

Method for removing signal contamination during significance estimation of a GstLAL analysis

Prathamesh Joshi^{1,2,*}, Leo Tsukada,^{1,2} and Chad Hanna^{1,2,3,4}

¹*Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802, USA*

²*Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, Pennsylvania 16802, USA*

³*Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, Pennsylvania 16802, USA*

⁴*Institute for Computational and Data Sciences, The Pennsylvania State University, University Park, Pennsylvania 16802, USA*



(Received 30 May 2023; accepted 18 September 2023; published 17 October 2023)

To evaluate the probability of a gravitational-wave candidate originating from noise, GstLAL collects noise statistics from the data it analyzes. Gravitational-wave signals of astrophysical origin get added to the noise statistics, harming the sensitivity of the search. We present the background filter, a novel tool to prevent this by removing noise statistics that were collected from gravitational-wave candidates. To demonstrate its efficacy, we analyze one week of LIGO and Virgo O3 data, and show that it improves the sensitivity of the analysis by 20%–40% in the high mass region, in the presence of 868 simulated gravitational-wave signals. With the upcoming fourth observing run of LIGO, Virgo, and KAGRA expected to yield a high rate of gravitational-wave detections, we expect the background filter to be an important tool for increasing the sensitivity of a GstLAL analysis.

DOI: [10.1103/PhysRevD.108.084032](https://doi.org/10.1103/PhysRevD.108.084032)

I. INTRODUCTION

The Laser Interferometer Gravitational-wave Observatory (LIGO) [1] and Virgo [2] Collaborations have revolutionized the field of gravitational-wave (GW) astronomy by detecting black hole and neutron star mergers [3–6]. The detections have allowed us to observe the Universe in new ways and have opened up new avenues of scientific inquiry. [7–10] The GstLAL GW search pipeline [11–14] (referred to as GstLAL hereafter) has been a significant contributor to this field. In particular, GstLAL’s ability to detect signals in low latency [15] has facilitated multimessenger observations [16].

GstLAL is a GW search pipeline that can process data from ground-based GW detectors, such as the Hanford and Livingston LIGO detectors, the Virgo detector, and the KAGRA detector [17], in near real time. It makes use of time-domain matched filtering to enable the detection of signals in noise-dominated data. It uses a likelihood ratio (LR) [18–20] as a ranking statistic for assigning significance to detections. GstLAL divides its template bank [21,22] into different “template bins” to reduce the computational cost of the analysis, and analyzes each one separately. Some of these techniques are also used by other search pipelines, such as PyCBC [23–25], MBTA [26,27], SPIIR [28,29], and IAS [30,31].

The fourth observing run of the LIGO Scientific, Virgo, and KAGRA Collaborations (O4) is set to begin in May 2023 [32] and promises to provide improved detector sensitivity. GstLAL will continue to play an essential role in the detection of new GW candidates. As such, it is necessary to keep refining the analysis pipeline to reap the benefits of improved detector sensitivity to detect even more and new types of candidates. The background filter is one such new feature to this end.

This paper is structured as follows. In Sec. II, we introduce the LR used by GstLAL, in particular the $\rho - \xi^2$ histograms that GstLAL uses to evaluate one term of the likelihood ratio, and how the presence of GW signals in the data can cause “contamination” of these histograms. In Sec. III, we describe how the background filter works, and how it removes this contamination. Finally, in Sec. IV, we describe the analyses we performed to evaluate the performance of the background filter, and the impact it has on the sensitivity of a GstLAL analysis.

II. SIGNAL CONTAMINATION

A. Likelihood ratio

GstLAL is a matched-filtering based GW search pipeline which uses a likelihood ratio statistic to rank GW candidates [18,19]. The LR is defined as

*prathamesh.joshi@ligo.org

$$\mathcal{L} = \frac{P(\vec{O}, \vec{\rho}, \vec{\xi}^2, \vec{t}, \vec{\phi}, \vec{\theta} | \mathcal{H}_s)}{P(\vec{O}, \vec{\rho}, \vec{\xi}^2, \vec{t}, \vec{\phi}, \vec{\theta} | \mathcal{H}_n)}, \quad (1)$$

where the numerator is the probability of obtaining a GW candidate with parameters $(\vec{O}, \vec{\rho}, \vec{\xi}^2, \vec{t}, \vec{\phi}, \vec{\theta})$ under the signal hypothesis ($\mathcal{H}_s$) and the denominator is the probability of obtaining the same candidate under the noise hypothesis ($\mathcal{H}_n$). $\vec{O}$ is the subset of GW detectors that the

candidate was found in, $\vec{\rho}$ is the set of matched-filtering signal-to-noise-ratios (SNRs) of those detectors, $\vec{\xi}^2$ is the set of ξ^2 -signal-based-veto values, $\vec{t}, \vec{\phi}$ are the times and phases with which the candidate was found in the detectors, and $\vec{\theta}$ is the template that recovered the candidate, which also represents a set of intrinsic parameters (masses and spins).

The LR can be factorized as

$$\mathcal{L} = \frac{P(\vec{\theta} | \mathcal{H}_s) \times P(t_{\text{ref}}, \phi_{\text{ref}} | \vec{\theta}, \mathcal{H}_s) \times P(\vec{O} | t_{\text{ref}}, \mathcal{H}_s) \times P(\vec{\rho}, \vec{\Delta t}, \vec{\Delta \phi} | \vec{O}, t_{\text{ref}}, \mathcal{H}_s) \times P(\vec{\xi}^2 | \vec{\rho}, \vec{\theta}, \mathcal{H}_s)}{P(t_{\text{ref}}, \vec{\theta} | \mathcal{H}_n) \times P(\vec{O} | t_{\text{ref}}, \vec{\theta}, \mathcal{H}_n) \times P(\vec{\Delta t}, \vec{\phi} | \vec{O}, \mathcal{H}_n) \times P(\vec{\rho}, \vec{\xi}^2 | t_{\text{ref}}, \vec{\theta}, \mathcal{H}_n)} \quad (2)$$

For a comprehensive explanation of Eq. (2) and every individual term in the LR, readers are referred to [18]. For the purpose of this paper, we are only concerned with the last term in the denominator, $P(\vec{\rho}, \vec{\xi}^2 | t_{\text{ref}}, \vec{\theta}, \mathcal{H}_n)$ (hereby referred to as the $\rho - \xi^2$ noise LR term).

