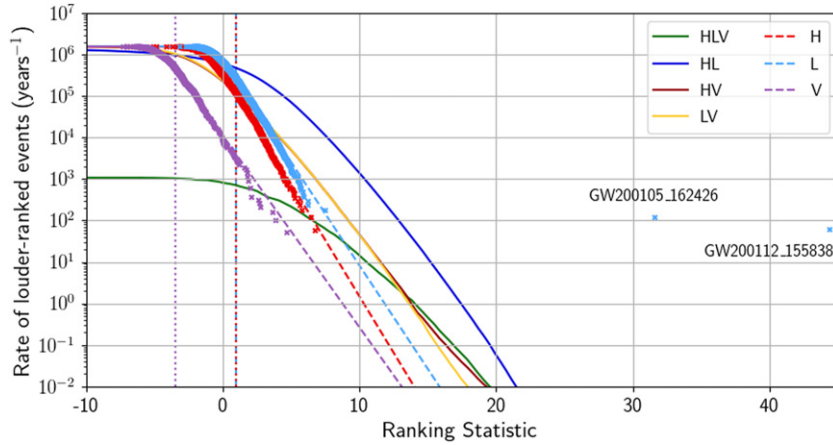
		PyCBC-broad	PyCBC-BBH
LIGO-Hanford	Maximum	2202 (0.51%)	21 (0.12%)
	Median	214 (0.05%)	14 (0.08%)
LIGO-Livingston	Maximum	1324 (0.31%)	31 (0.18%)
	Median	287 (0.07%)	20 (0.12%)
Virgo	Maximum	19 747 (4.61%)	261 (1.53%)
	Median	122 (0.03%)	1 (0.01%)

The observing runs are split into shorter sections, which we call chunks, for analysis during an offline search, and these chunks are analysed independently of one another to allow for changes in detector sensitivity and data quality throughout the observing run.

This cut means that the example trigger shown in figure 4 from a template with  $\alpha = 0.81$  would not be used further in the analysis.

### 2.3. Extrapolation of false alarm rate

In order to extrapolate the false alarm rate beyond the limit of one event per live time of the search, we fit the number of single-detector events above a particular ranking statistic to a



**Figure 6.** Rate of louder-ranked events in the exclusive background vs ranking statistic for the PyCBC-broad analysis of data between 2020-01-04 17:06:58 and 2020-01-13 10:28:01. For single-detector events we see the extrapolated background rate above the threshold ranking statistic, and events with the rate of equal-or-louder ranked events as scatter points. The lower thresholds for inclusion in the exponential fit are shown as dashed or dotted vertical lines at 1 for LIGO-Hanford and LIGO-Livingston, and at  $-3.5$  for Virgo. The FAR for single-detector events is calculated by finding the extrapolated rate of louder events, the dotted line, at the ranking statistic of the event. For contrast, we have plotted scatter points which would be the FAR when counting equal-or-higher ranked events. It may seem somewhat counter-intuitive for the coincident events to have higher FARs, however this is as the distribution of single-detector event ranking statistic peaks around lower values, as the ranking statistic is lower given that we believe they are less likely to be real events. A single-detector event with a high ranking statistic comparable to those found by coincident searches is therefore extremely rare and has a low FAR.

falling exponential. This is similar to the falling exponential used to estimate the rate of triggers at a particular reweighted SNR described above.

We fit this exponential for events above a certain ranking statistic threshold. This threshold must be chosen carefully; we must ensure that enough events are used in the fitting to obtain a confident event rate and exponential decay constant, and balance this consideration against avoiding the effects of clustering the events at lower ranking statistic. For example, in our analysis of O3 data in section 4, we use fit thresholds of 1 for H and L events, and of  $-3.5$  for V events, as seen in figure 6; this different threshold is mainly due to the sensitivity term  $R_{\sigma,i}$  being much lower in Virgo than in the LIGO detectors, down-ranking the candidate events. In order to obtain accurate FAR estimates for lower-significance candidate events, we revert to counting the number of louder-ranked events below this threshold. We do not use an explicit upper threshold on the ranking statistic for inclusion in the fit, but any triggers within  $0.1$  s of a (coincident or single-detector) event with FAR below 33.3 per year are removed from the fits.

By using the extrapolation of the ranking statistic–false alarm rate relationship, we are now able to estimate significance beyond the limit of one per live time. The extrapolation method we have detailed here is for single-detector events only, and so we now consider it as part of a wider search including coincident events.

### 3. Using the estimated significance in a wider search context

Once we have calculated the FARs for the events in their detector combinations, these are combined with results from all possible combinations. The best-ranked event from any detector combination at a given time is used, and so through the ‘look elsewhere’ effect, we will find more false alarms than we would looking in a single detector combination.

Often this effect would be included through multiplication by a trials factor, but that would not be appropriate here given the vastly different rate of false alarms in the different possible combinations of detectors. We follow the same method as [16], and obtain an overall FAR by summing FARs at the ranking statistic of the candidate event in all available combinations at that time. This means that a candidate event in HLV time would have the FAR at its ranking statistic combined from the H, L, V, HL, HV, LV and HLV FARs, but a candidate event in LV time would only have the LV, L and V FARs added together.

We consider what happens in this procedure in the context of single-detector events.

For a FAR based on an exponential fit with no limits at high statistic, highly-ranked events will get negligible FAR. At high statistic, the coincident backgrounds are much higher than the estimated backgrounds for single-detector events, as seen in figure 6. At very high statistic, if a coincident event is ranked louder than all background events, its FAR is set to one per background time; the amount of time which the time-shifted background explores. The background time depends on the analysis time and the number of time shifts, and is of the order of tens of millenia for both PyCBC-broad and PyCBC-BBH.

An event with very low FAR is therefore naturally mitigated by the combination with the coincident backgrounds when other detector combinations are available. Additionally, the single-detector backgrounds do not have a large effect on the significance estimates of confidently-detected coincident events.

For single-detector events in single-detector time, we set a limit on FAR corresponding to the largest available coincident background time, to prevent unbelievably small FARs. This cut-off only affects the very loudest events, and once we get to low values of FAR, we do not require more confidence, as we are already almost certain that the event is real. In catalogue papers, FARs are often quoted as being below a particular cutoff; in 3-OGC and 4-OGC, this was a FAR of one per 100 years, and in GWTC-2.1 and GWTC-3, this was one per  $10^5$  years.

By combining the extrapolated FAR of single-detector events with the FAR from coincident events at the same ranking statistic, we have included the single-detector events in the PyCBC offline search, ensuring that the results remain sensible for all events.

### 4. Results from the third LIGO observing run

In order to test our method, we compare to a coincident-only PyCBC search, similar to the analyses used in GWTC-2.1 from O3a [33] and GWTC-3 from O3b [6].

In each of these papers, there were two analyses, PyCBC-broad and PyCBC-BBH, which differ slightly from one another. The PyCBC-BBH analysis uses a much smaller template bank focussed on the stellar-mass black hole region in which we have found many events already, whereas PyCBC-broad is a much wider analysis, including binary neutron star (BNS) and neutron star-black hole binary (NSBH) events, and is able to recover events which may be outside of the parameter space in which we have already seen events.

In addition, the PyCBC-BBH analysis uses an explicit model of the black hole mass population, and so the ranking statistic has an additional term of  $-\frac{11}{3} \log(\mathcal{M}_i/\mathcal{M}_{\text{ref}})$ , where  $\mathcal{M}_i$  is the chirp mass [42, 43] of the triggered template and  $\mathcal{M}_{\text{ref}} = 20M_{\odot}$  is a reference chirp mass

[44, 45]. This term should be truncated above  $\mathcal{M} = 40M_{\odot}$ , however due to an error in implementation this was not applied in the GWTC-2.1 and GWTC-3 results [6]. This bug up-ranked many triggers, and as single-detector events are particularly sensitive to issues in the ranking statistic, the bug was corrected for this work in all our PyCBC-BBH analyses.

