

## Optimized PyCBC search for gravitational waves from intermediate-mass black hole mergers

Koustav Chandra<sup>1,\*</sup>, V. Villa-Ortega<sup>2</sup>, T. Dent<sup>2</sup>, C. McIsaac<sup>3</sup>, Archana Pai,<sup>1</sup> I. W. Harry<sup>3</sup>,  
G. S. Cabourn Davies<sup>3</sup> and K. Soni<sup>4</sup>

<sup>1</sup>*Department of Physics, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India*

<sup>2</sup>*IGFAE, Campus Sur, Universidade de Santiago de Compostela,  
ES-15782 Santiago de Compostela, Spain*

<sup>3</sup>*University of Portsmouth, Portsmouth PO1 3FX, United Kingdom*

<sup>4</sup>*Inter-University Centre for Astronomy and Astrophysics, Pune 411007, India*

 (Received 2 June 2021; accepted 23 July 2021; published 25 August 2021)

The detection of intermediate-mass black hole (IMBHs) i.e., those with mass  $\sim 100\text{--}10^5 M_{\odot}$ , is an emerging goal of gravitational-wave (GW) astronomy with wide implications for cosmology and tests of strong-field gravity. Current PyCBC-based searches for compact binary mergers, which match filtered the detector data against a set of template waveforms, have so far detected or confirmed several GW events. However, the sensitivity of these searches to signals arising from mergers of IMBH binaries is not optimal. Here, we present a new optimized PyCBC-based search for such signals. Our search benefits from using a targeted template bank, stricter signal-noise discriminators, and a lower matched-filter frequency cut-off. In particular, for a population of simulated signals with isotropically distributed spins, we improve the sensitive volume-time product over previous PyCBC-based searches at an inverse false alarm rate of 100 years, by a factor of 1.5 to 3 depending on the total binary mass. We deploy this new search on Advanced LIGO-Virgo data from the first half of the third observing run. The search does not identify any new significant IMBH binaries but does confirm the detection of the short-duration GW signal, GW190521, with a false alarm rate of 1 in 727 years.

DOI: [10.1103/PhysRevD.104.042004](https://doi.org/10.1103/PhysRevD.104.042004)

### I. INTRODUCTION

What started as a trickle back in September 2015 [1] has turned into a veritable deluge [2,3] of gravitational-wave (GW) detections at the end of the first half of the third observing run (O3a) of advanced GW detectors. As a result, the gravitational wave catalogs [2–8] now list more than 50 compact binary coalescences (CBCs). This rapid surge in GW events is largely due to the upgrades of the Advanced LIGO [9] and Advanced Virgo [10] instruments in the interim between the observing runs. The search algorithms, namely cWB [11–15], GstLAL [16–18], PyCBC [19–21], MBTA [22,23], and SPIIR [24] also have improved significantly enabling them to better separate gravitational signal from the detector noise. As a consequence, the second gravitational wave transient catalog (GWTC-2) not only reports the detection of a plethora of stellar mass compact binaries [3,25] but also details the detection of a couple of relatively asymmetric-mass compact binaries [26,27], and a signal, GW190521, from a putative intermediate-mass black hole (IMBH) binary [28,29].

The signal GW190521 is quite different from other reported GW events. The premerger phase of the signal was barely observable within the detectors' bandwidth, allowing for its multivalent interpretation [28–35]. If we assume that the signal is from a quasicircular black hole binary, then the heavier component of the binary has a high probability of populating the pair-instability mass gap, and the remnant, with mass  $\sim 142 M_{\odot}$ , is the first conclusively observed intermediate-mass black hole [28,29].

The event was first identified on May 21, 2019 at 03:02:29 UTC by the low latency matched-filter search for compact binaries, PyCBC-LIVE [36,37], and the weakly modeled transient search, cWB, in its IMBH configuration [38]. The two searches reported the event with an inverse false alarm rate (IFAR) of 8.34 years and 28.0 years, respectively. The event was also reported by two other low-latency searches, namely GstLAL and SPIIR, though with lower statistical significance.

Low latency searches are designed to generate rapid alerts to enable prompt follow-up of gravitational-wave triggers for potential electromagnetic or other counterpart signals. Hence, the statistical significance estimated by low-latency pipelines is limited due to the available amount of empirically generated background. Therefore, most

\*koustav.chandra@iitb.ac.in

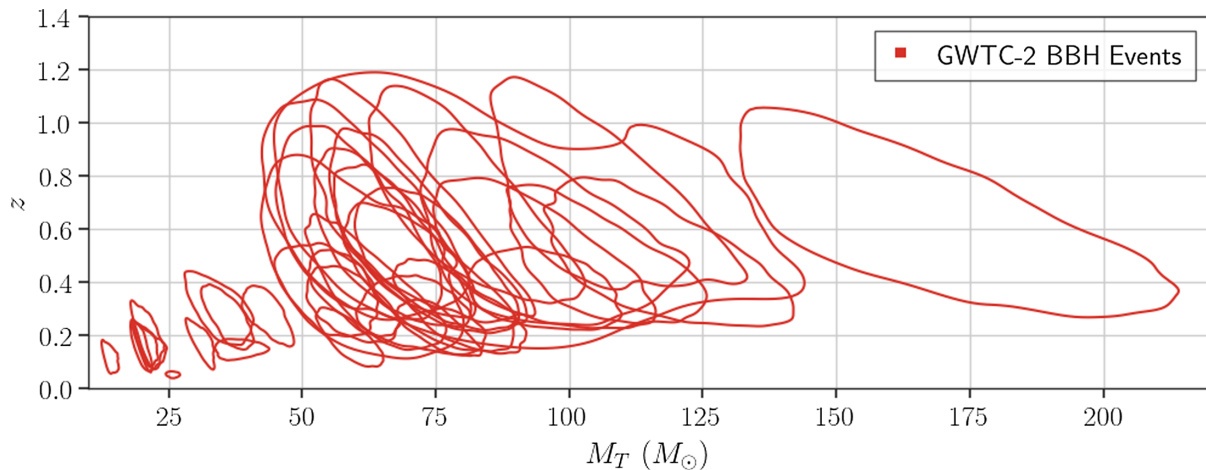


FIG. 1. Posterior probability distribution of the source frame total mass  $M_T$  and redshift  $z$  for the 37 confident BBH events announced in GWTC-2 [red]. The contours enclose 90% of the posterior probability, inferred under the default parameter estimation priors used in the GWTC-2 publications [3]. For several high-mass mergers, the redshift extends to  $z \sim 1$ : note that the template masses required to detect such signals scale as  $(1+z)M_T$ .

search pipelines rescrutinize the strain data with improved calibration, data quality vetoes, and signal-noise discriminators, which are generally unavailable during low-latency searches. Further, off-line searches estimate the background over a larger amount of data, and can thus provide a better estimate of a candidate’s statistical significance. In particular, when such an analysis was carried out at the time around GW190521, the search pipelines reported a maximum IFAR of 4900 years [28,38].

However, contrary to other searches, the off-line PyCBC-based search, PyCBC-broad, identified the event with an IFAR of less than a year. This is primarily due to the strong disagreement between the event and the best-matching template used in the search [39]. This search also suffered from an inherent “look-elsewhere effect” as it tries to find CBCs signals over a broad parameter space from binary neutron stars up to IMBH binaries. The other PyCBC-based search, namely PyCBC-BBH, did not detect the event due to its tight tuning to signals closely matching stellar-mass black hole binary (BBH) templates.

The two gravitational-wave transient catalogs to date also show some interesting trends concerning binary masses. In the first LIGO-Virgo catalogue, GWTC-1 [2], a lack of systems with  $m_1 \gtrsim 45 M_\odot$  is evident. However, more recently GWTC-2 lists several systems in this range, indicating additional structure in the BBH population distribution as discussed in Abbott *et al.* [40]. In Fig. 1 we show the inferred source total masses  $M_T$  and redshifts  $z$  for 37 confidently detected BBHs announced in GWTC-2. The lack of low-mass events at high  $z$  is due to a selection effect as the amplitude of emitted GW scales with mass. Already, several high-mass BBHs are detectable at  $z \sim 1$ ; thus, the template masses required to detect such systems are significantly larger than the source masses due to redshifting of the GW signal [41]. Thus, even for (relatively

heavy) stellar-mass BBH ( $\gtrsim 50 M_\odot$ ), an effective IMBH binary search is required.

We expect these trends to continue during upcoming runs when GW interferometers achieve higher global network sensitivity to BBH mergers [42]. The improvement will facilitate the detection of potentially hundreds of BBHs per year, reaching out to larger redshifts [43]. Also, within the turn of the next decade, we will have more sensitive detectors such as Cosmic Explorer [44] and the Einstein Telescope [45,46] that will help detect more BBH mergers at even further distances (possibly up to  $z \sim 20$  and higher). The launch of the Laser Interferometer Space Antenna (LISA) will potentially allow us to look even deeper and for more massive BBHs [47]. Such binaries may be heavier than those observed so far [48] and so would produce remnants in the lower ( $10^2 M_\odot - 10^3 M_\odot$ ) to mid ( $10^3 M_\odot - 10^4 M_\odot$ ) IMBH mass range [49]. Detecting such systems in multiband GW window [50,51] will help us to estimate better the population distribution of low and mid mass range IMBHs, which is currently lacking. It will also enable us to understand better the formation mechanism of IMBH systems and to perform more detailed tests of general relativity in the strong-field regime [52]. And as mentioned above, GW signals from high redshift binaries will have higher detector frame masses, up to an order of magnitude or larger than the source masses. So, regardless of whether we detect any more IMBHs, we still need to extend the target space of current CBC searches and tune them to prevent missing these potential signals.