B. The $\rho - \xi^2$ histograms

The $\rho - \xi^2$ noise LR term is calculated in a data-driven way. GstLAL creates a histogram for each detector and template bin in $\rho - \xi^2$ space, called $\rho - \xi^2$ background histograms, and populates it with the (ρ, ξ^2) values of noise events found in that template bin during the analysis. Then, the $\rho - \xi^2$ noise LR term can be calculated by evaluating the probability density function represented by the histograms at the relevant $(\vec{\rho}, \vec{\xi}^2)$ value.

Since the $\rho - \xi^2$ noise LR term assumes the noise hypothesis, we need to populate the histograms with events originating only from noise, as compared to events originating from GW candidates. To a large degree, this is achieved by requiring those events to be recovered only in one detector (called a single detector or single event in contrast to a coincident event) during a time when more than one detector was producing data (called coincident time in contrast to single time). This is because we expect GW signals to be correlated across detectors, but not noise events.

Despite this, GW signals can sometimes enter the $\rho - \xi^2$ histograms. The reason might be astrophysical in origin, e.g., the GW source is located in the blind spot of all but one detector, or it might be terrestrial, e.g., only one detector is sensitive enough to pick up the GW signal. In addition, GW signals that are recovered as coincident events in one template bin sometimes also get recovered as a single event, with a lower ρ and LR in other neighboring template bins, which do not contain templates with high ρ for that GW signal. We say a bin has a good match with a GW signal if it has templates with a high ρ for that GW signal, and that it has a bad match otherwise. GW events

being recovered as coincident events in one bin and as single events in others is demonstrated in Fig. 1 for GW200129_065458, a known GW candidate reported in GWTC-3 [6]. The candidate is recovered as a coincident event in bin 818, with which it has the best match. It is also recovered in bin 838 as a Livingston single with a lower ρ , since its match with that bin is not as good. As a result, the candidate will be added to the background histogram of bin 838. Gravitational wave signals entering the background histograms are commonly called the signal contamination of the $\rho - \xi^2$ background histograms. The contamination caused by GW200129_065458 in the background histogram of bin 838 is shown in Fig. 2. Since the GW signal gets added to the background histogram, it occupies a region in $\rho - \xi^2$ space typical of signals, but not of noise. As a result, we see a protrusion to the histogram, which is generally how signal contamination manifests visually.

Signal contamination can result in the $\rho - \xi^2$ histograms not accurately reflecting the noise characteristics of the data, and as a result, the $\rho - \xi^2$ noise LR term will not be calculated correctly. In general, it can cause the $\rho - \xi^2$ noise LR term for GW candidates to be evaluated higher than its true value, leading to lower LR values of candidates. In short, signal contamination can lower the sensitivity of the GW search.

III. REMOVING CONTAMINATION WITH THE BACKGROUND FILTER

To prevent any loss in sensitivity due to signal contamination, we need to selectively remove the events in the $\rho - \xi^2$ background histograms which originate from GW signals. The background filter is a way to track the background in a time-dependent fashion so that we only use events from times not corresponding to GW events to populate the background histograms. In this paper, we will describe the working of the background filter when GstLAL is running in the low-latency online mode, in

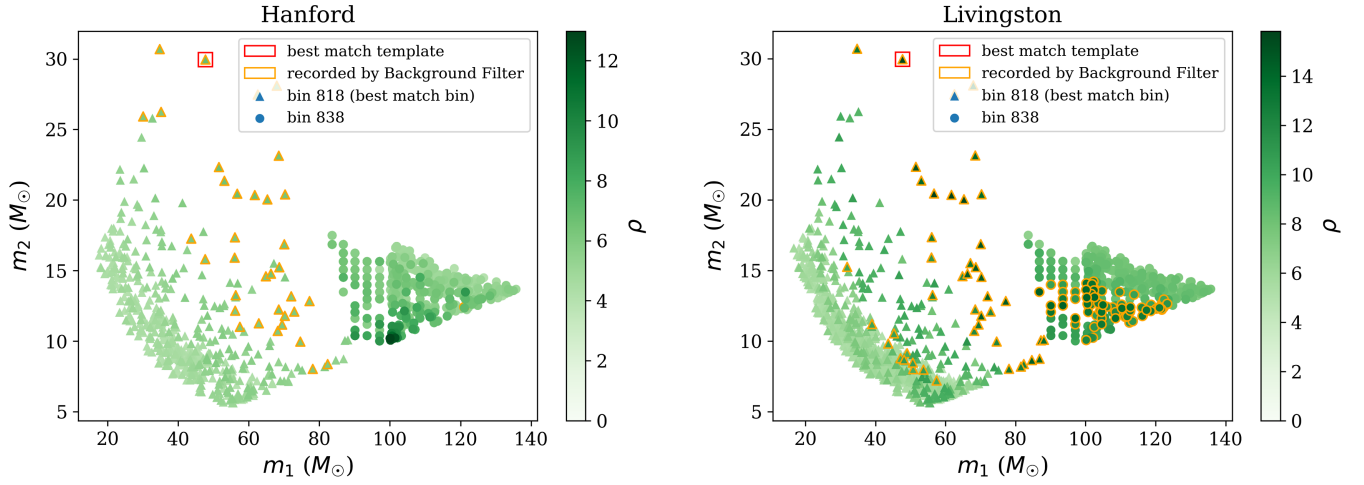


FIG. 1. An example of an event (GW200129_065458) having templates with high match in multiple template bins. Bin 818 has the best match with the GW candidate, and recovers it in both Hanford and Livingston as a coincidence. Bin 838 has a lower match than bin 818, causing it to recover the candidate as a Livingston single. This will lead to the candidate being added to the $\rho - \xi^2$ background histogram of bin 838, causing signal contamination for bin 838. This is shown in Fig. 2. The events passing the ρ and ξ^2 constraints, and hence recorded by the background filter, are outlined in orange.

which data is analyzed and results are produced in near real time [33].

A. Recording events

The strategy of the background filter is to record the events that are likely to have originated from GW signals, and then after verification by the user, subtract them from the background histograms. To associate events with a GW candidate, we need to record the time at which they were found in the data, apart from their $\vec{\rho}$ and ξ^2 values. This increases the dimensionality of the parameters we need to

store, and thus could potentially impact the memory and storage used during analysis. To prevent this, we record events only if they pass certain constraints placed on their ρ , ξ^2 , and time parameters.