In the GWTC papers, in order to prevent background contamination by single-detector events, background events were removed if they contained triggers within  $\pm 0.1$  s of events found by other searches with FAR below one per 100 years, provided that they did not form coincidences in the PyCBC searches at any significance. This one-per-hundred-years limit was designed to match the hierarchical removal stage of the PyCBC analysis.

In GWTC-2.1, the events used for trigger removal came from the GWTC-2 catalogue [5]. For GWTC-3, the events used for trigger removal came from events found by the GstLAL search in that paper.

In 4-OGC [8], the  $p_{\text{astro}}$  calculation for candidate events in a single detector utilised only the background when both LIGO detectors are observing in order to minimise possible signal contamination. Coincident events do not have the single-detector signals removed from the background, as the primary figure of merit for that work is  $p_{\text{astro}}$ , which remains high for BBH events at the point where background contamination becomes an issue. The FAR is a more suitable figure of merit for events where the signal distribution is less understood, and we see in 4-OGC that the single-detector BNS and NSBH events GW190425 and GW200105\_162426 are recovered with  $p_{\text{astro}} \sim 0.5$ , limited by the signal rates in that region.

To compare coincident-only analyses with the singles-included search, we re-analysed the data using the same method as in GWTC-2.1 and GWTC-3, but did so without removing the single-detector event triggers listed above from the background, this is denoted as the *coincident-only* search. This means that the results we compare to are not as accurate or optimistic as the results from those analyses, however they are more representative of independent offline PyCBC analyses than as presented in the GWTC catalogues.

In addition to the non-removal of single detector events from the background, the results do not exactly match the results in the GWTC papers for two reasons; the GWTC papers use probability of astrophysical origin  $p_{\text{astro}}$  as the threshold for inclusion, and the PyCBC results in those papers state the inclusive FAR, but the FAR used in this paper is exclusive.

The inclusive false alarm rate is defined under the assumption that the event of interest, and any event ranked lower is noise, and so all triggers at the time of the event are included in the background. The FAR given by an inclusive background is estimated by successively removing the triggers within a 0.1 s window of the foreground events in descending order of FAR, calculating FAR for each event before removing nearby triggers from the background. We do this for all events with FAR below one per 100 years. The exclusive false alarm rate is built under the assumption that all the events we find are signals, and so all triggers at the time of events are excluded from the background. We therefore remove all triggers within a 0.1 s window of any event with FAR less than one per 0.03 years for this calculation.

The inclusive FAR is not useful for single-detector signals, as it will always be less than the live-time of the search. One could assume a distribution of signal ranking statistics in order to extrapolate the inclusive FAR beyond this limit [34, 36], but in this work, we do not assume any signal distribution properties. The analyses are split into chunks with live times of around  $\sim$  one week for PyCBC-broad analysis and  $\sim$  one month for the PyCBC-BBH analysis, and so the un-extrapolated inclusive FAR would always be found to be insignificant.

Tables 2 and 3 give the events found in O3a and O3b respectively using this search technique, compared to the results of the coincident-only analyses.

In tables 2 and 3, we see that we manage to recover all of the single-detector signals from GWTCs 2.1 and 3, and all but GW190424\_180648 from 3-OGC and 4-OGC.

**Table 2.** Results from the PyCBC-broad and PyCBC-BBH analyses with singles-included and coincident-only searches in O3a. We list the FARs and network SNRs from each event for all analyses. Included are events with FAR less than two per year in any analysis, except those for which the FARs do not differ between the searches; these events are included in table A1 in appendix A. Events with names in bold were found according to the FAR criterion by this work, but did not reach the same criterion in the coincident analyses by the same search. Events with network SNR and FAR given in italics are included as they are at the same time as one which meets the criterion for inclusion. Instruments are given according to the initials of the detectors involved, H, L or V, e.g. a HL event comes from LIGO-Hanford and LIGO-Livingston, but not Virgo. In some cases, the set of instruments which contributed to the events differs between the analyses; for these events, we have listed the largest group of instruments which triggered in the instruments column. Where a subset of the instruments listed in the instruments column triggered to form the event, this is indicated with a †.

Event	Instruments Active, triggered	Search including single-detector events				Coincident-only search			
		PyCBC-broad		PyCBC-BBH		PyCBC-broad		PyCBC-BBH	
		FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR
<b>GW190814_211039</b>	LV, LV	$<2.5 \times 10^{-5}$	†20.3	$<1.7 \times 10^{-5}$	†20.3	$<2.5 \times 10^{-5}$	19.5	—	—
<b>GW190425</b>	LV, L	$1.8 \times 10^{-5}$	12.2	—	—	—	—	—	—
<b>GW190630_185205</b>	LV, LV	$<2.1 \times 10^{-5}$	†14.6	$<3.0 \times 10^{-5}$	15.1	—	—	0.012	15.1
<b>GW190708_232457</b>	LV, L	$<2.1 \times 10^{-5}$	12.7	$<3.0 \times 10^{-5}$	12.7	—	—	—	—
GW190925_232845	HV, HV	76	9.0	$6.8 \times 10^{-5}$	9.9	71	9.0	$1.0 \times 10^{-4}$	9.9
GW190915_235702	HLV, HL	$2.9 \times 10^{-4}$	13.1	$<7.6 \times 10^{-5}$	13.1	$7.0 \times 10^{-4}$	13.1	$<7.6 \times 10^{-5}$	13.1
GW190828_065509	HLV, HL	$2.8 \times 10^{-4}$	10.6	$<7.6 \times 10^{-5}$	10.6	$2.9 \times 10^{-4}$	10.6	$<7.6 \times 10^{-5}$	10.6
GW190728_064510	HLV, HL	$<8.7 \times 10^{-5}$	13.1	$<8.4 \times 10^{-5}$	13.1	$<8.7 \times 10^{-5}$	13.1	0.040	13.1
GW190720_000836	HLV, HLV	$1.4 \times 10^{-4}$	†10.7	$<8.4 \times 10^{-5}$	11.4	$1.4 \times 10^{-4}$	†10.7	0.068	11.4
GW190727_060333	HLV, HL	0.0046	11.4	$<8.4 \times 10^{-5}$	11.1	0.0047	11.4	0.096	11.1
GW190517_055101	HLV, HL	0.0075	10.5	$<1.2 \times 10^{-4}$	10.3	0.0087	10.5	$1.4 \times 10^{-4}$	10.3
GW190706_222641	HLV, HL	$3.4 \times 10^{-4}$	11.7	$<1.2 \times 10^{-4}$	12.6	$3.4 \times 10^{-4}$	11.7	0.13	12.6
GW190701_203306	HLV, HLV	0.058	11.9	$<1.2 \times 10^{-4}$	11.7	0.059	11.9	0.33	11.7
GW190503_185404	HLV, HL	0.035	12.2	$<1.3 \times 10^{-4}$	12.2	0.037	12.2	$<1.3 \times 10^{-4}$	12.2

(continued on next page)

Table 2. Continued.