We make such an attempt in this work. Here, we propose a new PyCBC-based search which is optimized to detect both GW signals from IMBH binaries and highly redshifted heavy stellar-mass BBH mergers. We construct a bank of nonprecessing waveforms simulating BBH mergers with component detector frame masses  $m_{1,2}(1+z) \geq 40 M_\odot$ ,

total detector frame masses  $M_T(1+z) \in [100, 600] M_\odot$  and mass ratios  $q \in [1, 10]$ . Because of our choices in the target search space, our search is susceptible to a larger number of noise triggers, requiring us to impose stricter noise vetoes and discriminators [53,54]. We find that, while keeping the basic structure of the PyCBC search algorithm, our modifications increase the search sensitivity, in terms of sensitive volume-time, for a simulated population of binaries, by an average factor of  $\sim 2$  at an IFAR of 100 years, compared to the PyCBC-broad search used in Abbott *et al.* [28].

We organize our paper as follows. Section II serves as an overview of the multidetector PyCBC analysis. Optimizations and changes to this analysis are presented in Sec. III. In Sec. IV, we describe how these modifications combine to improve the sensitivity of the PyCBC analysis to IMBH mergers. Because of the changes made, the search now recovers the heavier mass black hole binary mergers from the third observing run with higher significance, and we have summarized these findings in Sec. V. Finally, in Sec. VI we conclude the paper by indicating the future optimization prospects in the upcoming observing runs.

## II. REVIEW OF EXISTING PyCBC DETECTION METHODS FOR CBC SYSTEMS

The strain that an incident GW signal produces in the interferometric detector is a weighted linear combination of the gravitational wave polarization  $h_+(t)$  and  $h_\times(t)$

$$s(t) = F_+(\alpha, \delta, \psi)h_+(t; \vec{\zeta}, \vec{\kappa}) + F_\times(\alpha, \delta, \psi)h_\times(t; \vec{\zeta}, \vec{\kappa}), \quad (1)$$

where the weights  $F_{+/\times}$  represent the detector beam pattern functions. These functions depend on the source's sky location  $(\alpha, \delta)$  with respect to the detector and on the polarization angle  $\psi$ . If the signal is from a quasicircular binary, then  $\vec{\zeta}$  in the above equation describes the intrinsic parameters of the binary, which comprises of the individual component masses  $m_i$  and the dimensionless spin  $\vec{\chi}_i$  of the binary.  $\vec{\kappa}$ , on the other hand, consists of the luminosity distance  $D_L$  to the source, the inclination angle  $\iota$  of the binary with respect to the observer, and the merger time and phase  $(t_c, \varphi_c)$  of the signal.

Current PyCBC-based GW searches primarily target signals coming from mergers of binaries consisting of a pair of neutron stars, a pair of black holes, or a neutron-star black hole pair. So, in the subsequent subsections, we briefly review the steps involved in detecting these well-modeled CBC signals.

### A. Preprocessing of detector data

The calibrated detector output is an evenly spaced time-series data containing information about strain due to different sources. This makes it amenable to standard

techniques of time series analysis. The key foundation of prevalent gravitational wave time-series analysis is a linear model of the data  $d(t)$ :  $d(t) = s(t) + g(t) + n_G(t)$ , where  $s(t)$  is the GW signal strain,  $n_G(t)$  is the strain due to random (Gaussian) detector noise, and  $g(t)$  is the strain due to non-Gaussian terrestrial transients, also known as ‘glitches’.

For most of the time, the data  $d(t)$  is just  $n_G(t)$  which is usually assumed to be additive, wide-sense stationary colored Gaussian noise. This presupposition allows one to fully describe the noise's statistical properties in terms of its power spectral density (PSD),  $S_n(f)$ . PyCBC-based searches estimate the PSD via Welch's method and use it to whiten the detector data so that any excess power due to the presence of either  $s(t)$  or  $g(t)$  is more prominent.

However, the PSD estimate can be significantly affected during periods of poor detector performance. Therefore, the PyCBC toolkit [55] uses data quality information to zero out flagged times of badly contaminated data before evaluating PSD. Further, to reduce the number of false alarms caused due to high-amplitude glitches, the search algorithm internally identifies times of large deviation from Gaussian noise in the whitened strain data and window them out using a procedure known as autogating [19]. The PyCBC algorithm also vetoes out flagged times of additional glitches that can generate loud triggers but do not contaminate the surrounding data to affect the performance of the PSD estimation. This data wrangling process ensures better search performance but cannot completely purge the data from all glitches  $g(t)$ .

### B. Matched filtering

Most current CBC searches assume that the incident GW signal can be adequately represented by the dominant harmonic of a quasicircular nonprecessing binary. The form of such a signal is well modeled in the framework of general relativity. So, after the preprocessing step, PyCBC-based searches use matched filtering [56] to find whether or not there is a GW signal buried in the detector data [19,57,58]. As the parameters of the incident signal are not known *a priori*, CBC searches construct and search over a discrete bank of templates spanning the search space.

The detector response due to the dominant harmonic of a nonprecessing quasicircular binary in the frequency domain can be written as

$$\tilde{h}(f; \vec{\theta}) = \frac{F}{D_L} \tilde{h}_{22}(f; \vec{\zeta}) e^{i\phi}, \quad (2)$$

provided the stationary phase approximation holds [56,59]. Here,  $\tilde{h}_{22}$  is the geocentric waveform, parametrized completely by  $\vec{\zeta}$ , which forms the set of filter waveforms used in the search.  $F$ , on the other hand, describes the effective antenna pattern function

$$F = \sqrt{F_+^2 \left( \frac{1 + \cos^2 \iota}{2} \right) + F_\times^2 \cos^2 \iota}, \quad (3)$$

while

$$\phi = \arctan \left( \frac{2F_\times \cos \iota}{F_+ (1 + \cos^2 \iota)} \right) + \varphi_c + 2\pi f \Delta t \quad (4)$$

represents the effective merger phase, with  $\Delta t$  denoting the arrival time of the signal to the detector relative to a geocentric time reference.

The above presupposition also allow us to decompose the template strain,  $h_{22}$ , into two orthonormal components,  $\hat{h}_c$  and  $\hat{h}_s$ , with an overall effective amplitude. These can be used to define the log-likelihood ratio,  $\ln \mathcal{L}$ , between the signal and noise hypothesis given  $g(t) = 0$ . It also allows one to maximize the log-likelihood ratio over the effective template phase and amplitude, giving [60]

$$\ln \mathcal{L}_{\max}(t; \vec{\xi}) = \frac{1}{2} \rho^2(t; \vec{\xi}). \quad (5)$$

Here,  $\rho(t)$  denotes the matched-filter signal-to-noise ratio (SNR)

$$\rho(t; \vec{\xi}) = \sqrt{\langle d | \hat{h}_s \rangle^2 + \langle d | \hat{h}_c \rangle^2}, \quad (6)$$

which is evaluated for each template independently in each detector.  $\langle d | h \rangle$  defines the noise-weighted inner product between the preprocessed data and the template in the frequency domain and is given as

$$\langle d | h \rangle(t) = 4 \Re e \int_{f_{\text{low}}}^{f_{\text{high}}} \frac{\tilde{d}(f) \tilde{h}^*(f)}{S_n(f)} e^{2i\pi f t} df. \quad (7)$$

Here,  $f_{\text{low}}$  and  $f_{\text{high}}$  represent the lower and upper frequency cutoff of the matched-filter operation, respectively. Typically,  $f_{\text{low}}$  is set to the frequency from which the filter waveform is generated, while  $f_{\text{high}}$  is set either as the final frequency cutoff of the template or the Nyquist frequency of the detector data, whichever is smaller. A peak in this SNR time-series above a specific threshold will correspond to an astrophysical trigger, provided there is no glitch in the data. Since we do not know *a priori* the parameters describing the gravitational wave in the data, current CBC searches carry out this matched-filter operation over all the templates in the bank. The template which best describes any signal present in the data gives the largest peak in the SNR time-series; hereafter, we will call such a filter waveform the best-matched template.

### C. Single detector noise discriminators

Matched filtering gives the optimal detection statistic provided there are no glitches in the data [61].

Unfortunately, this is not the case. Preprocessing of the calibrated detector data eliminates loud glitches, but the relatively softer glitches still pollute the detector data and form the majority of the single-detector triggers. To reduce the effect of these glitches on the search, the PyCBC toolkit uses several noise discriminators at the single detector level.

Firstly, the PyCBC toolkit discards triggers below some predetermined SNR threshold and cluster triggers within a fixed window of time—keeping only the triggers with the largest SNR. It also tries to suppress loud glitches  $g(t)$  by performing two different signal consistency tests;  $\chi_r^2$  [62] and  $\chi_{\text{sg}}^2$  [63]. The  $\chi_r^2$  test checks if the morphology of the candidate trigger in frequency domain is consistent with that of the best-matched template while the  $\chi_{\text{sg}}^2$  test is designed to down-rank triggers with excess power at frequencies exceeding those allowed by the best-matched template. In the presence of Gaussian noise, both  $\chi_r^2$  and  $\chi_{\text{sg}}^2$  follow a normalized  $\chi^2$  distribution. However, these statistics produce large values in the presence of glitches or when the signal at hand is not well modeled by the best-matched template. PyCBC-based searches use these test results to reweigh the trigger SNR [58,63]. They first calculate the “reweighted SNR” for the triggers via

$$\tilde{\rho} = \begin{cases} \rho \left[ \frac{1}{2} (1 + (\chi_r^2)^3) \right]^{-1/6} & \text{for } \chi_r^2 > 1 \\ \rho & \text{for } \chi_r^2 \leq 1 \end{cases} \quad (8)$$

and then further reweigh it by calculating

$$\hat{\rho} = \begin{cases} \tilde{\rho} (\chi_{\text{sg}}^2)^{-1/2} & \text{if } \chi_{\text{sg}}^2 > 4, \\ \tilde{\rho} & \text{otherwise.} \end{cases} \quad (9)$$

The above two amendments to the standard trigger SNR help reduce the apparent loudness of the glitches and hence increase the false alarm rate (FAR) of quieter short-duration glitches.