The ρ and ξ^2 constraints take the form of a bounding box in $\rho - \xi^2$ space, defined by $\rho > 6$ and $\xi^2/\rho^2 < 0.04$. Qualitatively, ξ^2 represents how well the data fits the template, with large values of ξ^2 meaning the data is dissimilar to the template. Since, in general, noise events will not fit the template well, they generally have ξ^2/ρ^2 values that are greater than those of signals. As a result, we only expect events originating from GW signals to fall inside the bounding box. This is shown in Fig. 1, where most of the high ρ events caused by GW200129_065458 pass the ρ and ξ^2 constraints, and are recorded by the background filter. The ρ and ξ^2 constraints are shown on top of a background histogram in Fig. 3.

The time constraint makes use of the GstLAL online analysis' ability to process data, generate events, assign LR, and upload them to the Gravitational Wave Candidate Event Database (GraceDB) [34] in near real time. A GW signal can create multiple contaminating events across template bins. Only a small subset gets uploaded to GraceDB, since the events are aggregated within some time windows across bins before uploading [33], and the remaining contaminating events lie both before and after the uploaded events in time. With this in mind, and in order to account for processing delays during a GstLAL online analysis, the background filter keeps a temporary record of events passing the ρ and ξ^2 constraints, which occurred in the last 5000 s. When an event is uploaded to GraceDB, the events in the 10 s window around it are found from the temporary record of the last 5000s, and are then recorded by the background filter permanently.

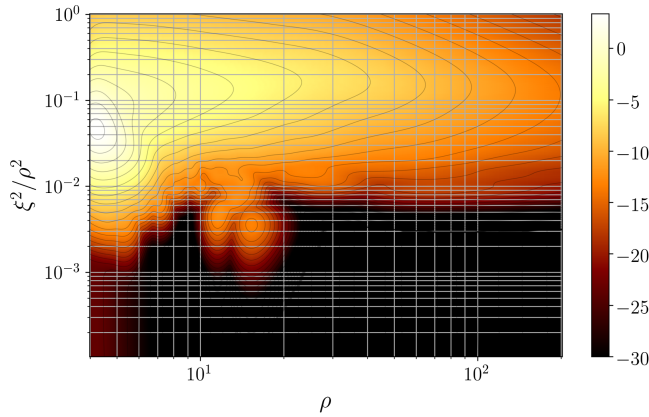


FIG. 2. An example of signal contamination in a $\rho - \xi^2$ histogram for Livingston. The contamination can be seen as a protrusion to the histogram at $(\rho, \xi^2/\rho^2) \sim (15, 0.004)$, a region usually occupied exclusively by GW signals. This contamination was caused by GW200129_065458 being recovered as a single event in this template bin, which is not the best match bin for that GW candidate, as demonstrated in Fig. 1. Note that kernel smoothing has been applied to this histogram.

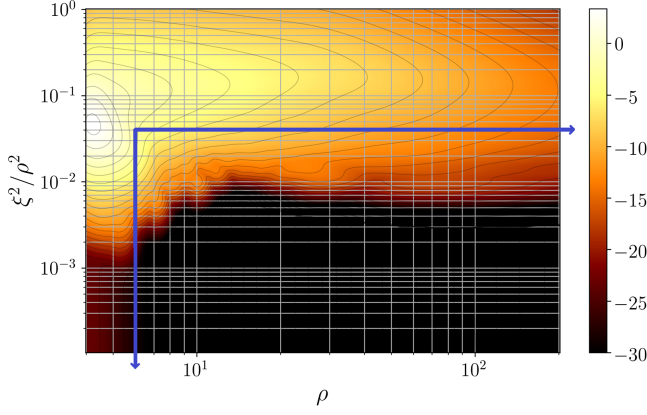


FIG. 3. The ρ and ξ^2 constraints for recording events. The bottom right area bounded by the blue lines is the area in which the background filter records events. If the events also pass the time constraint, the user can choose to remove them from the $\rho - \xi^2$ histogram. The result of doing so, to remove the contamination caused by GW200129_065458, is also shown. The same histogram, without using the background filter, and hence with contamination, is shown in Fig. 2. Note that kernel smoothing has been applied to this histogram.

The threshold for uploading an event to GraceDB differs among different GstLAL analyses, but it is often set to false alarm rate (FAR) < 1 per hour. That means all events recovered as a single event during coincident time, with $\rho > 6$, $\xi^2/\rho^2 < 0.04$, and falling in a 10 s interval around an event with FAR < 1 per hour, are recorded by the background filter.

The ρ and ξ^2 constraints, and the time constraint work together to ensure that only events originating from GW signals are recorded by the background filter in most cases. As a result, very few events are recorded by the background filter, in comparison to the number of events in the background histograms. This ensures that adding the background filter to a GstLAL analysis does not affect its memory or disk usage significantly. The choice of these constraints and their impact on the performance of a GstLAL analysis are discussed in Appendix A.

B. Removing contamination

As explained in Sec. II, since the $\rho - \xi^2$ noise LR term is calculated by evaluating the probability density function represented by the background histograms at the relevant $(\vec{\rho}, \vec{\xi}^2)$ value, we need the background histograms to accurately reflect the detector noise characteristics for that template bin. As much as possible, we need to take care not to let events originating from signals enter the background histograms. In addition, we must also make sure that events originating from noise are not removed from the background histograms by the background filter. In most cases, the ρ and ξ^2 constraints along with the time constraint are sufficient to ensure only events originating from signals are recorded by the background filter.

However, in rare cases, such as when the GstLAL analysis uploads a false positive to GraceDB (also called a “retraction”), these measures might not be enough. Out of an abundance of caution, we leave the decision of which events to remove from the background histograms to the user. At any point during a GstLAL online analysis, the user can choose to inform the analysis which events they are confident are GW candidates. The message is communicated to the analysis in real time using HTTP request methods, with the help of the PYTHON BOTTLE module [35]. Then, out of all the events that had been recorded by the background filter previously, it will subtract those which fall within a 10 s window of the given candidate from the background histograms. Thus, any contamination that candidate could have potentially caused is removed, and the LR of all future events is evaluated using the modified $\rho - \xi^2$ background histograms.

To summarize, for the background filter to subtract an event from the background histograms, three conditions need to be met:

- (1) The event needs to pass the ρ and ξ^2 constraints.
- (2) The event needs to pass the time constraint.
- (3) The user needs to inform the GstLAL analysis that there was a GW candidate nearby the event in time.