Event	Instruments Active, triggered	Search including single-detector events				Coincident-only search			
		PyCBC-broad		PyCBC-BBH		PyCBC-broad		PyCBC-BBH	
		FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR
<b>GW190620_030421</b>	LV, L	13	11.2	$3.5 \times 10^{-4}$	11.0	—	—	—	—
GW190930_133541	HL, HL	0.017	9.9	$8.2 \times 10^{-4}$	10.1	0.018	9.9	0.0015	10.1
GW190421_213856	HL, HL	5.9	10.2	0.0014	10.2	5.9	10.2	0.0025	10.2
GW190602_175927	HLV, HL	1.5	11.9	0.0026	11.9	0.27	11.9	0.0038	11.9
GW190513_205428	HLV, HL	19	11.7	0.0028	11.8	19	11.7	0.0041	11.8
GW190521_030229	HLV, HL	0.41	13.7	0.0031	13.7	0.42	13.7	0.0046	13.7
<b>GW190910_112807</b>	LV, L	$1.9 \times 10^4$	11.5	0.0054	13.4	—	—	—	—
GW190803_022701	HLV, HL	80	8.7	0.039	8.7	80	8.7	0.24	8.7
GW190413_134308	HLV, HL	20	9.4	0.044	9.3	21	9.4	0.070	9.3
GW190413_052954	HLV, HL	170	8.6	0.13	8.6	170	8.6	0.19	8.6
GW190719_215514	HL, HL	$1.5 \times 10^3$	8.1	0.24	8.0	$1.4 \times 10^3$	8.1	0.28	8.0
GW190725_174728	HLV, HL	0.44	9.1	0.65	8.8	0.44	9.1	0.95	8.8
GW190413_214724	HV, HV	220	9.0	0.46	8.6	210	9.0	1.1	8.6
GW190805_211137	HLV, HL	$2.0 \times 10^3$	8.1	0.51	8.3	$1.9 \times 10^3$	8.1	0.79	8.3
GW190731_140936	HL, HL	$2.1 \times 10^3$	8.2	0.60	8.2	$2.0 \times 10^3$	8.2	0.68	8.2
GW190514_065416	HL, HL	$1.4 \times 10^3$	8.5	0.84	8.5	$1.3 \times 10^3$	8.5	0.96	8.5
GW190816_213147	LV, LV	$1.8 \times 10^3$	8.4	1.0	8.4	920	8.4	1.3	8.4

**Table 3.** Results from the PyCBC-broad and PyCBC-BBH analyses with singles-included and coincident-only searches in O3b. We list the FARs and network SNRs from each event for all analyses. The event inclusion criteria and table format are the same as in table 2.

Event	Instruments Active, triggered	Search including single-detector events				Coincident-only search			
		PyCBC-broad		PyCBC-BBH		PyCBC-broad		PyCBC-BBH	
		FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR
<b>GW200112_155838</b>	LV, L	$<1.9 \times 10^{-5}$	18.1	$1.4 < \times 10^{-5}$	18.3	—	—	—	—
<b>GW200105_162426</b>	LV, L	$<1.9 \times 10^{-5}$	13.1	—	—	—	—	—	—
GW191222_033537	HL, HL	0.0020	11.5	$2.6 \times 10^{-5}$	11.5	0.0020	11.5	$2.6 \times 10^{-5}$	11.5
GW200208_130117	HLV, HLV	0.18	†9.6	$<7.8 \times 10^{-5}$	10.8	0.18	†9.6	$<7.8 \times 10^{-5}$	10.8
GW191215_223052	HLV, HL	0.0014	10.4	$8.8 \times 10^{-5}$	10.3	0.0015	10.4	0.44	10.3
GW200115_042309	HLV, HL	$<1.2 \times 10^{-4}$	10.9	$3.3 \times 10^6$	†5.9	$<1.2 \times 10^{-4}$	10.9	—	—
GW200128_022011	HL, HL	0.60	9.9	$1.3 \times 10^{-4}$	10.0	0.62	9.9	$2.3 \times 10^{-4}$	10.0
GW200219_094415	HLV, HL	1.7	10.0	$3.4 \times 10^{-4}$	10.0	1.7	10.0	$6.2 \times 10^{-4}$	10.0
GW200202_154313	HLV, HL	8.4	†9.7	$6.0 \times 10^{-4}$	10.8	—	—	0.0011	10.8
GW191105_143521	HLV, HL	0.0097	9.8	0.0018	9.8	0.012	9.8	0.0037	9.8
<b>GW200302_015811</b>	HV, H	$1.0 \times 10^4$	9.4	0.0031	10.5	—	—	—	—
GW191109_010717	HL, HL	0.087	13.2	0.013	14.9	0.095	13.2	0.018	14.9
GW200316_215756	HLV, HL	0.19	9.4	0.027	9.4	0.20	9.4	0.34	9.4
GW191103_012549	HL, HL	4.7	9.3	0.075	9.3	4.8	9.3	0.095	9.3
GW191230_180458	HLV, HL	52	9.6	0.098	9.9	52	9.6	0.74	9.9
GW200209_085452	HLV, HL	540	9.3	0.39	9.3	540	9.3	0.46	9.3
GW200308_173609	HLV, HL	770	8.0	0.45	8.0	760	8.0	1.0	8.0
GW191204_110529	HL, HL	980	9.0	0.61	9.0	980	9.0	0.75	9.0
GW191127_050227	HLV, HL	20	9.5	0.87	8.7	20	9.5	1.00	8.7
GW191126_115259	HL, HL	22	8.6	0.92	8.6	22	8.6	1.0	8.6
GW191118_212859	LV, LV	1.4	9.2	—	—	1.2	9.2	—	—
GW200311_103121	HL, HL	1.3	9.3	—	—	1.3	9.3	—	—
<b>200218_201552</b>	V, V	$1.0 \times 10^3$	10.3	1.7	10.5	—	—	—	—



GW190424\_180648 is a LIGO-Livingston-only event recovered with  $p_{\text{astro}} = 0.81$  in 3-OGC but is not found with any significance by this analysis.

Event names encode the event time according to the convention GWYYMMDD\_hhmmss, for example the event GW200112\_155838 was found at time 2020-01-12 15:58:38 UTC. The event 200218\_201552 does not have the ‘GW’ prefix, as it comes from a glitch in the detector, as discussed later. The network SNR is calculated from the sum of squares of SNRs from the triggers which form the event, and does not account for detectors that did not produce triggers contributing to the event.

The events listed in bold come from a few broad categories of events. Firstly we see seven single-detector events; GW190425, GW190620\_030421, GW190708\_232457, GW190910\_112807, GW200105\_162426, GW200112\_155838 and GW200302\_015811, which are assigned significant FAR for the first time by PyCBC searches.

Secondly, there are events which were found by either the PyCBC-broad or PyCBC-BBH search in the coincident-only search, but which are found in the other search as a single-detector event; this was the case for GW190630\_185205, which is newly found in the PyCBC-broad search. The FAR of GW190630\_185205 is also significantly improved in the PyCBC-BBH search for a different reason, as the triggers around the event GW190708\_232457 are removed from the background.

The event GW190814\_211039 is newly found in the PyCBC-BBH search as a single event. Though the FAR of GW190814\_211039 does not change in the PyCBC-broad search, as it is already ranked higher than the loudest background event, the event changes from being an LV coincident event with ranking statistic 45.15 to an L single-detector event with ranking statistic 65.98.

Finally we see 200218\_201552 which, as can be seen in figure 7(i), is caused by a glitch in the Virgo detector. This event has ranking statistic 7.02, which would have FAR of over ten per year for LIGO-Hanford and LIGO-Livingston singles, and hundreds per year for any coincident events. As a result, 200218\_201552’s significance is boosted by the fact that it is a Virgo single-detector event in Virgo-only time and does not have backgrounds from other event types suppressing it.