The detector data also contains broadband nonstationary detector noise, which affects the detector’s sensitivity in the order of tens of seconds. To alleviate the effects of this noise, the PyCBC algorithm rescales  $\hat{\rho}$  with an estimate of the short-term variation in the PSD,  $v_s(t)$ , [64]. When combined with  $\hat{\rho}$ , this gives the single-detector statistic of the trigger

$$\check{\rho} = \hat{\rho} v_s(t)^{-0.5}. \quad (10)$$

### D. Multi-detector coincident candidates

After generating single-detector triggers from each detector, templated searches perform a coincidence test to identify candidate events. This involves checking whether triggers from the same template are observed



across multiple detectors with time-delays consistent with the light travel time between detectors.

Under the assumption that the noise is independent between detectors, this coincidence test helps us to further separate signal triggers from the noise background. We perform this test for all possible two or three detector combinations. To determine the significance of the surviving multidetector candidates, PyCBC searches assigns them a ranking statistic based on the relative rate densities of signal and noise events [20,21]

$$\Lambda_s = -\ln A_{\{a\}} - \sum_a \ln r_{a,i}(\check{\rho}_a) + \ln \frac{p(\vec{\Omega}|S)}{p(\vec{\Omega}|N)} + 3 \ln \frac{\sigma_{\min}}{\sigma_{\text{ref}}}. \quad (11)$$

The first two terms on the right-hand side estimate the rate of coincident noise triggers in a combination of detectors labeled by  $\{a\}$ , representing the allowed time window  $A$  for coincidences in these detectors and the noise-rate estimate  $r$  in each detector and each template  $i$  at single-detector statistic values  $\check{\rho}_a$ . The noise models  $r_{a,i}(\check{\rho}_a)$  are obtained by fitting (detector- and template-dependent) exponential functions to the observed distributions of single-detector triggers. The probability densities  $p(\vec{\Omega}|S)$  and  $p(\vec{\Omega}|N)$  describe the distributions over extrinsic parameters  $\vec{\Omega}$ —the differences in times-of-arrival, phase, and the relative amplitude of the recovered triggers—expected under the signal or noise hypothesis respectively. The last term provides an estimate of network-sensitive volume based on the least sensitive detector involved in this coincident event, compared to a constant reference sensitivity across detector combinations. This term then provides a natural weighting in favor of more sensitive detector combinations. Therefore, the ranking statistic exploits the information encapsulated in the relative phases, amplitudes, arrival times to the different detectors, noise model, and expected distribution of signal SNR.

### E. Estimation of statistical significance

The final step in GW searches is the determination of the statistical significance of a coincident candidate. PyCBC-based searches do it empirically by first generating background triggers using the method of time shifts; shift the data from one of the detectors with respect to other detector(s) by a duration greater than the light travel time between them and look for artificial coincidences [19,21]. This set of coincidences form the search background, and PyCBC-based searches assign a FAR to an apparent event by comparing its ranking statistics with those of the simulated background triggers. The method, by construction, guarantees that the background candidates thus produced cannot be entirely due to an astrophysical signal. By repeating this procedure, searches can generate noise background worth more than  $10^4$  years with just a few days

of data. To minimize background contamination due to signal triggers, the PyCBC toolkit removes triggers from the background within a certain time window around events with large IFAR.<sup>1</sup>

However, there can be chance coincidences between a signal and a noise trigger. This can significantly contaminate the search background, especially when one of them is loud. The PyCBC toolkit, therefore, hierarchically removes loud triggers to minimize the impact of aforementioned astrophysical contamination while still estimating an unbiased IFAR [65].

Currently, PyCBC-based searches analyze data from all possible two- and three-detector combinations. So it is possible to obtain coincidences of the same event in different detector combinations. The PyCBC pipelines therefore only keep the event with the highest-ranking statistics across detector combinations within a given time-shift window to prevent double counting. Further, while calculating an event’s overall statistical significance, PyCBC searches add the FAR from all contributing detector combinations at the ranking statistic of the event, thus accounting for the multiple possible detector combinations.

## III. NEW METHODS FOR OBSERVING INTERMEDIATE-MASS BLACK HOLE BINARIES

At present, two PyCBC-based searches are looking for GW signals from compact binary coalescence as recently presented in Abbott *et al.* [3], Nitz *et al.* [4,5], and Abbott *et al.* [28]. The first of the two covers a broad range of parameters ranging from binary neutron stars to intermediate-mass black hole (IMBH) binaries, using a hybrid geometric-stochastic template bank [66]; we will refer to this search as “PyCBC-broad”. The other search, namely PyCBC-BBH, employs a subset of the templates used in PyCBC-broad and targets specifically binaries having redshifted component masses greater than  $5 M_{\odot}$  and mass ratio  $q \in [1, 3]$ . Thus, both these analyses are not precisely intended to identify IMBH binary mergers, and their sensitivities to such signals are correspondingly penalized. Further, both the searches use template banks that are devoid of filter waveforms with a duration less than 150 ms measured from  $\sim 20$  Hz, which further limits their sensitivity towards shorter duration signals [28,67–69].

This section introduces a new PyCBC-based search that is specifically designed to look for short duration signals arising from putative IMBH binaries. Owing to its construction, the search is also sensitive to highly redshifted stellar-mass BBH mergers.

### A. Search space and template bank

The black hole mergers we are mostly targeting form the lower mass range ( $10^2$ – $10^3$ ) of IMBH binaries. Signals

<sup>1</sup>IFAR = 1/FAR.

emitted by these systems are extremely short lived ( $\lesssim 1$  s) within the current detector bandwidth, making them difficult to detect. However, since they are a result of CBC mergers, they should share morphological resemblances with other observed GW signals. Hence like most templated searches, we assume that these signals are adequately described by the dominant harmonic of a quasicircular, nonprecessing binary. This allows us to design and use a template bank that is entirely parametrized by  $\vec{\zeta}$ .

In what follows, we discuss the salient features of our proposed template bank and how we construct it.

### 1. Bounds on mass-spin space

For this work, we construct a bank of filter waveforms imitating BBHs with detector frame (redshifted) total mass,  $M_T(1+z)$ , between or close to  $100 M_\odot$  and  $600 M_\odot$ . We limit our templates to a mass ratio between  $1 \leq q = m_1/m_2 \leq 10$ , also imposing that the secondary component has a mass  $m_2 > 40 M_\odot$ . Binaries with more unequal mass ratios range will emit signals with significant higher harmonics, making dominant-mode templates ill equipped at detecting them [70,71]. Further, systems with detector-frame total mass greater than  $600 M_\odot$  will emit very few GW cycles within the detector bandwidth; the matched-filter method is only expected to be useful if the number of cycles is significantly  $> 1$ . Also, as shown in Abbott *et al.* [72], the distance reach to such sources is small because the detector sensitivity degrades rapidly towards lower frequency (see Fig. 2), thus the expected rate of detections is likely to fall off rapidly at higher masses. Further, for such massive systems, the  $(2, \pm 2)$  mode is mostly out of the detector bandwidth, implying that templates including only the dominant harmonic will not appropriately recover the signals. We also conducted preliminary analyses with template waveforms having  $M_T(1+z) > 600 M_\odot$  and found their noise distributions extended to significantly higher amplitudes and were harder to model or suppress, resulting in a poorer overall search sensitivity.

Since we are ignorant about the spin distribution of the IMBH population, we allow  $\chi_1$  and  $\chi_2$  of the templates to take on any values between  $\pm 0.998$ . We don't allow our templates to have misaligned spins<sup>2</sup> as it will violate our underlying assumption of the signal.

Our search domain motivates us to choose the reduced-order representation of spinning effective one-body model, SEOBNRv4 [76,77] as the template-bank approximant. Because of the restriction in the waveform model, we bound the dimensionless spin between the specified range and not between theoretical bounds of  $\pm 1$ .

<sup>2</sup>Due to spin-spin and spin-orbit coupling, misaligned spins cause orbital precession [75].

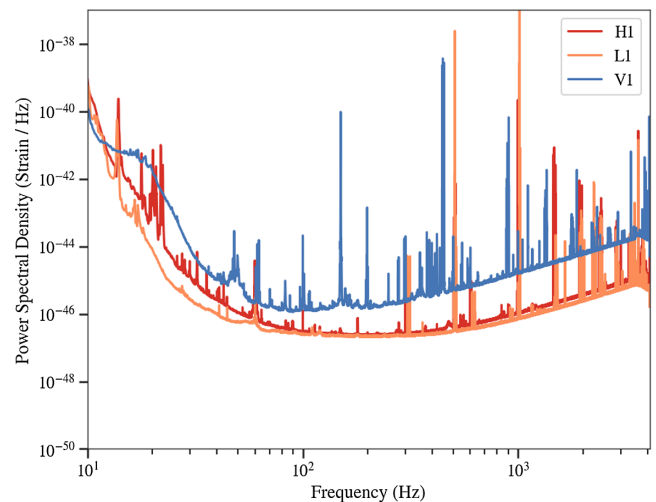


FIG. 2. PSDs estimated using data from each of the contributing detectors used in the first half of the third observing run. The data is now publicly available at the GWOSC website [73,74]. We use the harmonic mean of these PSDs to model the detector noise while constructing the bank.