The first two conditions are sufficient for the background filter to record an event and save it to disk, but all three are necessary (and sufficient) for the event to be subtracted from the background histograms. For O4, we have decided that the criteria for informing the analysis of a GW candidate, and hence for removing the background events associated with it are

- (1) The event should have a FAR $\leq$ one per 5 months, which is the public alert threshold for significant events [36].
- (2) The event should not be a retraction, i.e., the GraceDB event should not have the “ADVNO” label applied to it.

This criteria was chosen keeping in mind the tradeoff between removing contamination from as many GW candidates as possible, and not removing noise events from the background. This choice and its effects on the sensitivity improvement caused by the background filter are discussed in Appendix B.

In Fig. 2, signal contamination caused by GW200129_065458 is shown. The same $\rho - \xi^2$ histogram, but with the background filter used to remove that contamination, is shown in Fig. 3.

IV. RESULTS

A. Analysis methods

To test the effect of signal contamination on the sensitivity of a GstLAL analysis, and the ability of the background filter to remove the contamination, we analyze a week of O3 data [37], from Apr 18 2019 16:46 UTC to

Apr 26 2019 17:14 UTC, in three different ways. First, we perform a control run without any GW signals. Next, to simulate the effect of GW signals, we add “blind injections.” The concept of blind injections and the set of blind injections that we used are explained in the following subsection. Finally, we perform a “rerank” with the background filter enabled, in which LRs are recomputed and significance assignment is done again, but the filtering stage of the GstLAL analysis is taken from the blind injection analysis, since the ρ and ξ^2 values of analyzed events are not affected by the background filter, only the LRs and the FARs are. Hence, the rerank corresponds to the case with blind injections present and the background filter being used.

As noted in Sec. III, our policy during O4 for selecting GW candidates to inform the analysis about, is that the candidate should have a FAR $\leq$ one per 5 months. This is also what is done during the rerank, and hence, the background filter is only applied to events passing the one per 5 months threshold.

This chunk of data contains two known GW candidates reported in GWTC-2.1 [5] and elsewhere [6,38], GW190421_213856 and GW190425. However, since we use the background filter only on the times of the blind injections, any contamination and subsequent loss in sensitivity caused by either of the two candidates will be present in all three analyses that we perform, and will not affect the evaluation of the performance of the background filter.

B. Simulation set

Blind injections are simulated GW signals that are added to the data which we analyze and collect background events from (in contrast to regular injections, from which we do not collect background events). We use a set of 868 blind injections distributed across the binary black hole (BBH), binary neutron star (BNS), neutron star-black hole (NSBH), and intermediate-mass black hole (IMBH) parameter spaces. The blind injection set comprises three subsets, a BNS subset, a BBH subset, and a broad subset, with the BNS subset containing half of the total blind injections, and the BBH and broad subsets containing a quarter each. The BNS subset has component masses distributed uniformly from 1 to $3M_\odot$, and the z components of dimensionless spin (which are parallel to the orbital angular momentum of the binary) distributed uniformly from -0.05 to 0.05 . The BBH subset has component masses distributed uniformly from 5 to $50M_\odot$ and the z components of dimensionless spin distributed uniformly from -0.99 to 0.99 . The broad subset spans all four parameter spaces mentioned above. It is distributed uniformly in the log of the component masses from 1 to $300M_\odot$ and has the z components of dimensionless spin distributed uniformly from -0.99 to 0.99 . In addition to the definitions of the BNS and BBH parameter spaces provided above, and the implied NSBH parameter

space definition, for the purpose of this paper, we shall consider the parameter space with either component mass greater than $50M_\odot$ to be the IMBH space. The distribution of the blind injection set in the two component masses can be seen in Fig. 4.

A point to note is that even though 868 blind injections may sound high, most of these are too quiet to be recovered, as shown in Fig. 4, and hence will not cause any contamination. The result of the analysis shows that only 190 blind injections are recovered with a FAR ≤ 1 per 5 months. We will also see later that BNS and NSBH template bins are not affected by signal contamination to a significant degree. As a result, only the BBH and IMBH injections will contribute to contaminating the background. Given the high number of GW candidate events we expect to detect in O4, this is a reasonable representation of the total amount of signal contamination we expect to see.

We also perform an injection campaign to calculate the sensitivity of the analysis, both with and without the application of the background filter. The injection set is distributed similarly to the blind injection set, but with a total size of 86,606 injections. It is important to note that the injections and blind injections are analyzed separately, with the blind injections affecting injection recovery only through the background events they add to the $\rho - \xi^2$ background histograms.

C. Sensitivity improvements

In order to estimate the sensitivity of a search, we use the sensitive volume-time (VT) as a measure. The volume that we analyze is determined by the efficiency of recovering injections at a given FAR and redshift, and T is the live time of the analysis. We calculate VT separately for injections falling in four different chirp mass bins. The first is from 0.5 to $2M_\odot$, the second from 2 to $4.5M_\odot$, the third from 4.5

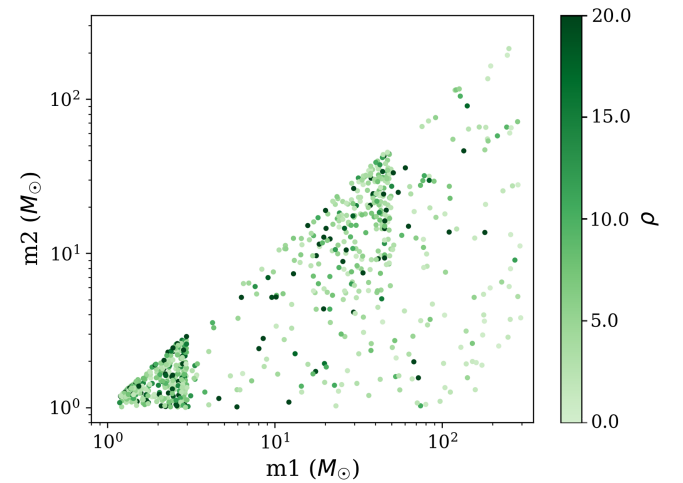


FIG. 4. The distribution of component masses of the blind injection set, colored by injected ρ . Blind injections are used to replicate the contamination caused by GW signals in the data.

to $45M_{\odot}$, and the final one is from 45 to $450M_{\odot}$. The reason for calculating VT separately for different mass bins is so that we have an idea about how sensitive the analysis is for different source categories, with the four mass bins roughly corresponding to BNS, NSBH, BBH, and IMBH source categories, respectively.