No events lost significance in order to move above the two-per-year threshold for inclusion in the results.

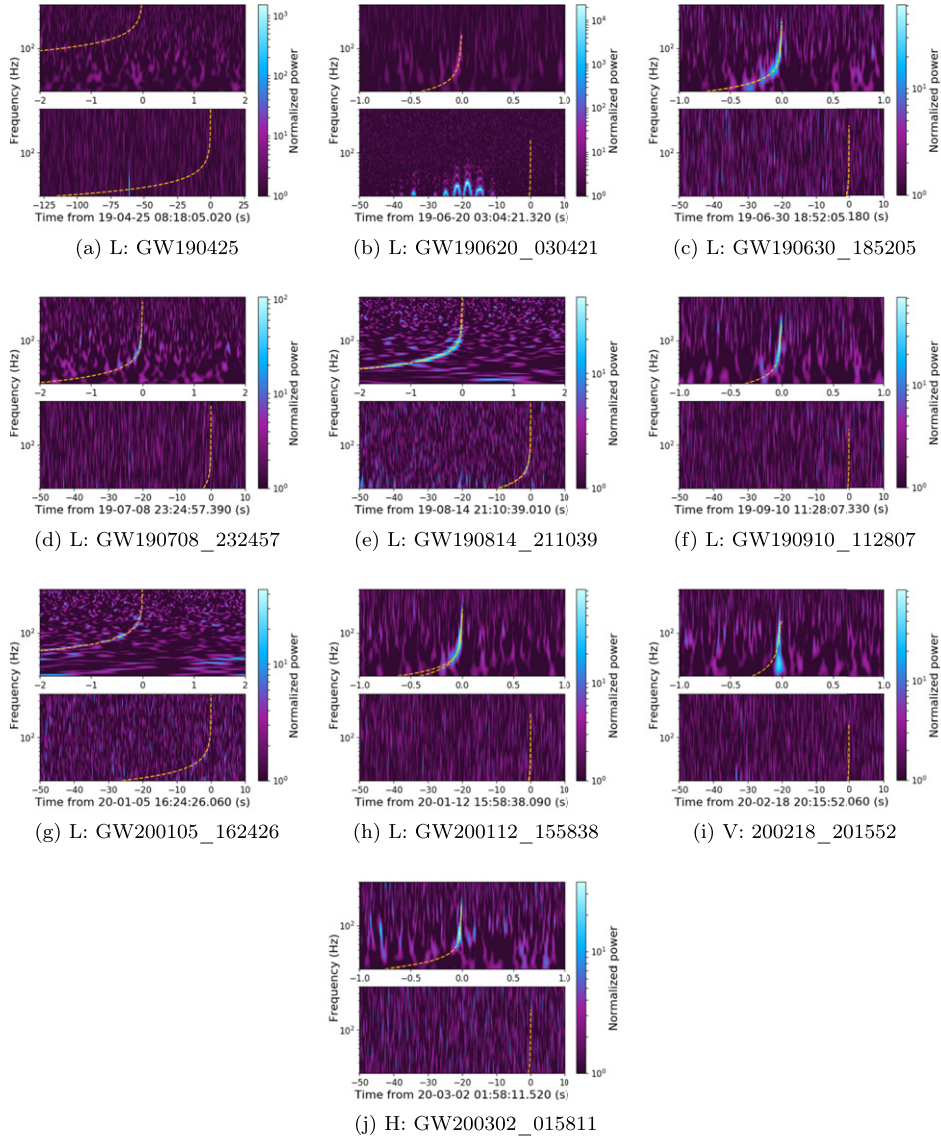
Figure 7 shows time–frequency plots of the events newly-found by this work. We see that a few of the events are strong enough to be seen by eye in the plots, these are the strong signals we would be concerned about missing in a coincident-only search.

We have seen that using the method described here, we are able to recover signals seen in a single detector, as well as improving our estimates of the significance of other signals by removing the triggers close to single-detector events from the background estimates of coincident events.

## 5. Search sensitivity

We have seen the results on O3 data in section 4, and here we consider injected signals in the data. By injecting signals from various parts of the CBC parameter space, we can assess the change in sensitivity to different signals. By comparing which injections were recovered in each search, we can estimate the change in the number of events we would expect to find. The injections we use are the same as those used in the GWTC-2.1 and GWTC-3 catalogues, and their distributions are described fully in the appendix of the GWTC-3 paper [6].

First we will discuss the change in the sensitivity of the search, and then discuss the situations which contribute to the change in sensitivity. To do this, we compare the sensitive

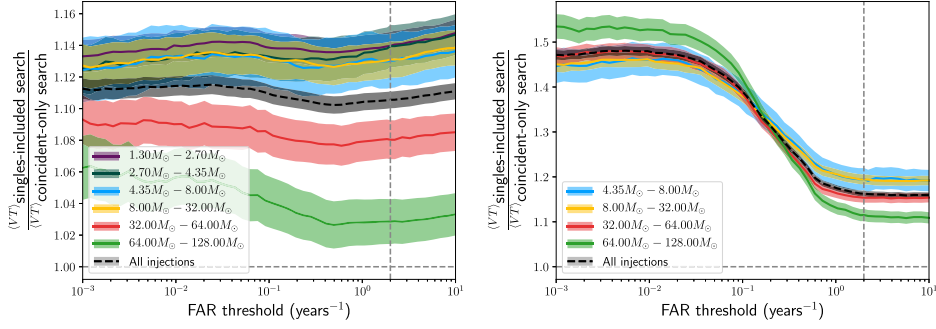


**Figure 7.** Time-frequency plots for the single-detector events found by the singles-included analyses. The subcaption indicates the detector and which event is plotted. The time–frequency track of the template(s) which triggered from the single-detector searches is (are) overlaid.

volume–time ( $\langle VT \rangle$ ) of each search, which is a measure proportional to the number of signals we would expect to see in a search.

Figure 8 shows the ratio of the sensitive volume–time ( $R_{VT}$ ) of each analysis, where

$$R_{VT} = \frac{\langle VT \rangle_{\text{singles-included search}}}{\langle VT \rangle_{\text{coincident-only search}}}, \quad (4)$$



**Figure 8.** The ratio of the  $\langle VT \rangle$  of the singles-included and coincident-only searches at different FAR thresholds for the PyCBC-broad (left) and PyCBC-BBH (right) analyses. Injections in different chirp mass bins are denoted by colour, and the  $\langle VT \rangle$  for the whole population is also included as the black dashed line. Uncertainty bands are based on Poisson counting uncertainty for the found injections.

**Table 4.** Sensitive  $\langle VT \rangle$  ratio for the PyCBC-broad and PyCBC-BBH searches for signals in various chirp mass bins, and averaged for all injections.