### 2. Choice of lower frequency cutoff

When constructing the bank and calculating the matched-filter SNR, we use a lower frequency cutoff of 15 Hz. We make this choice to balance out a couple of practical considerations. The quadrupolar GW frequency at merger scales inversely with the binary's detector frame total mass [78]. Therefore, particularly for heavier CBCs, a lower starting frequency for matched filtering equates to better search sensitivity. However, as illustrated in Fig. 2, the noise PSD of the current advanced detectors increases steeply below 15 Hz, indicating that a lower starting frequency would not effectively improve the search performance. Furthermore, we observe that lower values of the cutoff frequencies resulted in significantly higher matched-filter SNRs caused by short-duration glitches [79]. Current signal-noise discriminators, such as  $\chi_r^2$  and  $\chi_{sg}^2$ , often fail to distinguish these transients from astrophysical population appreciably, especially when the best-matched templates are comparatively short lived. So, we restrict the matched-filter calculation to the aforementioned lower frequency cutoff to limit the SNR of these artifacts. We set the higher frequency cutoff to the maximum frequency of the template both while creating the bank and calculating the matched-filter SNR.

### 3. Template duration cut

The above choices in mass and spin space make our search susceptible to a variety of glitches since short-duration CBC templates are well mimicked by glitches. To demonstrate this, we perform a search over  $\sim 32$  days of O3a science data using a template bank covering the specified target space and obeying the frequency constraints. As can be seen in Fig. 3 the proportion of the noise

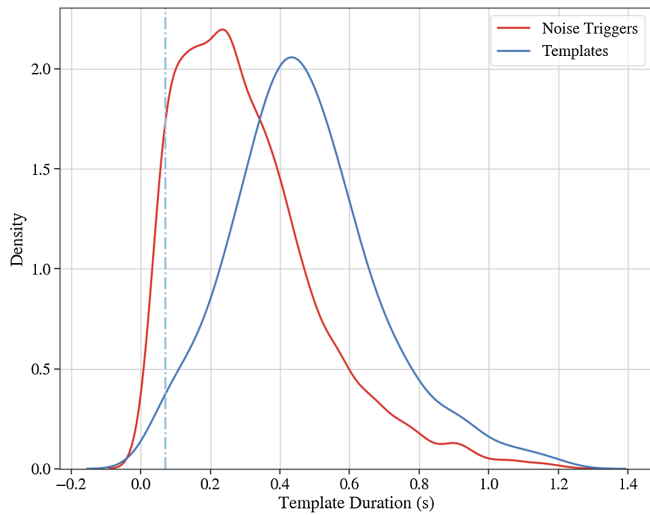


FIG. 3. Kernel density estimate (KDE) plot showing the distribution of noise triggers [red] with  $\rho > 6$  in O3a LIGO-Livingston data while using a template bank without duration threshold. We also show in blue the template duration distribution. The dashed dotted represents the position of the proposed 70 ms with respect to the distribution of the noise triggers. The template duration in the negative zone is a KDE artifact.

triggers increases with a decrease in template duration. After comparing search sensitivity for representative stretches of data using different settings, we empirically excluded any templates having a duration of less than 70 ms (measured from 15 Hz) in the detector bandwidth.

#### 4. Stochastic template bank approach

Having selected the bounds of our template bank, we now seek to create the bank efficiently. Generally, a template bank can be constructed in two ways—geometrically [80–86] or stochastically [87–90], where a hybrid of both methods may be applied to cover a large parameter space [91]. The geometrical approach uses a quasiregular lattice and is particularly successful when applied to low-mass systems, for which the merger and ringdown occur outside of the detector bandwidth. Since the postinspiral phase of the waveform is crucial for our target space, we use the stochastic template placement scheme with a minimal match criterion of 0.99.

The basic idea of the stochastic template scheme is the following. We begin by randomly selecting a potential template waveform,  $h_T$  for our bank  $\mathfrak{B}$ . Initially,  $\mathfrak{B}$  is empty, so we add  $h_T$  directly to it. If however there are any templates in the bank, we evaluate the match,  $\mathcal{M}(h_T, h_i) = \max_{\Delta\phi_c, \Delta t_c} \langle \hat{h}_T | \hat{h}_i \rangle$ , between the proposed template and all templates  $h_i \in \mathfrak{B}$ . Or in other words, we evaluate the effectualness or the fitting factor (FF) [92] of  $\mathfrak{B}$ :

$$\text{FF} = \max_{h_i \in \mathfrak{B}} \mathcal{M}(h_T, h_i). \quad (12)$$

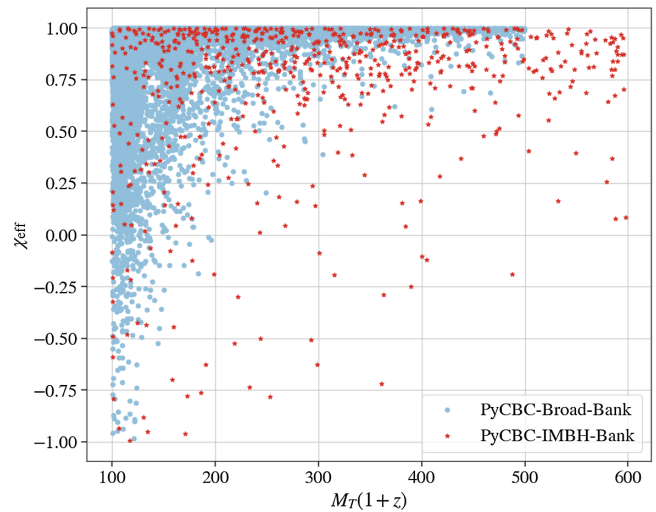


FIG. 4. The distribution of templates in the redshifted total mass and effective spin space. In red are the templates for the bank used in this work, and in light blue, we represent the subset of templates with redshifted total mass greater than  $100 M_\odot$ , as used in the PyCBC-broad search of Abbott *et al.* [3]. The density of the PyCBC-broad templates in the comparatively lower total mass region is larger due to a higher mass ratio cut of 98 and lower minimum component mass requirements. Despite that, the PyCBC-broad bank has a coarser coverage over shorter-duration templates because of the duration cut of 150 ms (measured from  $\sim 20$  Hz). Note that the PyCBC-broad template bank has no templates beyond  $M_T(1+z) = 500 M_\odot$  as opposed to PyCBC-IMBH bank.

If the FF of the new template is below the minimal match of 0.99, then the bank  $\mathfrak{B}$  is deemed ineffectual, and we add the new template  $h_T$  to  $\mathfrak{B}$ . This generates a new seed bank  $\mathfrak{B}' = \mathfrak{B} + h_T$  for the next iteration, and we repeat this process until the rate of rejection of proposed templates reaches a set limit.

The seemingly high choice of minimal match value ensures an adequate number of templates in the high-mass region of our search space where the signals lack premerger power. Figure 4 compares the proposed bank used in this work with that of the one used by the PyCBC-broad in Abbott *et al.* [3,28] in terms of total mass and effective inspiral spin [93,94].<sup>3</sup>

Before performing the search, as a final check, we evaluate the template bank’s performance in terms of effective fitting factor [95]

$$\text{FF}_{\text{eff}} = \sqrt[3]{\frac{\sum_i \text{FF}^3(h_i) \rho_{\text{opt}}^3(h_i)}{\sum_i \rho_{\text{opt}}^3(h_i)}}, \quad (13)$$

<sup>3</sup>The effective inspiral spin  $\chi_{\text{eff}} = [(m_1\chi_1 + m_2\chi_2)/m_1 + m_2]$  ranges between  $\pm 1$  and gives a measure of the alignedness of the out-of-plane component spin with the direction of orbital angular momentum.

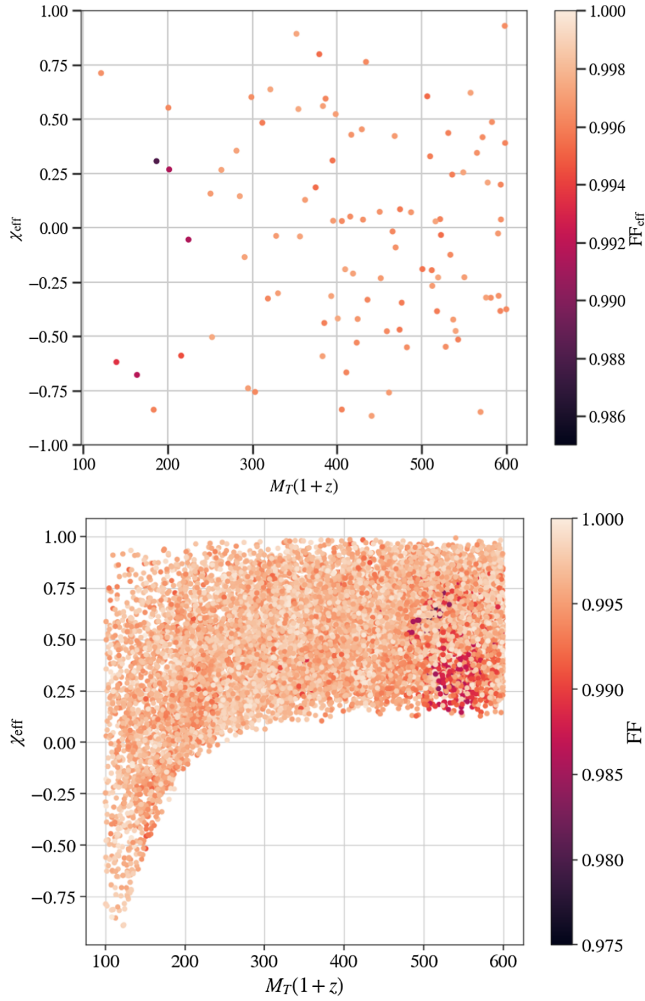


FIG. 5. Top: The effective fitting factor of 100 simulated signals as a function of total mass and effective spin. Note that even though some simulated points lie outside the bank limits, owing to the duration threshold, they still have good  $\text{FF}_{\text{eff}}$ . Bottom: Plot showing the fitting factor of 15,600 simulated nonprecessing signals lying within the target search space.

for a hundred different nonprecessing simulated signals containing only the dominant harmonic and lying within the search target space. Here,  $\rho_{\text{opt}}(h_i) = \sqrt{\langle h_i | h_i \rangle}$  defines the SNR recovered if the template is an exact replica of the incident signal. As can be seen in Fig. 5, the  $\text{FF}_{\text{eff}} \gtrsim 0.99$  for all the points. Additionally, we also calculate the recovered fitting factor of approximately 15,600 nonprecessing simulated quadrupolar harmonic signals lying within our target space. The edge in the bottom panel plot is owing to the duration threshold of 70 ms. Here too, we observe that all of the injections have  $\text{FF} \geq 0.975$  signifying that the bank, summarized in Table I, meets the design expectations.