Comparing the blind injection analysis with the control run, signal contamination due to the presence of blind injections causes a small ($\sim 5\%$) decrease in VT in the two lowest mass bins, but causes a significant ($\sim 20\%$ – 30%) decrease in VT in the two highest mass bins. This is shown in Fig. 5. High mass templates have a greater match with their neighboring templates, and with themselves across time, as compared to low mass templates. We hypothesize that this causes a single high mass GW signal to be recovered multiple times across template bins and time with suboptimal ρ , leading to more signal contamination in the high mass template bins than in the low mass ones. This is discussed in more detail in Appendix C.

Next, to check the efficacy of the background filter in removing contamination, we compare the VT of the rerank to the VT of the control run. Despite the presence of blind injections in the data, the background filter mitigates the effect they have on the background histograms, and sensitivities of all four mass bins are close to the same as what they were in the control run. This is shown in Fig. 6. This represents a 20%–40% increase in the sensitivities of the two high mass bins in the case of the rerank, as compared to that of the blind injection analysis. We can conclude that the background filter is successful in removing close to all of the contamination that the blind injections cause. Since the number of blind injections we used was a high estimate of the number of GW events we expect to see in O4, this means that by using the background filter, we do not expect signal contamination to be a

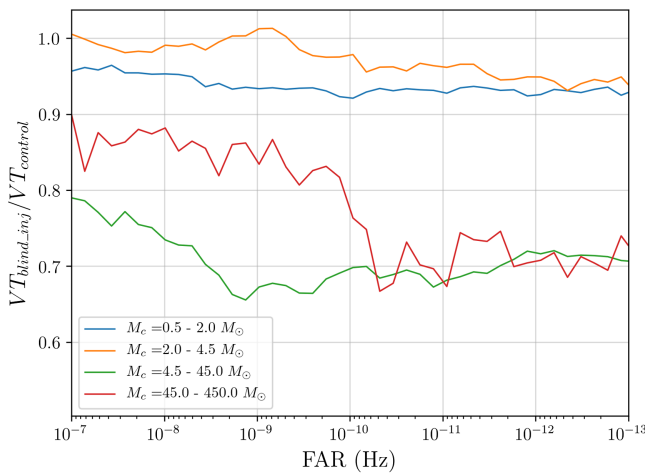


FIG. 5. The decrease in VT caused by signal contamination due to the presence of blind injections in the data. The two highest mass bins are the most affected. The presence of GW signals will also have a similar effect.

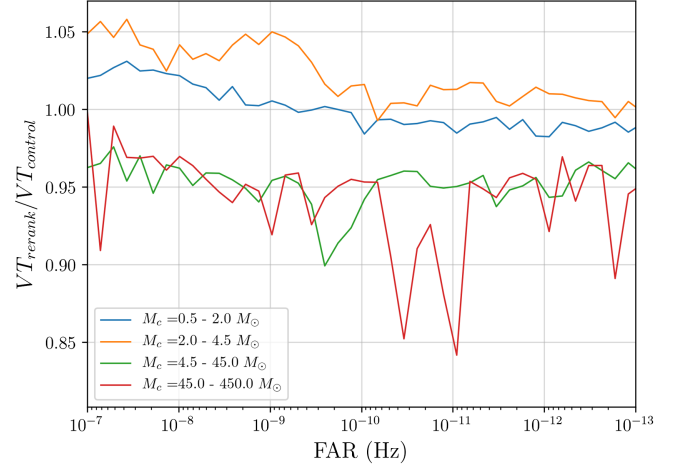


FIG. 6. The VT of the rerank, which has blind injections with the background filter applied, compared to that of the control run, which has neither. The fact that all four lines are close to 1 tells us that the background filter is successful in removing nearly all of the contamination caused by the presence of the blind injections in the data. The peaks and dips in the highest mass bin curve are explained by the smaller number of injections in this bin as compared to other bins, leading to greater variance.

significant problem during O4. To test our readiness for O4, GstLAL has participated in a mock data challenge, where an online analysis is run over 40 days of O3 data [33]. This chunk of data contains nine GW candidates. The background filter was deployed in this analysis, and applied to all nine candidates. It was able to remove all instances of signal contamination we had previously seen, as verified by visual inspection.

V. CONCLUSION

GstLAL constructs $\rho - \xi^2$ background histograms to calculate the $P(\vec{\rho}, \vec{\xi}^2 | t_{\text{ref}}, \vec{\theta}, \mathcal{H}_n)$ term in the likelihood ratio. However, GW signals in the data can cause the background histograms to be incorrectly constructed. This is called signal contamination, and it leads to the sensitivity of the GstLAL analysis being lowered.

The background filter is a novel way to remove the contamination. It records the events that populate the background histograms which satisfy two constraints. The first is that the event must fall in an area in $\rho - \xi^2$ space consistent with GW signals. The second is that it must fall in a 10 s window around a significant event. The user then identifies which of the significant events originate from GW signals. The user communicates this to the GstLAL analysis in real time, and then the events recorded by the background filter corresponding to the times identified by the user are subtracted from the background histograms. Thus, signal contamination is removed from the background histograms.

To test the efficacy of the background filter, we ran a GstLAL analysis over a week of O3 data, with simulated

gravitational-wave signals injected into the data. We found that signal contamination primarily affects the high mass bins. The sensitivity of these bins decreased by 20%–30% due to the presence of the gravitational-wave signals. By applying the background filter, we were able to increase the sensitivity close to what it was without the injected gravitational-wave signals. This shows that the background filter is effective in removing nearly all the signal contamination. With a high rate of gravitational-wave events expected during O4, the background filter will be an important tool in improving the sensitivity of the GstLAL analysis.

ACKNOWLEDGMENTS

The authors are grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants No. PHY-0757058 and No. PHY-0823459. This material is based upon work supported by NSF’s LIGO Laboratory which is a major facility fully funded by the National Science Foundation. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation (NSF) and operates under cooperative agreement No. PHY-1764464. The authors are grateful for computational resources provided by the Pennsylvania State University’s Institute for Computational and Data Sciences (ICDS) and by the California Institute of Technology, and support by Grants No. NSF PHY-2011865, No. NSF OAC-2103662, No. NSF PHY-1626190, No. NSF PHY-1700765, No. NSF PHY-2207728, and No. NSF PHY-2207594. This research has made use of data or software obtained from the Gravitational Wave Open Science Center [39], 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 Science and Technology Facilities Council (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, and Spain. K. A. G. R. A. is supported by Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan Society for the Promotion of Science (JSPS) in Japan; National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea; Academia Sinica (AS) and National Science and Technology Council (NSTC) in Taiwan.