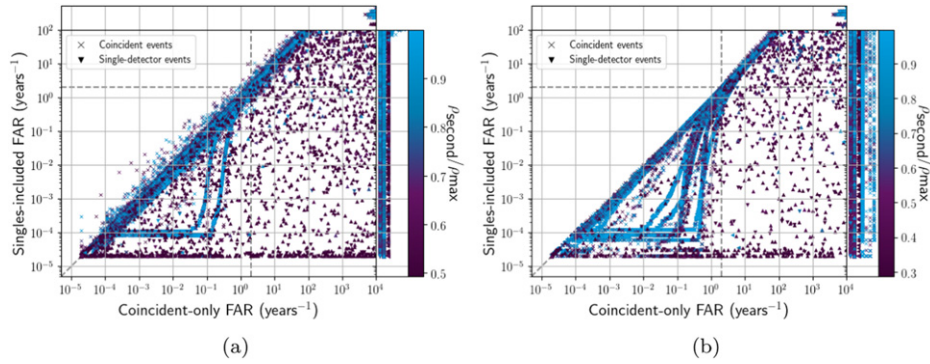
Chirp mass bin	PyCBC-broad $R_{\langle VT \rangle}$	PyCBC-BBH $R_{\langle VT \rangle}$
$1.30M_{\odot} < \mathcal{M} \leq 2.70M_{\odot}$	$1.14 \pm 0.01$	--
$2.70M_{\odot} < \mathcal{M} \leq 4.35M_{\odot}$	$1.14 \pm 0.01$	--
$4.35M_{\odot} < \mathcal{M} \leq 8.00M_{\odot}$	$1.13 \pm 0.02$	$1.20 \pm 0.02$
$8.00M_{\odot} < \mathcal{M} \leq 32.00M_{\odot}$	$1.13 \pm 0.01$	$1.19 \pm 0.01$
$32.00M_{\odot} < \mathcal{M} \leq 64.00M_{\odot}$	$1.08 \pm 0.01$	$1.15 \pm 0.01$
$64.00M_{\odot} < \mathcal{M} \leq 128.00M_{\odot}$	$1.03 \pm 0.01$	$1.12 \pm 0.01$
All injections	$1.11 \pm 0.01$	$1.16 \pm 0.01$

which is estimated by counting the number of injected signals and the number of recovered signals as a function of FAR.

We use chirp mass ( $\mathcal{M}$ ) bins to show the effect of including single-detector events on the search in different parts of the parameter space. The bins  $1.30M_{\odot} < \mathcal{M} \leq 2.70M_{\odot}$  and  $2.70M_{\odot} < \mathcal{M} \leq 4.35M_{\odot}$  match the BNS and NSBH chirp mass bins used for the  $p_{\text{astro}}$  calculations for the PyCBC searches in [6].

We see that  $R_{\langle VT \rangle}$  is highest for the lowest-mass bins of injections; this is because the longer-duration waveforms in these bins ensure that the trigger distributions are closer to those from Gaussian noise, as they are not influenced so much by the glitches to which the single-detector search is more susceptible. At lower FARs, we see that the PyCBC-BBH  $R_{\langle VT \rangle}$  increases significantly, this is due to the background contamination by strong single-detector signals, which given the longer analysis time of the PyCBC-BBH chunks, affect more time (and therefore injections) than the shorter PyCBC-broad analyses.

Table 4 shows the  $\langle VT \rangle$  ratio  $R_{\langle VT \rangle}$  for the PyCBC-broad and PyCBC-BBH searches at the FAR threshold of two per year used in the results of section 4. We see the results in different mass bins, and averaged for all injections. For all injections in the PyCBC-broad search, we see an increase in sensitive  $\langle VT \rangle$  by a factor of  $1.11 \pm 0.01$ , and by a factor of  $1.16 \pm 0.01$  in the PyCBC-BBH search.



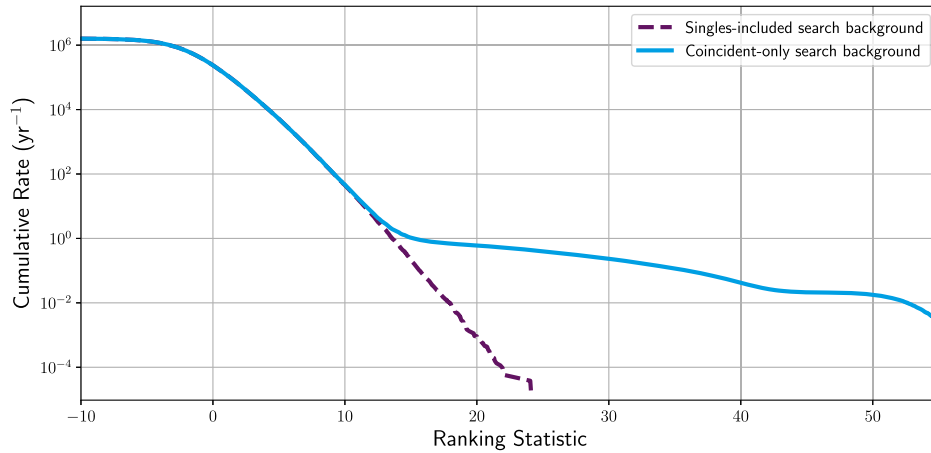
**Figure 9.** Recovered FAR values for injections from the singles-included analysis compared to the FAR of the same injections from the coincident-only analysis for the O3 PyCBC-broad search (left) and PyCBC-BBH search (right). The shape of each scatter point indicates whether the injection was recovered as a coincident or single-detector event in the singles-included analysis. Each scatter point is coloured according to the ratio of the second largest optimal SNR versus the maximum optimal SNR over the detectors. The dashed lines indicate the two-per-year cutoff used to list events in section 4, meaning that events in the lower-right section of the plot would be newly-found by the singles-included search. We see the effect of removing events from the coincident background, as well as the additional events recovered using the singles-included analysis. The panels at the side and top of the plot show events which were either completely missed by the search or above the FAR limits of the plot.

Figure 9 compares the recovered FAR of injections from the coincident-only analysis and singles-included analysis, for the PyCBC-broad search (left) and PyCBC-BBH search (right). We see that though there are a few injections found with better FAR in the coincident-only search, most of the injections are found with improved FAR by including the single-detector events. We also see the many newly-found events which were completely missed by the coincident search in the side panel.

A useful metric for assessing whether a signal is likely to be found is the optimal SNR. The optimal SNR is the SNR which would be recovered by an exact-match template in zero noise given the power spectral density (PSD) at the time of the event. The detector with the maximum optimal SNR would be decisive in a search involving single-detector events, however for coincident-only searches, the decisive optimal SNR would be the second largest SNR over the set of operating detectors. We use the ratio of these two optimal SNRs,  $\rho_{\text{second}}/\rho_{\text{max}}$ , in figure 9, and the biggest improvement is for events with the lowest value of this ratio, with darker scatter points. Events with a low value of this ratio are loud in one detector, but not in the next-loudest operating detector.

We additionally see arcs of injections with significantly improved recovered FAR; these come from the analysis chunks containing loud single-detector events. For PyCBC-broad, we see two arcs, corresponding to the chunk containing GW200105\_162426 and GW200112\_155838, and the one containing GW190630\_185205. The higher FARs in these arcs in the coincident-only search come from the presence of loud single-detector events in the background.

Figure 10 shows the LV background of the analysis chunk containing GW200105\_162426 and GW200112\_155838, where the triggers from the single-detector events in LIGO-Livingston match to random noise in Virgo, and form significant events in the background.



**Figure 10.** The LV ‘exclusive’ background of the analysis containing GW200105\_162426 and GW200112\_155838, from the coincident-only search (blue, solid) and from the singles-included search where the single-detector signals are removed from the background (purple, dashed).

The result is effectively a shelf in the false alarm rate, where events cannot be seen with false alarm rates below one per a few years unless they are very significant. We see in figure 9(a) that this means that events from PyCBC-broad are prevented from becoming more significant between FARs of around one per 10 years and one per ten thousand years, and that this effect begins at slightly higher FAR in PyCBC-BBH as in figure 9(b).

By injecting signals into the data, we have seen that by including single-detector events, we gain sensitivity of the search to all signals by a factor of  $1.11 \pm 0.01$  in the PyCBC-broad search and  $1.16 \pm 0.01$  in the PyCBC-BBH search, and by up to a factor of  $1.20 \pm 0.02$  for parts of the parameter space.