We note that applying the duration cut leads to only  $\sim 1\%$  fewer templates in the bank but improves the search sensitivity. We demonstrate this in Sec. IV when discussing sensitivity computation. Also, as most of the power in the

TABLE I. Summary of the Template Bank used for the search.

| Parameter                 |                                      |
|---------------------------|--------------------------------------|
| $M_T(1+z)$                | (100 $M_{\odot}$ , 600 $M_{\odot}$ ) |
| $q = m_1/m_2$             | (1, 10)                              |
| $m_i$                     | (40 $M_{\odot}$ , 540 $M_{\odot}$ )  |
| $\chi_{i,z}$              | (-0.998, 0.998)                      |
| $f_{\text{low}}$          | 15 Hz                                |
| Minimum match             | 0.99                                 |
| Minimum template duration | 70 ms                                |
| Waveform model            | SEOBNRv4_ROM                         |
| Number of templates       | 630                                  |

waveforms comes from the merger-ringdown portion of the template, the designed search is sensitive to a wide variety of short-duration signals arising from mergers producing IMBHs [30].

### B. Improved rejection of instrumental transients

As discussed in Sec. II, an important part of search algorithms for detecting GW signals is reducing the number of false alarms caused due to glitches. Relative to previous PyCBC searches, our search uses a lower duration threshold and starts filtering from a comparatively low low-frequency cutoff. This makes it susceptible to more glitches, as illustrated in Fig. 3. There is also an absence of data quality flags for Virgo data segments which seriously hampers the search performance. Hence, while preprocessing the data, we choose to use an autogating threshold of  $50\sigma$  from Gaussian noise to gate out loud  $g(t)$ . The threshold used is half of that used in Nitz *et al.* [4,5]. We note that this can also excise out signals of similar loudness, but observed GW signals from recent BBH mergers have so far not caused such large deviations in whitened detector data.

Even then, while performing the single-detector analysis with our proposed bank, we find many loud background triggers that our signal-noise discriminator does not efficiently differentiate. As illustrated in the top panel of Fig. 6, these triggers have a large  $\tilde{\rho}$  even when having relatively high  $\chi_r^2$  and hence are not efficiently suppressed. Further, some background triggers also have large  $v_s(t)$ , signifying that they are within data segments affected by nonstationary broadband noise. So, we choose to discard any trigger whose  $\chi_r^2$  or  $v_s(t)$  is greater than 10. The choice is motivated by the separation in the distribution of background triggers and the injection population. We use the SEOBNRv4 waveform model to generate quadrupolar quasicircular nonprecessing injections, plotted in red. In blue, we have nonprecessing injections containing the full symphony of a signal and they are generated using the reduced-order representation of the SEOBNRv4HM waveform model [96,97]. We generate the remaining orange marked injections with the NRSur7dq4 [98] waveform



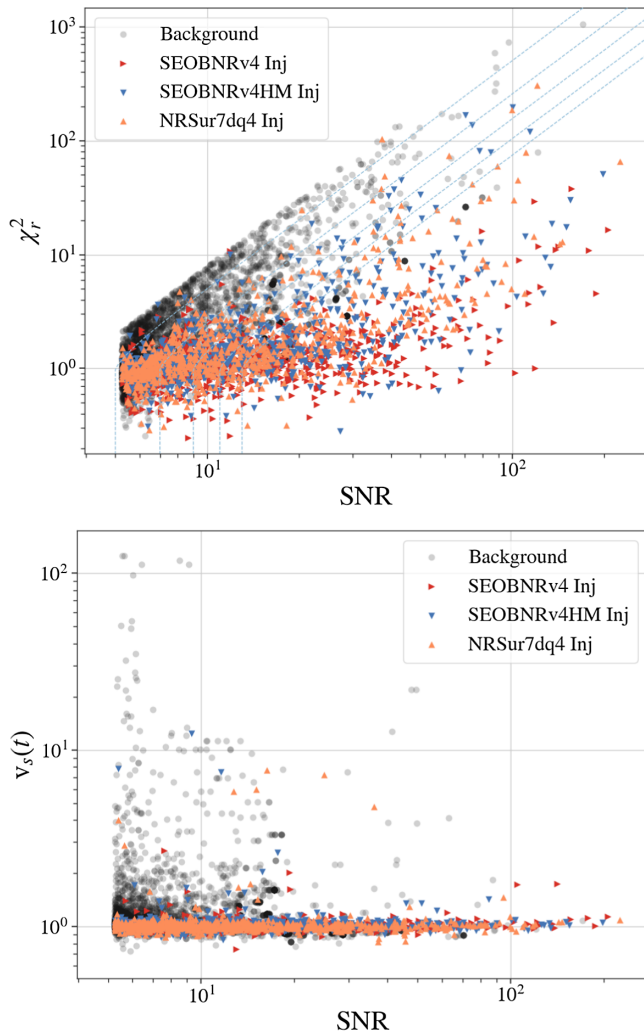


FIG. 6. Top: Distribution of SNR and  $\chi_r^2$  of single-detector triggers from the representative portion of LIGO-Livingston data from the third observing run. Bottom: Distribution of SNR and short-term variation in the PSD,  $v_s(t)$  from the representative portion of same LIGO-Livingston data from the third observing run. In both the panels, the black dots represent the background triggers which are coincidentally observed in Livingston and Virgo data while the other colored markers correspond to the recovered injections. The dashed lines in the top panel are contours of  $\tilde{\rho} = 5, 7, 9, 11, 13$  respectively.

model. They contain both the quadrupolar as well as higher-order harmonics of a precessing binary. We provide additional details of these injections in Sec. IV.

As can be seen, the above constraint rejects approximately  $\sim 3\%$  of the total injections used. But most of these injections have large single-detector SNR ( $\gtrsim 30$ ); given current detector sensitivity, we are unlikely to see IMBH signals with such loud SNRs. The distribution of our injections in luminosity distance, and thus in SNR, is also not astrophysical, giving relatively more high-SNR signals than in a realistic case, thus we expect the fraction of actual signals rejected to be smaller.

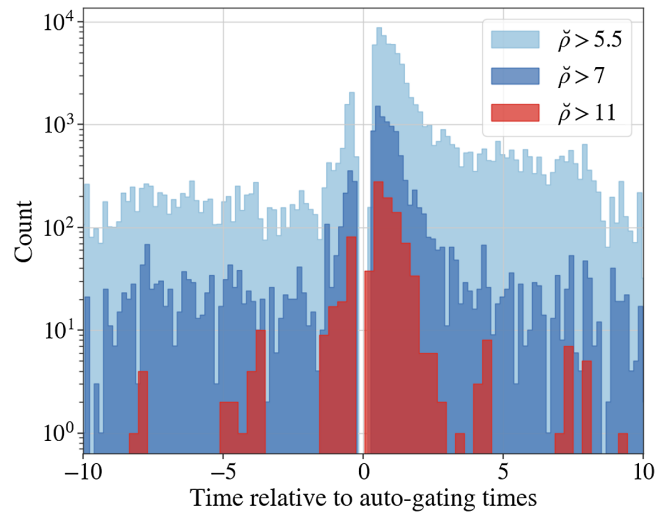


FIG. 7. Histogram of triggers close to autogated glitches in the Livingston detector. Triggers are removed around the central time within a window of  $\pm 0.5$  s around an autogate, but we see here that there is a significantly increased trigger rate around autogates which is not removed.

In addition to these veto choices, we also remove triggers from a wider window around autogated times of loud non-Gaussianity. During preliminary analysis of O3a data, we noticed several instances where the loud glitch prompting the autogate was correlated with nearby, lower frequency glitches that had relatively high SNRs in IMBH templates. We can see the effect of this correlation in the histogram of Fig. 7, which shows an elevated relative rate of triggers with high  $\tilde{\rho}$  in short-duration templates within a few seconds of autogated glitches. This effect is mainly visible in both LIGO detectors' data (increasing up to two orders of

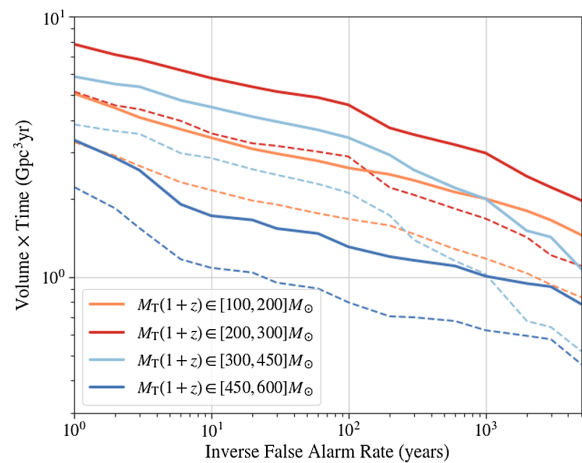


FIG. 8. Plot showing the sensitive volume-time, VT, that the search can survey before [dashed] and after [solid] applying the duration threshold on filter waveforms. We assess the improvement using a representative population of GWs from a simulated generically spinning BHs in a binary system injected in  $\sim 30$  days of O3a data.