APPENDIX A: CHOICE OF CONSTRAINTS AND THEIR IMPACT ON PERFORMANCE

With the constraints described in Sec. III, the background filter does not consume too many resources. When looking at a month-long GstLAL analysis, we found that on average, it adds $\sim$ bytes to kilobytes to the data products stored by a GstLAL analysis for every template bin. We did not see any significant increase to the memory used by the GstLAL analysis either. Figure 6 shows that with these constraints, the background filter is effective in removing close to all contamination.

To check if there is any improvement to the sensitivity upon loosening the ρ and ξ^2 constraints, we performed the same analysis as described in Sec. IV, but with the ρ and ξ^2 constraints changed to record events with $\rho > 6$ and $\xi^2/\rho^2 < 0.4$. This broader bounding box for recording events did not have any noticeable effect on the sensitivity. This is shown in Fig. 7. However, loosening the constraints did increase memory usage of the GstLAL analysis to a noticeable degree.

We do not expect that loosening the $\rho > 6$ constraint or the time constraint would increase sensitivity, since the extra events collected by changing these constraints would be no more significant than noise. This discussion, along with Fig. 6 shows us that the existing constraints used by the background filter satisfy all our requirements.

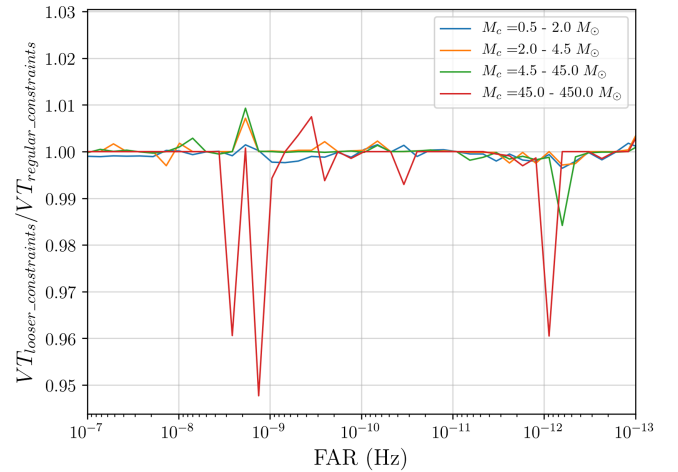


FIG. 7. The VT with the background filter recording events with looser constraints, as compared to the VT with the background filter recording events with the regular constraints. The fact that the VT ratios for all four mass bins are close to 1 shows that loosening the constraints does not improve sensitivity. The peaks and dips in the highest mass bin curve are explained by the smaller number of injections in this bin as compared to other bins, leading to greater variance. Both analyses included the 868 blind injections described in Sec. IV. The regular constraints are described in Sec. III, whereas the looser constraints are described in Appendix A.

APPENDIX B: CRITERIA FOR REMOVING EVENTS FROM THE BACKGROUND, AND ITS EFFECT ON SENSITIVITY

In Sec. III, we saw that in order for an event to be removed from the background, it needs to pass the ρ and ξ^2 constraints, the time constraint, and the user needs to inform the analysis that there was a GW candidate at the time of the event. This last condition was implemented as an additional check that the event being removed does actually originate from a GW candidate. For the background histograms to accurately model the $\rho - \xi^2$ noise LR term, only events originating from noise must populate the $\rho - \xi^2$ histograms. Events originating from GW candidates entering the histograms and events originating from noise being removed from the histograms will both cause the $\rho - \xi^2$ noise LR term to not be evaluated correctly, and will cause a lowering of sensitivity.

As a result, it is important to choose the criteria for informing the analysis of a GW candidate correctly. For O4, we have chosen this criteria to be that the candidate has a FAR ≤ 1 per 5 months, and that it is not retracted. The FAR threshold is low enough that it is highly unlikely that the candidate is not astrophysical in origin without being a retraction, while simultaneously being high enough to not exclude too many contaminating (and hence, loud) GW candidates.

To test this, we set up a rerank similar to the one described in Sec. IV, but instead of applying the background filter to blind injections with FAR ≤ 1 per 5 months, we applied it to all blind injections. This rerank

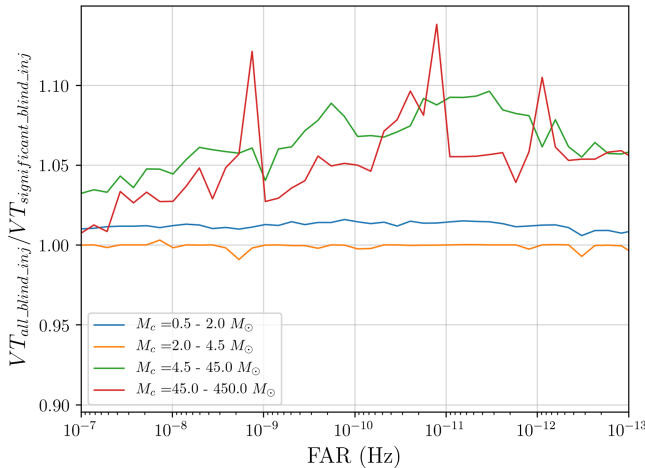


FIG. 8. The VT with the background filter removing contamination from all blind injections, as compared to the VT with the background filter removing contamination from blind injections with FAR ≤ 1 per 5 months. The first represents the best-case performance of the background filter, whereas the second represents the current (O4) performance, designed to prevent noise events from being removed from the background. This graph shows that the current system recovers almost all of the lost sensitivity due to signal contamination.

represents the biggest improvement in sensitivity possible from the background filter. The result is shown in Fig. 8. We can see that the FAR threshold of 1 per 5 months works well, and out of the 20%–30% lost sensitivity due to signal contamination from the blind injections, it manages to recover most of it (see Fig. 5, Fig. 6), with only around 5% not being recovered in the heavier mass bins.