## 6. Conclusions

We present a method for extrapolating the false alarm rate of gravitational wave candidate events which do not form a coincidence for use in PyCBC searches. This method adapts the coincident ranking statistic for use with single-detector events, and fits the number of events with that ranking statistic or higher to a falling exponential model, extrapolating the number of higher-ranked events using this exponential fit. We have shown how this extrapolated FAR is used in the wider context of a GW search analysis.

We have assessed the ability of this method to digest single-detector events within a search. We recover seven single-detector events in O3 with a false alarm rate less than two per year; these events correspond to known events found by other pipelines. Only one glitch was identified as a marginally significant single-detector event during O3, showing that we have balanced the requirement to find events with the need to avoid non-Gaussian transient glitches. The total time–volume sensitivity of the PyCBC-broad search increases by a factor of  $1.11 \pm 0.01$  at a false alarm rate of one per two years compared to completely ignoring single-detector events, and the PyCBC-BBH search sensitivity increases by a factor of  $1.16 \pm 0.01$ .

## 7. Further work

We have presented methods to extrapolate the FAR for events in a single detector. This method is optimal in Gaussian noise, and although we have given a description of how we can mitigate non-Gaussian glitches in the search, they can always show up as outliers, and therefore be assigned low FARs, as evidenced by the reported significance of 200218\_201552.

Data being closer to Gaussian will then help this method even more than in a coincident search. Ongoing efforts to understand, reduce and mitigate glitches [40, 46–51], will therefore help to improve the ability of the search to find events.

Working out the inclusive FAR of events needs careful consideration, as we have so far only considered exclusive FAR. The inclusive FAR, as noted previously, is one per live time for the loudest-ranked single-detector event. The FAR is added as if we have seen an event in each combination at the ranking statistic of the event—this would result in the addition of one per live time for all active detectors, and all events would then have insignificant FARs. The solution may be to simply state that the inclusive FAR is one per live time of the detector it was found in; as the event we consider is, through clustering, the loudest event at that time, we do not have any higher-ranked events at that time. However this would not follow the same process we have used up to now, and would still not be useful for finding events with useful significance.

There are also ways we can make improvements to the ranking of events in order to help us to further discriminate signals from noise. We currently have no way to include any information from detectors which were active but did not trigger. For example, a high-SNR event in LIGO-Hanford is unlikely to be real if LIGO-Livingston was active but did not trigger. Inclusion of these terms may be possible as part of an extension to the prior histograms used to calculate  $p(\vec{\Omega}|S)$ , the probability of the given extrinsic parameters given that the trigger is a signal.

## Acknowledgments

The authors acknowledge the Science and Technology Funding Council (STFC) for funding through Grants ST/T000333/1 and ST/V005715/1. We are grateful to Frédérique Marion and Thomas Dent for help and comments on the manuscript. This material is based upon work supported by National Science Foundation (NSF)'s LIGO Laboratory which is a major facility fully funded by the NSF. This research has made use of data or software available from the Gravitational Wave Open Science Centre ([gw-openscience.org](http://gw-openscience.org)), a service of LIGO Laboratory, the LIGO Scientific Collaboration, the Virgo Collaboration, and KAGRA. LIGO Laboratory and Advanced LIGO are funded by the United States NSF as well as the STFC of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain. The construction and operation of KAGRA are funded by Ministry of Education, Culture, Sports, Science and Technology (MEXT), and Japan Society for the Promotion of Science (JSPS), National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea, Academia Sinica (AS) and the Ministry of Science and Technology (MoST) in Taiwan. The authors are grateful for the computational resources and data provided by the LIGO Laboratory

**Table A1.** Events with FAR unchanged between the coincident and singles-included analyses, and events with FAR > 2 in both the single-included and coincident searches, but which are reported in GWTC-2.1, GWTC-3, 3-OGC or 4-OGC. The horizontal line indicates the point where events above were recovered but unchanged, the ones below the line were not recovered in this analysis. Events are not included if the FAR is infinite in all searches presented here, which was the case for GW190424\_180648. Table columns and details are the same as in table 2.

Event	Instruments Active, triggered	Search including single-detector events				Coincident-only search			
		PyCBC-broad		PyCBC-BBH		PyCBC-broad		PyCBC-BBH	
		FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR
GW190707_093326	HL, HL	<1.0 × 10 <sup>-5</sup>	13.0	<2.0 × 10 <sup>-5</sup>	13.0	<1.0 × 10 <sup>-5</sup>	13.0	<2.0 × 10 <sup>-5</sup>	13.0
GW200225_060421	HL, HL	<1.2 × 10 <sup>-5</sup>	12.3	<1.5 × 10 <sup>-5</sup>	12.3	<1.2 × 10 <sup>-5</sup>	12.3	<1.5 × 10 <sup>-5</sup>	12.3
GW191204_171526	HL, HL	<1.5 × 10 <sup>-5</sup>	16.9	<1.3 × 10 <sup>-5</sup>	16.9	<1.5 × 10 <sup>-5</sup>	16.9	<1.3 × 10 <sup>-5</sup>	16.9
GW191216_213338	HV, HV	<1.4 × 10 <sup>-5</sup>	18.3	<1.4 × 10 <sup>-5</sup>	18.3	<1.4 × 10 <sup>-5</sup>	18.3	<1.4 × 10 <sup>-5</sup>	18.3
GW190521_074359	HL, HL	<1.8 × 10 <sup>-5</sup>	24.1	<2.5 × 10 <sup>-5</sup>	24.1	<1.8 × 10 <sup>-5</sup>	24.1	<2.5 × 10 <sup>-5</sup>	24.1
GW200129_065458	HLV, HV	<2.6 × 10 <sup>-5</sup>	16.3	<1.9 × 10 <sup>-5</sup>	16.2	<2.6 × 10 <sup>-5</sup>	16.3	<1.9 × 10 <sup>-5</sup>	16.2
GW191129_134029	HL, HL	<2.7 × 10 <sup>-5</sup>	12.9	<2.6 × 10 <sup>-5</sup>	12.9	<2.7 × 10 <sup>-5</sup>	12.9	<2.6 × 10 <sup>-5</sup>	12.9
GW200311_115853	HLV, HLV	<7.3 × 10 <sup>-5</sup>	17.0	<8.4 × 10 <sup>-5</sup>	17.4	<7.3 × 10 <sup>-5</sup>	17.0	<8.4 × 10 <sup>-5</sup>	17.4
GW190512_180714	HLV, HL	<7.5 × 10 <sup>-5</sup>	12.4	<1.2 × 10 <sup>-4</sup>	12.4	<7.5 × 10 <sup>-5</sup>	12.4	<1.2 × 10 <sup>-4</sup>	12.4
GW190828_063405	HLV, HL	<9.0 × 10 <sup>-5</sup>	13.9	<7.6 × 10 <sup>-5</sup>	15.9	<9.0 × 10 <sup>-5</sup>	13.9	<7.6 × 10 <sup>-5</sup>	15.9
GW190924_021846	HLV, HL	<8.7 × 10 <sup>-5</sup>	12.5	<7.6 × 10 <sup>-5</sup>	12.5	<8.7 × 10 <sup>-5</sup>	12.5	<7.6 × 10 <sup>-5</sup>	12.5
GW200224_222234	HLV, HLV	<8.6 × 10 <sup>-5</sup>	19.2	<8.4 × 10 <sup>-5</sup>	18.6	<8.6 × 10 <sup>-5</sup>	19.2	<8.4 × 10 <sup>-5</sup>	18.6
GW190519_153544	HLV, HL	<1.1 × 10 <sup>-4</sup>	13.2	<1.2 × 10 <sup>-4</sup>	13.2	<1.1 × 10 <sup>-4</sup>	13.2	<1.2 × 10 <sup>-4</sup>	13.2
GW190412	HLV, HL	<1.2 × 10 <sup>-4</sup>	17.4	<1.3 × 10 <sup>-4</sup>	18.0	<1.2 × 10 <sup>-4</sup>	17.4	<1.3 × 10 <sup>-4</sup>	18.6
GW190408_181802	HLV, HL	2.7 × 10 <sup>-4</sup>	13.2	<1.3 × 10 <sup>-4</sup>	13.7	<2.7 × 10 <sup>-4</sup>	13.2	<1.3 × 10 <sup>-4</sup>	13.7
GW200210_092254	HLV, HL	16	9.0	2.2	9.0	17	9.0	2.4	9.0
GW190916_200658	HLV, HL	8.5 × 10 <sup>3</sup>	8.0	3.6	8.0	8.4 × 10 <sup>3</sup>	8.0	4.1	8.0
GW191219_163120	HLV, HL	3.9	9.0	—	—	3.9	9.0	—	—
GW200216_220804	HLV, HL	960	9.1	4.3	9.1	960	9.1	4.6	9.1
GW190929_012149	HLV, HL	120	9.4	5.1	8.6	120	9.4	5.8	8.6
GW191224_043228	HLV, HL	—	—	7.3	8.7	—	—	8.6	8.7
GW190527_092055	HL, HL	—	—	7.9	8.4	—	—	8.3	8.4
GW200306_093714	HL, HL	3.4 × 10 <sup>3</sup>	7.9	10	8.0	3.4 × 10 <sup>3</sup>	7.9	11	8.0