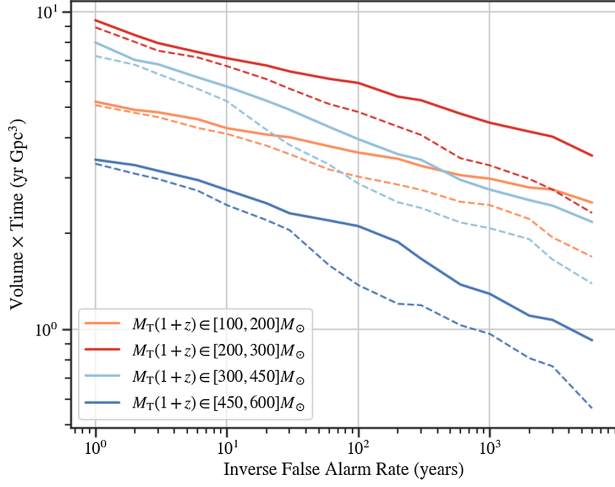


FIG. 9. Plot showing the sensitive VT before [dashed] and after [solid] the removal of additional triggers around autogated times.

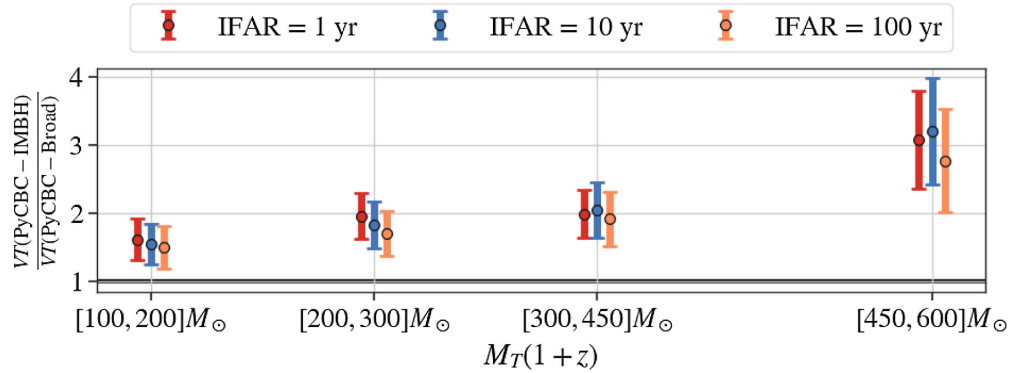
magnitude) being much subtler for Virgo data. As a first approach to this issue, we veto all triggers between one second before and 2.5 seconds after the central autogating times of Advanced LIGO data. The tuning limits were chosen empirically and are subject to future improvements. This refinement yields additional suppression of high-

statistic background events, leading to an increase in sensitive volume-time, which implies that a GW signal belonging to the target space will be detected with an improved significance. We provide details of this study in the next section.

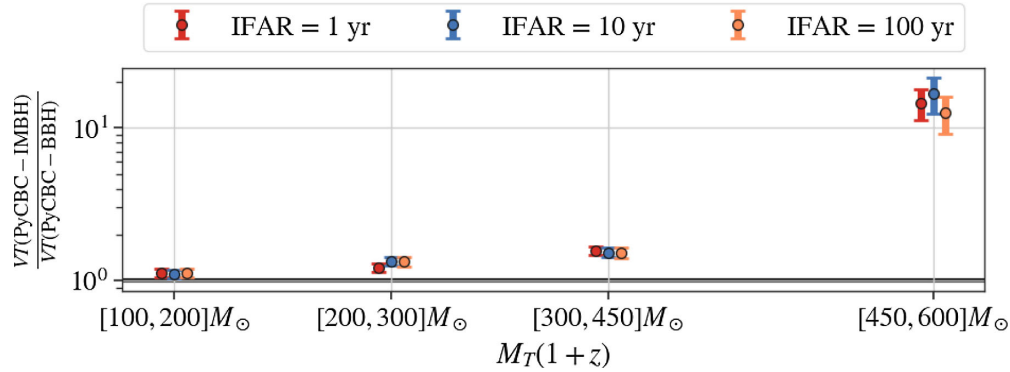
#### IV. EVALUATING THE IMPROVED SEARCH TECHNIQUE

To motivate the various choices made and to quantify the benefit of our restricted analysis as compared to previous PyCBC-based analysis, we resort to an injection campaign; add simulated signals to real or simulated GW data as part of the analysis pipeline without any physical actuation and then estimate the sensitivity of the pipeline to the signals added. For our purpose, we inject the simulated signals in  $\sim 30$  days of O3a data, and we use the sensitive volume-time product, VT, and the sensitive distance as the metrics of the search sensitivity. We estimate these sensitivity metrics using a Monte Carlo method with importance sampling (See Sec. 4 of Usman *et al.* [19]).

As mentioned in Sec. III B, we use SEOBNRv4, SEOBNRv4HM, and NRSur7dq4 as the waveform model for our signals. The simulated binaries have  $M_T(1+z)$  in the range  $[100, 600] M_\odot$  and  $q$  between  $[1, 10]$  except



(a) Plot showing the ratio of sensitive VT, between PyCBC-IMBH and PyCBC-broad search.



(b) Plot showing the ratio of sensitive VT, between PyCBC-IMBH and PyCBC-BBH search.

FIG. 10. Plots showing the improvement in sensitivity at different total mass range at different IFAR. Both the plots show a significant increment in VT, particularly for signals from relatively heavier binaries.

when using precessing injections during which we restrict to  $q \leq 4$  binaries owing to waveform constraints. The binary spins are isotropically distributed in spin directions as allowed by respective waveform models, and we uniformly distribute the simulated signals in  $M_T(1+z)$  and  $m_1/(m_1+m_2)$  space. We also place them uniformly in distances between bounds of 0.5 Gpc and 11 Gpc and distribute them isotropically all over the sky and binary orientation parameters.

Figure 8 compares the sensitivity between searches with and without the duration threshold of filter waveforms discussed in Sec. III A. As can be seen, the search using the duration threshold of 70 ms performs better than the search not using it, and the improvement in sensitive VT ranges between a factor of 1.1 to 1.3 at an IFAR of 100 years, depending on the total mass of the injection.

The search also benefits from the improved gating setting discussed in Sec. III B. Due to this, we see that the search performance improves significantly, especially for comparatively heavier BBHs (see Fig. 9).

Combined, all the proposed optimizations significantly improve the search performance as compared to the already-existing PyCBC-broad search. We illustrate this in Fig. 10(a) where one can see a clear increment in sensitive VT between a factor of  $1.50 \pm 0.31$  and  $2.77 \pm 0.75$  at an IFAR of 100 years depending on the total mass of the system, demonstrating the algorithm's ability to distinguish high-mass signals from noisy transients.

Finally, we also compare the sensitivity of the PyCBC-IMBH search with that of the PyCBC-BBH search used in

Abbott *et al.* [3] in Fig. 10(b) for our set of injections. We find that the sensitive VT at an IFAR of 100 years varies between a factor of  $1.12 \pm 0.07$  for systems with a total mass between  $[100, 200] M_\odot$  to  $12.61 \pm 3.52$  for  $[450, 600] M_\odot$  systems, clearly demonstrating the improvement due to detection optimization for the IMBH binaries. For signals with  $M_T(1+z) > 450 M_\odot$ , the match between the template and signal is quite poor in the PyCBC-BBH search resulting in high  $\chi_r^2$  values, which lead to the majority of even high-SNR signals being entirely missed.

## V. RESULTS

Using this optimized analysis, we perform a search on the O3a data collected during the observation period between April 1, 2019 15:00 UTC and October 1, 2019 15:00 UTC. Table II summarizes the search results and tabulates them in the descending order of significance in terms of IFAR. We analyzed the data separately after subdividing them into five independent blocks of duration  $\sim 30$  days. The bulk of the events listed are known BBHs from below the mass range of the template bank, meaning they have been found with reduced significance in templates with around  $100 M_\odot$ . We only state the lower limits in IFAR of the top six events due to finite-background statistics. Apart from the IFAR and event name, we also report the GPS time,  $\Lambda_s$ , the parameters of the best-matched templates and the value of matched-filter SNR of the signal in the corresponding detector. Among all the signals

TABLE II. Candidate events from the PyCBC-IMBH search for compact binary mergers in O3a data. Candidates are sorted by IFAR evaluated for the entire bank of templates. Finite background statistics limits the IFAR estimated for the top six events, and hence we state their lower limits. We list the parameters of the template associated with each candidate. The masses are given in the detector frame while the SNR corresponds to the matched-filter SNR defined in Eq. (6). Note that the listed parameters are not intended as a rigorous estimation of the source parameters, and the reader should refer to Abbott *et al.* [3] for the same.