APPENDIX C: DIFFERING IMPACTS OF SIGNAL CONTAMINATION OF THE SENSITIVITIES OF TEMPLATE BINS

As discussed in Sec. IV, signal contamination only causes a 5% decrease in the VT of low mass template bins, such as the BNS and NSBH bins, whereas it causes a 20%–30% decrease in the VT of high mass template bins, such as the BBH and IMBH bins. This is despite the fact that there are more blind injections in the low mass parameter spaces than in the high mass ones. We conjecture two reasons for this, the first relating to how the correlation among neighboring templates changes with mass, and the second relating to how the correlation of a templates with itself across time changes with mass. For the remainder of this section, we shall treat BNS template bins as representative of all low mass bins, and IMBH template bins as representative of all high mass bins.

The “bank correlation function” of a template measures how well it matches with other templates in the template bank. This calculation is similar to how ρ is calculated, except that the match is calculated between two templates with no time shift between them. To see how the bank correlation function of templates changes with mass, we took 5 BNS template bins (corresponding to ~ 5000 templates), calculated the bank correlation of every combination of templates, and plotted the average bank correlation function in descending order of template match. We then did the same for 5 IMBH template bins. The results are shown in Fig. 9. The fact that the BNS bank correlation function drops sharply as compared to the IMBH one means that there are many IMBH templates across template bins that can recover a given IMBH GW signal with a lower ρ than the best template, but only a few BNS templates that can recover a given BNS GW signal. This means a high mass GW signal will create many events, increasing the probability of signal contamination. This is not a problem for the GW candidates reported by GstLAL, since “event clustering” [11] ensures that only the best candidate in an 8 s window survives.

The “autocorrelation function” of a template measures how well it matches with a time-shifted version of itself, similar to how ρ calculates the match between the data and a time-shifted template. The autocorrelation function of a typical BNS template and a typical IMBH template are shown in Fig. 10. The IMBH autocorrelation function has multiple secondary peaks ~ 5 –10 ms away from the primary one. We conjecture that in the case of an IMBH GW

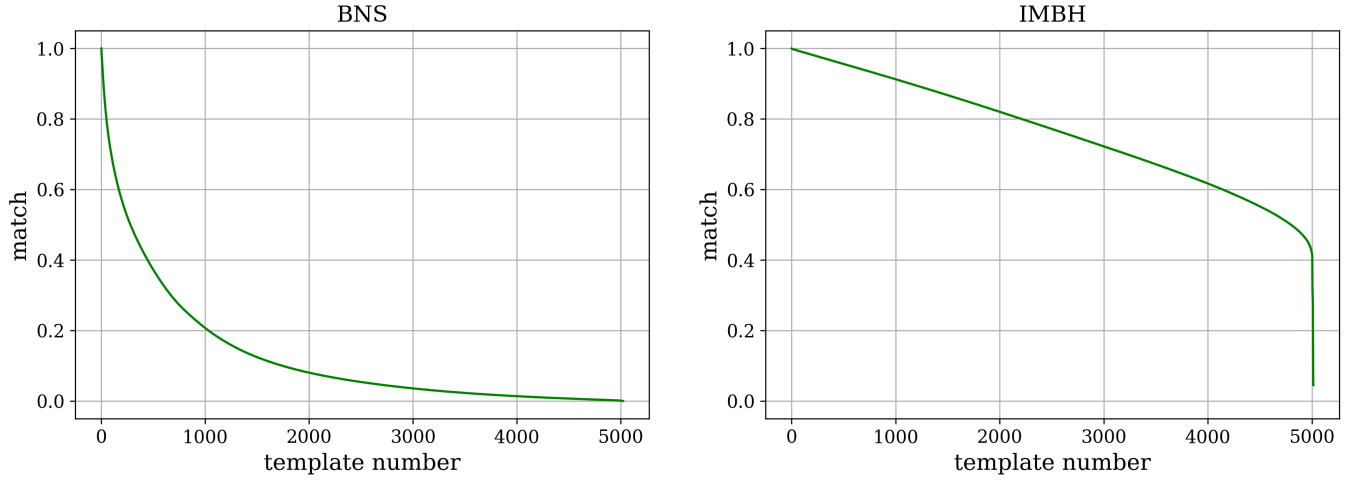


FIG. 9. The average bank correlation function of a BNS template in descending order of template match, as compared to that of an IMBH template, calculated for the five closest template bins. Since IMBH templates correlate well with other IMBH templates across template bins, an IMBH GW signal will be recovered by multiple template bins, increasing the probability of signal contamination. This is not the case for BNS template bins, and it is more likely a BNS GW signal will be recovered by only one template bin, resulting in fewer cases of signal contamination.

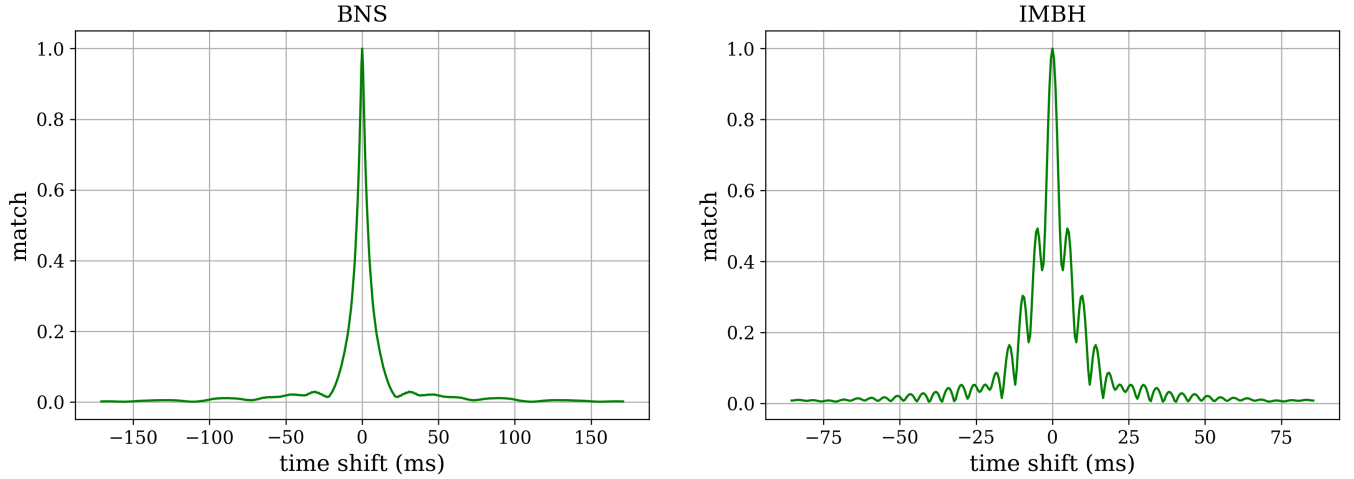


FIG. 10. The autocorrelation function of a BNS template, as compared to that of an IMBH template. Since there are multiple peaks in autocorrelation function of the IMBH template, a quiet IMBH GW signal could be recovered in different detectors at different times, corresponding to the different peaks in the IMBH autocorrelation function. This will result in the GW signal being recovered as multiple single detector events rather than a single coincident event, which leads to signal contamination. Since the BNS autocorrelation function does not have multiple peaks, signal contamination is less likely for BNS template bins.