(continued on next page)

Table A1. Continued.

Event	Instruments Active, triggered	Search including single-detector events				Coincident-only search			
		PyCBC-broad		PyCBC-BBH		PyCBC-broad		PyCBC-BBH	
		FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR	FAR (years <sup>-1</sup> )	Network SNR
GW200208_222617	HLV, HL	$1.6 \times 10^6$	†7.0	12	7.9	—	—	13	7.9
GW200214_223307	HLV, HL	$7.8 \times 10^3$	7.9	14	7.6	$7.7 \times 10^3$	7.9	14	7.6
GW200220_124850	HL, HL	—	—	14	7.9	—	—	14	7.9
GW191201_054144	HV, HV	30	9.1	730	8.4	28	9.1	610	8.4
GW190427_180650	HLV, HL	$1.4 \times 10^3$	8.9	40	9.0	$1.4 \times 10^3$	8.9	42	9.0
GW190926_050336	HLV, HL	$6.4 \times 10^4$	8.0	40	7.8	$6.2 \times 10^4$	8.0	42	7.8
GW190403_051519	HLV, HL	—	—	42	8.0	—	—	45	8.0
GW200305_084739	HLV, HL	—	—	55	7.7	—	—	58	7.7
GW200210_005122	HLV, HL	430	8.2	57	8.2	420	8.2	58	8.2
GW200322_091133	HLV, HL	$1.4 \times 10^3$	8.1	62	7.8	$1.4 \times 10^3$	8.1	66	7.8
GW200106_134123	HLV, HL	$7.7 \times 10^4$	7.6	84	7.6	$7.4 \times 10^4$	7.6	87	7.6
GW190426_190642	HLV, HL	$1.2 \times 10^4$	9.9	100	9.6	$1.2 \times 10^4$	9.9	110	9.6
GW200220_061928	HLV, HL	—	—	160	7.6	—	—	160	7.6
GW200318_191337	HLV, HL	—	—	250	7.6	—	—	250	7.6
GW190805_105432	HL, HL	—	—	360	7.8	—	—	360	7.8
GW191113_071753	HLV, HL	$9.9 \times 10^3$	8.4	730	8.6	$1.1 \times 10^4$	8.4	730	8.6
GW200129_114245	HLV, HL	$1.2 \times 10^5$	7.6	790	7.2	$1.1 \times 10^5$	7.6	790	7.2
GW200201_203549	HLV, HL	$1.0 \times 10^3$	8.3	—	—	$1.0 \times 10^3$	8.3	790	7.2
GW190917_114630	HLV, HL	$1.5 \times 10^3$	8.4	—	—	$1.5 \times 10^3$	8.4	—	—



**Table A2.** Template parameters for events which were found in the singles-included analysis but not by the coincident-only analysis.

Event	PyCBC-broad				PyCBC-BBH			
	$\mathcal{M}$	$\eta$	$\chi_{\text{eff}}$	Dur (s)	$\mathcal{M}$	$\eta$	$\chi_{\text{eff}}$	Dur (s)
GW200112_155838	44.41	0.25	0.52	0.24	29.25	0.25	-0.24	0.36
GW190425	1.49	0.25	0.01	67.24	—	—	—	—
GW200105_162426	3.61	0.17	-0.15	15.63	—	—	—	—
GW190708_232457	15.60	0.20	0.20	1.34	15.60	0.20	0.20	1.34
GW190620_030421	27.97	0.12	0.15	0.41	41.61	0.20	0.33	0.24
GW200302_015811	9.80	0.06	-0.80	2.24	24.88	0.19	-0.30	0.47
GW190910_112807	17.33	0.15	-0.95	0.74	40.48	0.24	-0.05	0.20
200218_201552	21.27	0.13	-0.25	0.60	67.48	0.23	0.94	0.18

and supported by National Science Foundation Grant Nos. PHY-0757058 and PHY-0823459. The authors also acknowledge the use of the IUCAA LDG cluster, Sarathi, for computational/numerical work. The PyCBC offline search software used was based on PyCBC release 1.18.1 [30]. The PyCBC search software is built upon LALSuite [52], numpy [53], SciPy [54], Astropy [55] and Pegasus [32].

### Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

### Appendix A. Further results

In addition to the results shown in section 4, we present in table A1 the FAR and network SNR for events from GWTC-2.1, GWTC-3, 3-OGC and 4-OGC which were not included in tables 2 and 3. These were not included in tables 2 or 3 either as they did not meet the FAR threshold in these analyses, or as the FAR did not change between the analyses.

It should be noted that the event GW200210\_092254 is indicated as an event not previously found in table A1, but this was found by both PyCBC searches in the GWTC-3 analysis with  $p_{\text{astro}} > 0.5$ , the criteria for inclusion in that work.

We also provide template parameters for the events found by the singles-included search but not by the coincident search in table A2.

The parameters in table A2 are not intended as parameter estimation analysis for the events; as all events here are given in GWTC-2.1 or GWTC-3, parameter estimation is given in those papers. Parameter estimation of intrinsic parameters for events where only a single detector was operating should be largely unaffected by observation in only one detector, however extrinsic parameters such as the localisation of the source on the sky would be significantly affected [56].