| Event name      | GPS time (s) | IFAR(yr)  | $\Lambda_s$ | $\rho_H$ | $\rho_L$ | $\rho_V$ | $m_1 (M_\odot)$ | $m_2 (M_\odot)$ | $\chi_{\text{eff}}$ |
|-----------------|--------------|-----------|-------------|----------|----------|----------|-----------------|-----------------|---------------------|
| GW190521_063405 | 1242459857.5 | >43070.23 | 72.55       | 12.00    | 21.14    | ...      | 60.74           | 40.32           | 0.12                |
| GW190828_063405 | 1251009263.8 | >14260.16 | 47.92       | 10.04    | 11.20    | ...      | 60.38           | 40.36           | 0.63                |
| GW190519_153544 | 1242315362.4 | >9103.08  | 44.02       | 8.37     | 10.51    | ...      | 100.2           | 62.82           | 0.72                |
| GW190706_222641 | 1246487219.3 | >8873.62  | 34.81       | 8.73     | 8.36     | ...      | 149.40          | 44.49           | 0.81                |
| GW190727_060333 | 1248242632.0 | >8105.10  | 31.58       | 8.37     | 8.11     | ...      | 60.32           | 40.27           | 0.14                |
| GW190701_203306 | 1246048404.6 | >5191.03  | 28.64       | 6.01     | 8.76     | 5.85     | 67.68           | 47.02           | -0.48               |
| GW190915_235702 | 1252627040.7 | 2612.10   | 27.87       | 8.41     | 7.50     | ...      | 73.87           | 49.39           | 0.75                |
| GW190517_055101 | 1242107479.8 | 1153.60   | 26.45       | 6.62     | 7.42     | ...      | 58.43           | 48.41           | 0.99                |
| GW190602_175927 | 1243533585.1 | 915.02    | 26.09       | 6.32     | 10.58    | ...      | 132.67          | 52.10           | 0.47                |
| GW190521        | 1242442967.5 | 726.65    | 25.78       | 8.27     | 11.54    | ...      | 195.46          | 43.06           | 0.50                |
| GW190503_185404 | 1240944862.3 | 397.78    | 27.31       | 9.00     | 7.66     | ...      | 60.32           | 40.27           | 0.15                |
| GW190421_213856 | 1239917954.3 | 163.39    | 22.66       | 7.88     | 6.25     | ...      | 58.41           | 53.35           | 0.05                |
| GW190408_181802 | 1238782700.3 | 62.66     | 24.01       | 6.70     | 8.09     | ...      | 82.52           | 40.07           | 0.98                |
| GW190513_205428 | 1241816086.8 | 19.97     | 20.94       | 8.83     | 7.20     | ...      | 55.39           | 46.03           | 0.76                |
| GW190413_134308 | 1239198206.7 | 7.10      | 20.32       | 5.46     | 7.50     | ...      | 79.54           | 56.15           | 0.04                |
| GW190803_022701 | 1248834439.9 | 5.86      | 19.69       | 5.68     | 6.57     | ...      | 57.33           | 42.78           | -0.08               |
| GW190929_012149 | 1253755327.5 | 3.02      | 19.01       | 6.04     | 6.79     | ...      | 123.03          | 40.67           | 0.15                |
| GW190413_052954 | 1239168612.5 | 1.79      | 18.24       | 5.14     | 6.65     | ...      | 71.14           | 43.83           | 0.23                |
| GW190731_140936 | 1248617394.6 | 1.56      | 17.09       | 5.61     | 6.00     | ...      | 58.41           | 53.35           | 0.05                |

reported, only GW190701\_203306 is observed as a three-detector event by this dedicated search.

### A. Redetection of GW190521

GW190521 was first reported in Abbott *et al.* [28] where the dedicated cWB-IMBH search recorded the event with IFAR 4900 years. However, the offline PyCBC-broad and PyCBC-BBH search failed to detect this signal with enough statistical significance, as mentioned before. Our optimized PyCBC-IMBH search recovers the signal with a network SNR of 14.1 and IFAR 727 years which is significantly larger than those reported in Abbott *et al.* [28] for a PyCBC-based search. This change is due to a culmination of all the modifications that we have made for our analysis. However, the improvement in IFAR is not a statistical statement about the search as a whole. The reader should refer to Sec. IV for the same.

## VI. CONCLUSION

We construct and validate a new optimized PyCBC-based search that is capable of detecting both highly redshifted BBH mergers and mergers producing low mass range IMBH remnants. We test the analysis against previous PyCBC-based searches using a set of generically spinning simulated BBHs which we inject at different times in the GW data collected during the O3a run. We find that our analysis can now better separate short-duration GW signals from instrumental artifacts because of targeted usage of filter waveforms, optimization of data quality vetoes and stricter signal-noise discriminators. As a result, the search sensitive VT at an IFAR of 100 years improves by a factor of  $1.50 \pm 0.31$  to  $2.77 \pm 0.75$  as compared to previous PyCBC-based searches, with the largest increase seen for binaries with total masses between [450–600]  $M_{\odot}$ . We also report all the signals with an IFAR > 1.5 years in our search. However, none of these signals are new, and they have already been reported in [3].

While we are primarily concerned here with comparing our optimized search for IMBH binaries with existing PyCBC searches covering or overlapping with this parameter space, ultimately if several such analyses are performed on the same data, it will be necessary to interpret the joint results consistently. Using a simplistic approach of selecting only the most significant (lowest FAR) candidates from any of the searches, the significance of a given candidate will be diluted by so-called ‘trials factors’, as different searches may yield partly or completely independent sets of noise events, resulting in a reduced sensitivity of the joint search at a fixed FAR threshold. An improved strategy may be obtained by accounting for the relative sensitivities of different searches to the target signal population [99]; in the limit where one search’s sensitivity is much larger than others, candidates which are not assigned high significance by that search would be discarded. More generally, the

impact of noise events in a joint search may be mitigated by imposing consistency conditions (derived from studying the recovery of a simulated signal population) on any detection candidates.

The main limitation of this new search is that it currently relies on waveforms that model the dominant harmonic of a quasicircular nonprecessing binary. A near-term development goal is to make the search capable of detecting CBC signals with significant higher-harmonic content and develop newer signal-noise discriminators that can better separate a glitch from a short-duration signal. These improvements will enable the search to detect heavier, asymmetric binaries.

Also, advanced GW detectors are expected to observe thousands of BBH mergers in the next few years, and the detection rates will go up by several orders of magnitude with the advent of third-generation observatories like the Einstein Telescope [45,46] and the Cosmic Explorer [44] and space-based observatories like LISA [50]. Their functioning will enable us to detect GW signals from even further BBH mergers whose detector-frame masses can be several orders of magnitude larger than their source mass. Some of these signals may be due to mergers of IMBHs lying above the pair-instability mass gap [43,49,100,101], and detection of such signals will enable us to understand the origins of these binaries better. Also, the detection of mass-gap remnants or components will provide us with a better understanding of the rate at which they are formed. Hence, there is a continuous need to expand the target search space depending on the detector sensitivity so as not to miss out on potential sources.

The open data is available in the Gravitational Wave Open Science Center (<https://www.gw-openscience.org/>) a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration.

## ACKNOWLEDGMENTS

The authors would like to thank Shasvath Kapadia, Khun Sang Phukon, Tito Dal Canton, and Juan Calderon Bustillo for their detailed comments and useful suggestions. The authors are grateful for the computational resources and data provided by the LIGO Laboratory and supported by National Science Foundation Grants No. PHY-0757058 and No. PHY-0823459. The authors also acknowledge the use of the IUCAA LDG cluster, Sarathi, for computational/numerical work. We acknowledge the usage. K. C. acknowledges the MHRD, Government of India, for the fellowship support. V. V. O. and T. D. acknowledge financial support from Xunta de Galicia (Centro singular de investigación de Galicia accreditation 2019–2022), by European Union ERDF, and by the ‘‘María de Maeztu’’ Units of Excellence Program No. MDM-2016-0692 and the Spanish Research State Agency. A. P. acknowledges the research support from the SERB-Power-fellowship Grant No. SPF/2021/000036, DST, India. G. S. C. D. and



I. W. H. acknowledge the STFC for funding through Grant No. ST/T000333/1. C.M. was supported by the STFC through the DISCnet Centre for Doctoral Training. K. S.

acknowledges the Inter-University Centre of Astronomy and Astrophysics (IUCAA), India, for the fellowship support. This document has LIGO DCC No. LIGO-P2100177.