signal with low ρ or in high noise, an IMBH template could recover the signal in different detectors at different times, corresponding to the different peaks in the IMBH autocorrelation function. This would cause the signal to be recovered as multiple single detector events, instead of a single coincident one, leading to signal contamination of the high mass template bins. Again, this is not a problem for the GW candidates reported by GstLAL, due to event clustering. Since all the secondary peaks in the autocorrelation function of an IMBH template lie well within an 8 s

window, multiple single detector events will be clustered out, and only the best one will survive.

For high mass bins, the bank correlation factor increases the probability of low ρ events getting created by a GW signal, and the autocorrelation factor increases the probability of signal contamination from those events. These two factors compound each other's effect, and as a result, we see a much higher impact of signal contamination in the high mass template bins than in the low mass ones.

- [1] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 074001 (2015).
- [2] F. Acernese *et al.* (Virgo Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).
- [3] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **9**, 031040 (2019).
- [4] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **11**, 021053 (2021).
- [5] R. Abbott *et al.* (The LIGO Scientific and The Virgo Collaborations), [arXiv:2108.01045](https://arxiv.org/abs/2108.01045) [Phys. Rev. D (to be published)].
- [6] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), [arXiv:2111.03606](https://arxiv.org/abs/2111.03606) [Phys. Rev. X (to be published)].
- [7] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **913**, L7 (2021).
- [8] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **103**, 122002 (2021).
- [9] The LIGO Scientific, the Virgo, and the KAGRA Collaborations, [arXiv:2304.08393](https://arxiv.org/abs/2304.08393).
- [10] R. Magee, A.-S. Deutsch, P. McClincy, C. Hanna, C. Horst, D. Meacher, C. Messick, S. Shandera, and M. Wade, *Phys. Rev. D* **98**, 103024 (2018).
- [11] C. Messick *et al.*, *Phys. Rev. D* **95**, 042001 (2017).
- [12] S. Sachdev *et al.*, [arXiv:1901.08580](https://arxiv.org/abs/1901.08580).
- [13] C. Hanna *et al.*, *Phys. Rev. D* **101**, 022003 (2020).
- [14] K. Cannon *et al.*, *SoftwareX* **14**, 100680 (2021).
- [15] <https://gcn.gsfc.nasa.gov/other/G298048.gcn3>.
- [16] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **848**, L12 (2017).
- [17] T. Akutsu *et al.*, *Prog. Theor. Exp. Phys.* **2021**, 05A101 (2020).
- [18] L. Tsukada *et al.*, *Phys. Rev. D* **108**, 043004 (2023).
- [19] K. Cannon, C. Hanna, and J. Peoples, [arXiv:1504.04632](https://arxiv.org/abs/1504.04632).
- [20] K. Cannon, C. Hanna, and D. Keppel, *Phys. Rev. D* **88**, 024025 (2013).
- [21] D. Mukherjee *et al.*, *Phys. Rev. D* **103**, 084047 (2021).
- [22] S. Sakon *et al.*, [arXiv:2211.16674](https://arxiv.org/abs/2211.16674).
- [23] T. D. Canton, A. H. Nitz, B. Gadre, G. S. C. Davies, V. Villa-Ortega, T. Dent, I. Harry, and L. Xiao, *Astrophys. J.* **923**, 254 (2021).
- [24] G. S. Davies, T. Dent, M. Tápai, I. Harry, C. McIsaac, and A. H. Nitz, *Phys. Rev. D* **102**, 022004 (2020).
- [25] A. H. Nitz, *Phys. Rev. D* **98**, 024050 (2018).
- [26] F. Aubin, F. Brighenti, R. Chierici, D. Estevez, G. Greco, G. M. Guidi, V. Juste, F. Marion, B. Mours, E. Nitoglia, O. Sauter, and V. Sordini, *Classical Quantum Gravity* **38**, 095004 (2021).
- [27] T. Adams, D. Buskulic, V. Germain, G. M. Guidi, F. Marion, M. Montani, B. Mours, F. Piergiovanni, and G. Wang, *Classical Quantum Gravity* **33**, 175012 (2016).
- [28] Q. Chu, *Phys. Rev. D* **105**, 024023 (2022).
- [29] Q. Chu, Low-latency detection and localization of gravitational waves from compact binary coalescences, Ph.D. thesis, The University of Western Australia, 2017.
- [30] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, *Phys. Rev. D* **100**, 023011 (2019).
- [31] B. Zackay, L. Dai, T. Venumadhav, J. Roulet, and M. Zaldarriaga, *Phys. Rev. D* **104**, 063030 (2021).
- [32] IGWN public alerts user guide: Observing capabilities, <https://emfollow.docs.ligo.org/userguide/capabilities.html>.
- [33] B. Ewing *et al.*, [arXiv:2305.05625](https://arxiv.org/abs/2305.05625).
- [34] B. Moe, P. Brady, B. Stephens, E. Katsavounidis, R. Williams, and F. Zhang, GraceDB: A Gravitational Wave Candidate Event Database (2014), <https://gracedb.ligo.org/>.
- [35] BOTTLE: PYTHON web framework: <https://bottlepy.org/docs/dev/>.
- [36] IGWN public alerts user guide, <https://emfollow.docs.ligo.org/userguide/analysis/index.html##alert-threshold>.
- [37] R. Abbott *et al.* (The LIGO Scientific, The Virgo, and The KAGRA Collaborations), *Astrophys. J. Suppl. Ser.* **267** 29 (2023).
- [38] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **892**, L3 (2020).
- [39] <https://gwosc.org/>.