### ORCID iDs

Gareth S Cabourn Davies  <https://orcid.org/0000-0002-4289-3439>

Ian W Harry  <https://orcid.org/0000-0002-5304-9372>

## References

- [1] Aasi J *et al* (LIGO Scientific Collaboration) 2015 *Class. Quantum Grav.* **32** 074001
- [2] Acernese F *et al* (Virgo Collaboration) 2015 *Class. Quantum Grav.* **32** 024001
- [3] Dal Canton T, Nitz A H, Gadre B, Cabourn Davies G S, Villa-Ortega V, Dent T, Harry I and Xiao L 2021 *Astrophys. J.* **923** 254
- [4] Abbott B P *et al* (LIGO Scientific, Virgo, Fermi GBM, INTEGRAL, IceCube, AstroSat Cadmium Zinc Telluride Imager Team, IPN, Insight-Hxmt, ANTARES, Swift, AGILE Team, 1M2H Team, Dark Energy Camera GW-EM, DES, DLT40, GRAWITA, Fermi-LAT, ATCA, ASKAP, Las Cumbres Observatory Group, OzGrav, DWF (Deeper Wider Faster Program), AST3, CAASTRO, VINROUGE, MASTER, J-GEM, GROWTH, JAGWAR, CaltechNRAO, TTU-NRAO, NuSTAR, Pan-STARRS, MAXI Team, TZAC Consortium, KU, Nordic Optical Telescope, ePESSTO, GROND, Texas Tech University, SALT Group, TOROS, BOOTES, MWA, CALET, IKI-GW Follow-up, H.E.S.S., LOFAR, LWA, HAWC, Pierre Auger, ALMA, Euro VLBI Team, Pi of Sky, Chandra Team at McGill University, DFN, ATLAS Telescopes, High Time Resolution Universe Survey, RIMAS, RATIR, SKA South Africa/MeerKAT) 2017 *Astrophys. J. Lett.* **848** L12
- [5] Abbott R *et al* (LIGO Scientific, Virgo) 2021 *Phys. Rev. X* **11** 021053
- [6] Abbott R *et al* (LIGO Scientific, VIRGO, KAGRA) 2021 arXiv:2111.03606
- [7] Nitz A H, Capano C D, Kumar S, Wang Y F, Kastha S, Schäfer M, Dhurkunde R and Cabero M 2021 arXiv:2105.09151
- [8] Nitz A H, Kumar S, Wang Y F, Kastha S, Wu S, Schäfer M, Dhurkunde R and Capano C D 2021 arXiv:2112.06878
- [9] Biwer C M, Capano C D, De S, Cabero M, Brown D A, Nitz A H and Raymond V 2019 *Publ. Astron. Soc. Pac.* **131** 024503
- [10] Berry C P L *et al* 2015 *Astrophys. J.* **804** 114
- [11] Cutler C and Flanagan E E 1994 *Phys. Rev. D* **49** 2658–97
- [12] Roulet J, Venumadhav T, Zackay B, Dai L and Zaldarriaga M 2020 *Phys. Rev. D* **102** 123022
- [13] Abbott B P *et al* (LIGO Scientific, Virgo) 2019 *Astrophys. J. Lett.* **882** L24
- [14] Abbott R *et al* (LIGO Scientific, VIRGO, KAGRA) 2021 arXiv:2111.03634
- [15] Abbott R *et al* (LIGO Scientific, VIRGO, KAGRA) 2021 arXiv:2112.06861
- [16] Davies G S, Dent T, Tápai M, Harry I, McIsaac C and Nitz A H 2020 *Phys. Rev. D* **102** 022004
- [17] Nitz A H, Dent T, Dal Canton T, Fairhurst S and Brown D A 2017 *Astrophys. J.* **849** 118
- [18] Usman S A *et al* 2016 *Class. Quantum Grav.* **33** 215004
- [19] Babak S *et al* 2013 *Phys. Rev. D* **87** 024033
- [20] Messick C *et al* 2017 *Phys. Rev. D* **95** 042001
- [21] Sachdev S *et al* 2019 arXiv:1901.08580
- [22] Aubin F *et al* 2021 *Class. Quantum Grav.* **38** 095004
- [23] Adams T, Buskulic D, Germain V, Guidi G M, Marion F, Montani M, Mours B, Piergiorganni F and Wang G 2016 *Class. Quantum Grav.* **33** 175012
- [24] Hooper S, Chung S K, Luan J, Blair D, Chen Y and Wen L 2012 *Phys. Rev. D* **86** 024012
- [25] Venumadhav T, Zackay B, Roulet J, Dai L and Zaldarriaga M 2020 *Phys. Rev. D* **101** 083030
- [26] Roy S, Sengupta A S and Thakor N 2017 *Phys. Rev. D* **95** 104045
- [27] Roy S, Sengupta A S and Ajith P 2019 *Phys. Rev. D* **99** 024048
- [28] Abbott B P *et al* (LIGO Scientific, Virgo) 2016 *Phys. Rev. D* **93** 122004  
Abbott B P *et al* (LIGO Scientific, Virgo) *Phys. Rev. D* **94** 069903 (addendum)
- [29] Klimenko S *et al* 2016 *Phys. Rev. D* **93** 042004
- [30] Nitz A *et al* 2021 gwastro/pycbc: Pycbc release 1.18.1 <https://doi.org/10.5281/zenodo.4849433>
- [31] Weitzel D, Bockelman B, Brown D A, Couvares P, Würthwein F and Fajardo Hernandez E 2017 Data access for LIGO on the OSG *Proc. Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact PEARC17* (New York: ACM) 24:pp 1–6
- [32] Deelman E *et al* 2015 *Future Gener. Comput. Syst.* **46** 17–35
- [33] Abbott R *et al* (LIGO Scientific, VIRGO) 2021 arXiv:2108.01045
- [34] Nitz A H, Dent T, Davies G S and Harry I 2020 *Astrophys. J.* **897** 169
- [35] Andres N *et al* 2022 *Class. Quantum Grav.* **39** 055002
- [36] Callister T A, Kanner J B, Massinger T J, Dhurandhar S and Weinstein A J 2017 *Class. Quantum Grav.* **34** 155007
- [37] Cabero M *et al* 2019 *Class. Quantum Grav.* **36** 15
- [38] Allen B 2005 *Phys. Rev. D* **71** 062001

- [39] Brown D A 2004 Searching for gravitational radiation from binary black hole MACHOs in the galactic halo *PhD Thesis* University of Wisconsin-Milwaukee
- [40] Soni S *et al* (LIGO) 2020 *Class. Quantum Grav.* **38** 025016
- [41] Accadia T *et al* 2010 *Class. Quantum Grav.* **27** 194011
- [42] Peters P 1964 *Phys. Rev.* **136** B1224–32
- [43] Blanchet L, Damour T, Iyer B R, Will C M and Wiseman A G 1995 *Phys. Rev. Lett.* **74** 3515–8
- [44] Dent T and Veitch J 2014 *Phys. Rev. D* **89** 062002
- [45] Nitz A H *et al* 2019 *Astrophys. J.* **891** 123
- [46] Davis D *et al* (LIGO) 2021 *Class. Quantum Grav.* **38** 135014
- [47] Robinet F, Arnaud N, Leroy N, Lundgren A, Macleod D and McIver J 2020 *SoftwareX* **12** 100620
- [48] Soni S *et al* 2021 *Class. Quantum Grav.* **38** 195016
- [49] Merritt J, Farr B, Hur R, Edelman B and Doctor Z 2021 *Phys. Rev. D* **104** 102004
- [50] Godwin P *et al* 2020 arXiv:2010.15282
- [51] McIsaac C and Harry I 2022 arXiv:2203.03449
- [52] LIGO Scientific Collaboration 2018 *LIGO Algorithm Library—LALSuite Free Software (GPL)*
- [53] van der Walt S, Colbert S C and Varoquaux G 2011 *Comput. Sci. Eng.* **13** 22–30
- [54] Virtanen P *et al* 2020 *Nat. Methods* **17** 261
- [55] Price-Whelan A M *et al* 2018 *Astron. J.* **156** 123
- [56] Fairhurst S 2011 *Class. Quantum Grav.* **28** 105021