- 
- [1] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 061102 (2016).
- [2] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **9**, 031040 (2019).
- [3] R. Abbott *et al.*, *Phys. Rev. X* **11**, 021053 (2021).
- [4] A. H. Nitz, C. Capano, A. B. Nielsen, S. Reyes, R. White, D. A. Brown, and B. Krishnan, *Astrophys. J.* **872**, 195 (2019).
- [5] A. H. Nitz, T. Dent, G. S. Davies, S. Kumar, C. D. Capano, I. Harry, S. Mozzon, L. Nuttall, A. Lundgren, and M. Tápai, *Astrophys. J.* **891**, 123 (2019).
- [6] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, *Phys. Rev. D* **101**, 083030 (2020).
- [7] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, *Phys. Rev. D* **100**, 023011 (2019).
- [8] A. H. Nitz, C. D. Capano, S. Kumar, Y.-F. Wang, S. Kastha, M. Schäfer, R. Dhurkunde, and M. Cabero, [arXiv:2105.09151](https://arxiv.org/abs/2105.09151).
- [9] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 074001 (2015).
- [10] F. Acernese *et al.* (Virgo Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).
- [11] S. Klimenko and G. Mitselmakher, *Classical Quantum Gravity* **21**, S1819 (2004).
- [12] S. Klimenko, S. Mohanty, M. Rakhmanov, and G. Mitselmakher, *Phys. Rev. D* **72**, 122002 (2005).
- [13] S. Klimenko, S. Mohanty, M. Rakhmanov, and G. Mitselmakher, *J. Phys. Conf. Ser.* **32**, 12 (2006).
- [14] S. Klimenko, G. Vedovato, M. Drago, G. Mazzolo, G. Mitselmakher, C. Pankow, G. Prodi, V. Re, F. Salemi, and I. Yakushin, *Phys. Rev. D* **83**, 102001 (2011).
- [15] S. Klimenko *et al.*, *Phys. Rev. D* **93**, 042004 (2016).
- [16] C. Messick *et al.*, *Phys. Rev. D* **95**, 042001 (2017).
- [17] S. Sachdev *et al.*, [arXiv:1901.08580](https://arxiv.org/abs/1901.08580).
- [18] C. Hanna *et al.*, *Phys. Rev. D* **101**, 022003 (2020).
- [19] S. A. Usman *et al.*, *Classical Quantum Gravity* **33**, 215004 (2016).
- [20] A. H. Nitz, T. Dent, T. Dal Canton, S. Fairhurst, and D. A. Brown, *Astrophys. J.* **849**, 118 (2017).
- [21] G. S. Davies, T. Dent, M. Tápai, I. Harry, C. McIsaac, and A. H. Nitz, *Phys. Rev. D* **102**, 022004 (2020).
- [22] T. Adams, D. Buskulic, V. Germain, G. Guidi, F. Marion, M. Montani, B. Mours, F. Piergiovanni, and G. Wang, *Classical Quantum Gravity* **33**, 175012 (2016).
- [23] F. Aubin *et al.*, *Classical Quantum Gravity* **38**, 095004 (2021).
- [24] Q. Chu, Low-latency detection and localization of gravitational waves from compact binary coalescences, Ph.D. thesis, The University of Western Australia, 2017.
- [25] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **892**, L3 (2020).
- [26] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **102**, 043015 (2020).
- [27] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **896**, L44 (2020).
- [28] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **125**, 101102 (2020).
- [29] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **900**, L13 (2020).
- [30] J. C. Bustillo, N. Sanchis-Gual, A. Torres-Forné, and J. A. Font, *Phys. Rev. Lett.* **126**, 201101 (2021).
- [31] V. Gayathri, J. Healy, J. Lange, B. O'Brien, M. Szczepanczyk, I. Bartos, M. Campanelli, S. Klimenko, C. Lousto, and R. O'Shaughnessy, [arXiv:2009.05461](https://arxiv.org/abs/2009.05461).
- [32] A. Palmese and C. J. Conscience, *Phys. Rev. Lett.* **126**, 181103 (2021).
- [33] B. Liu and V. Bromm, *Astrophys. J. Lett.* **903**, L40 (2020).
- [34] J. C. Bustillo, N. Sanchis-Gual, A. Torres-Forné, J. A. Font, A. Vajpeyi, R. Smith, C. Herdeiro, E. Radu, and S. H. W. Leong, *Phys. Rev. Lett.* **126**, 081101 (2021).
- [35] A. H. Nitz and C. D. Capano, *Astrophys. J. Lett.* **907**, L9 (2021).
- [36] A. H. Nitz, T. Dal Canton, D. Davis, and S. Reyes, *Phys. Rev. D* **98**, 024050 (2018).
- [37] T. Dal Canton, A. H. Nitz, B. Gadre, G. S. Davies, V. Villa-Ortega, T. Dent, I. Harry, and L. Xiao, [arXiv:2008.07494](https://arxiv.org/abs/2008.07494).
- [38] M. Szczepanczyk *et al.*, *Phys. Rev. D* **103**, 082002 (2021).
- [39] R. Abbott *et al.*, Supplement: Gw190521: A binary black hole merger with a total mass of  $150m_{\odot}$  (2020).
- [40] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **913**, L7 (2021).
- [41] A. Krolak and B. F. Schutz, *Gen. Relativ. Gravit.* **19**, 1163 (1987).
- [42] B. Abbott *et al.* (KAGRA, LIGO Scientific, and Virgo Collaborations), *Living Rev. Relativity* **23**, 3 (2020).
- [43] A. K. Mehta, A. Buonanno, J. Gair, M. C. Miller, E. Farag, R. J. deBoer, M. Wiescher, and F. X. Timmes, [arXiv:2105.06366](https://arxiv.org/abs/2105.06366).
- [44] D. Reitze *et al.*, *Bull. Am. Astron. Soc.* **51**, 035 (2019), <https://inspirehep.net/literature/1743201>.
- [45] M. Maggiore *et al.*, *J. Cosmol. Astropart. Phys.* **03** (2020) 050.
- [46] B. Sathyaprakash *et al.*, *Classical Quantum Gravity* **29**, 124013 (2012); **30**, 079501(E) (2013).
- [47] P. Amaro-Seoane *et al.* (LISA Collaboration), [arXiv:1702.00786](https://arxiv.org/abs/1702.00786).
- [48] M. Safarzadeh and W. M. Farr, *Astrophys. J. Lett.* **883**, L24 (2019).

- [49] A. Mangiagli, M. Bonetti, A. Sesana, and M. Colpi, *Astrophys. J. Lett.* **883**, L27 (2019).
- [50] A. Sesana, *Phys. Rev. Lett.* **116**, 231102 (2016).
- [51] K. Jani, D. Shoemaker, and C. Cutler, *Nat. Astron.* **4**, 260 (2020).
- [52] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **103**, 122002 (2021).
- [53] M. Cabero *et al.*, *Classical Quantum Gravity* **36**, 155010 (2019).
- [54] D. Davis, L. V. White, and P. R. Saulson, *Classical Quantum Gravity* **37**, 145001 (2020).
- [55] A. Nitz *et al.*, gwastro/pycbc: 1.18.0 release of pycbc (2021).
- [56] B. Sathyaprakash and S. Dhurandhar, *Phys. Rev. D* **44**, 3819 (1991).
- [57] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. Creighton, *Phys. Rev. D* **85**, 122006 (2012).
- [58] S. Babak *et al.*, *Phys. Rev. D* **87**, 024033 (2013).
- [59] L. S. Finn and D. F. Chernoff, *Phys. Rev. D* **47**, 2198 (1993).
- [60] L. S. Finn, *Phys. Rev. D* **46**, 5236 (1992).
- [61] L. A. Wainstein and V. D. Zubakov, *Extraction of Signals from Noise* (Dover Publications, New York, 1970), ISBN 0-486-62625-3.
- [62] B. Allen, *Phys. Rev. D* **71**, 062001 (2005).
- [63] A. H. Nitz, *Classical Quantum Gravity* **35**, 035016 (2018).
- [64] S. Mozzon, L. Nuttall, A. Lundgren, T. Dent, S. Kumar, and A. Nitz, *Classical Quantum Gravity* **37**, 215014 (2020).
- [65] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **6**, 041015 (2016); **8**, 039903(E) (2018).
- [66] S. Roy, A. S. Sengupta, and P. Ajith, *Phys. Rev. D* **99**, 024048 (2019).
- [67] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **100**, 064064 (2019).
- [68] J. C. Bustillo, F. Salemi, T. Dal Canton, and K. P. Jani, *Phys. Rev. D* **97**, 024016 (2018).
- [69] K. Chandra, V. Gayathri, J. C. Bustillo, and A. Pai, *Phys. Rev. D* **102**, 044035 (2020).
- [70] I. Harry, J. C. Bustillo, and A. Nitz, *Phys. Rev. D* **97**, 023004 (2018).
- [71] C. Mills and S. Fairhurst, *Phys. Rev. D* **103**, 024042 (2021).
- [72] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), arXiv:2105.15120.
- [73] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein, and B. Stephens, *J. Phys. Conf. Ser.* **610**, 012021 (2015).
- [74] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *SoftwareX* **13**, 100658 (2021).
- [75] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, *Phys. Rev. D* **49**, 6274 (1994).
- [76] A. Taracchini, A. Buonanno, Y. Pan, T. Hinderer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, A. H. Mroué, H. P. Pfeiffer *et al.*, *Phys. Rev. D* **89**, 061502 (2014).
- [77] A. Bohé *et al.*, *Phys. Rev. D* **95**, 044028 (2017).
- [78] M. Maggiore, *Gravitational Waves. Vol. 1: Theory and Experiments*, Oxford Master Series in Physics (Oxford University Press, New York, 2007).
- [79] S. Soni *et al.*, arXiv:2103.12104.
- [80] B. J. Owen, *Phys. Rev. D* **53**, 6749 (1996).
- [81] B. J. Owen and B. Sathyaprakash, *Phys. Rev. D* **60**, 022002 (1999).
- [82] S. Babak, R. Balasubramanian, D. Churches, T. Cokelaer, and B. S. Sathyaprakash, *Classical Quantum Gravity* **23**, 5477 (2006).
- [83] T. Cokelaer, *Phys. Rev. D* **76**, 102004 (2007).
- [84] D. A. Brown, I. Harry, A. Lundgren, and A. H. Nitz, *Phys. Rev. D* **86**, 084017 (2012).
- [85] I. W. Harry, A. H. Nitz, D. A. Brown, A. P. Lundgren, E. Ochsner, and D. Keppel, *Phys. Rev. D* **89**, 024010 (2014).
- [86] T. Dal Canton *et al.*, *Phys. Rev. D* **90**, 082004 (2014).
- [87] I. W. Harry, S. Fairhurst, and B. S. Sathyaprakash, *Classical Quantum Gravity* **25**, 184027 (2008).
- [88] S. Babak, *Classical Quantum Gravity* **25**, 195011 (2008).
- [89] I. W. Harry, B. Allen, and B. Sathyaprakash, *Phys. Rev. D* **80**, 104014 (2009).
- [90] G. M. Manca and M. Vallisneri, *Phys. Rev. D* **81**, 024004 (2010).
- [91] S. Roy, A. S. Sengupta, and N. Thakor, *Phys. Rev. D* **95**, 104045 (2017).
- [92] T. A. Apostolatos, *Phys. Rev. D* **52**, 605 (1995).
- [93] P. Ajith *et al.*, *Phys. Rev. Lett.* **106**, 241101 (2011).
- [94] L. Santamaria *et al.*, *Phys. Rev. D* **82**, 064016 (2010).
- [95] A. Buonanno, Y.-B. Chen, and M. Vallisneri, *Phys. Rev. D* **67**, 104025 (2003); **74**, 029904(E) (2006).
- [96] R. Cotesta, A. Buonanno, A. Bohé, A. Taracchini, I. Hinder, and S. Ossokine, *Phys. Rev. D* **98**, 084028 (2018).
- [97] R. Cotesta, S. Marsat, and M. Pürrer, *Phys. Rev. D* **101**, 124040 (2020).
- [98] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, *Phys. Rev. Research* **1**, 033015 (2019).
- [99] R. Biswas *et al.*, *Phys. Rev. D* **85**, 122009 (2012).
- [100] J. M. Ezquiaga and D. E. Holz, *Astrophys. J. Lett.* **909**, L23 (2021).
- [101] A. Tanikawa, H. Susa, T. Yoshida, A. A. Trani, and T. Kinugawa, *Astrophys. J.* **910**, 30 (2021).