# Searching for gravitational waves from binary coalescence

S. Babak,[1,2] R. Biswas,[3,4] P. R. Brady,[3] D. A. Brown,[5,6,3] K. Cannon,[7,6,3] C. D. Capano,[5,8] J. H. Clayton,[3]
T. Cokelaer,[1] J. D. E. Creighton,[3] T. Dent,[1,9] A. Dietz,[10,1,11] S. Fairhurst,[1,6,3] N. Fotopoulos,[6,3] G. González,[11]
C. Hanna,[12,6,11] I. W. Harry,[1,5] G. Jones,[1] D. Keppel,[9,13,6] D. J. A. McKechan,[1] L. Pekowsky,[5,14] S. Privitera,[6]
C. Robinson,[1,8] A. C. Rodriguez,[11] B. S. Sathyaprakash,[1] A. S. Sengupta,[15,6,1]
M. Vallisneri,[16] R. Vaulin,[17,3] and A. J. Weinstein[6]

[1]*Cardiff University, Cardiff CF24 3AA, United Kingdom*
[2]*Max-Planck-Institut für Gravitationsphysik—Albert-Einstein-Institut Am Mühlenberg 1,*
*14476 Potsdam-Golm, Germany*
[3]*University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201, USA*
[4]*The University of Texas at Brownsville and Texas Southmost College, Brownsville, Texas 78520, USA*
[5]*Syracuse University, Syracuse, New York 13244, USA*
[6]*LIGO Laboratory, California Institute of Technology, Pasadena, California 91125, USA*
[7]*Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario M5S 3H8, Canada*
[8]*University of Maryland, College Park, Maryland 20742, USA*
[9]*Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany*
[10]*The University of Mississippi, University, Mississippi 38677, USA*
[11]*Louisiana State University, Baton Rouge, Louisiana 70803, USA*
[12]*Perimeter Institute for Theoretical Physics, Ontario N2L 2Y5, Canada*
[13]*Leibniz Universität Hannover, D-30167 Hannover, Germany*
[14]*Center for Relativistic Astrophysics and School of Physics, Georgia Institute of Technology,*
*Atlanta, Georgia 30332, USA*
[15]*IIT Gandhinagar, VGEC Complex, Chandkheda Ahmedabad 382424, Gujarat, India*
[16]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA*
[17]*LIGO Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

We describe the implementation of a search for gravitational waves from compact binary coalescences in LIGO and Virgo data. This all-sky, all-time, multidetector search for binary coalescence has been used to search data taken in recent LIGO and Virgo runs. The search is built around a matched filter analysis of the data, augmented by numerous signal consistency tests designed to distinguish artifacts of non-Gaussian detector noise from potential detections. We demonstrate the search performance using Gaussian noise and data from the fifth LIGO science run and demonstrate that the signal consistency tests are capable of mitigating the effect of non-Gaussian noise and providing a sensitivity comparable to that achieved in Gaussian noise.

## I. INTRODUCTION

Coalescing binaries of compact objects such as neutron-stars (NSs) and stellar-mass black holes (BHs) are promising gravitational-wave (GW) sources for ground-based, kilometer-scale interferometric detectors such as LIGO [1], Virgo [2], and GEO600 [3], which are sensitive to waves of frequencies between tens and thousands of Hertz. Numerous searches for these signals were performed on data from the six LIGO and GEO science runs (S1–S6) and from the four Virgo science runs (VSR1–4) [4–14].

Over time, the software developed to run these searches and evaluate the significance of results evolved into a sophisticated pipeline, known as IHOPE. An early version of the pipeline was described in Ref. [15]. In this paper, we describe the IHOPE pipeline in detail for the first time. A large fraction of the pipeline has been developed to mitigate the effects of non-Gaussian artifacts in the output of the detectors. We demonstrate that the analysis successfully eliminates virtually all of the effects of non-Gaussian noise and achieves a sensitivity close to what is possible in simulated Gaussian data. This is a significant result as it means that the achieved [1,2] and predicted [16,17] detector sensitivities can be used to infer the rate of expected signals [18], even though the detector data is not Gaussian and stationary.

Compact binary coalescences (CBCs) consist of three dynamical phases: a gradual *inspiral*, which is described accurately by the post-Newtonian approximation to the Einstein equations [19]; a nonlinear *merger*, which can be modeled with numerical simulations (see Refs. [20–22] for recent reviews); and the final ringdown of the merged object to a quiescent state [23]. For the lighter NS-NS systems, only the inspiral lies within the band of detector sensitivity. Since CBC waveforms are

well modeled, it is natural to search for them by matched-filtering the data with banks of theoretical *template* waveforms [24].

The most general CBC waveform is described by seventeen parameters, which include the masses and intrinsic spins of the binary components, as well as the location, orientation, and orbital elements of the binary. It is not feasible to perform a search by placing templates across such a high-dimensional parameter space. However, it is astrophysically reasonable to neglect orbital eccentricity [25,26]; furthermore, CBC waveforms that omit the effects of spins have been shown to have acceptable phase overlaps with spinning-binary waveforms, and are therefore suitable for the purpose of detecting CBCs, if not to estimate their parameters accurately [27].

Thus, CBC searches so far have relied on nonspinning waveforms that are parameterized only by the component masses, by the location and orientation of the binary, by the initial orbital phase, and by the time of coalescence. Among these parameters, the masses determine the intrinsic phasing of the waveforms, while the others affect only the relative amplitudes, phases, and timing observed at multiple detector sites [28]. Since we filter the data from each detector separately, it follows that templates need to be placed only across the two-dimensional parameter space spanned by the masses [28]. Even so, past CBC searches have required many thousands of templates to cover their target ranges of masses. (We note that IHOPE could be extended easily to nonprecessing binaries with aligned spins. However, more general precessing waveforms would prove more difficult, as discussed in Refs. [29–32].)

In the context of stationary Gaussian noise, matched-filtering would directly yield the most statistically significant detection candidates. In practice, environmental and instrumental disturbances cause non-Gaussian noise transients (*glitches*) in the data. Searches must distinguish between the candidates, or *triggers*, resulting from glitches and those resulting from true GWs. The techniques developed for this challenging task include *coincidence* (signals must be observed in two or more detectors with consistent mass parameters and times of arrival), *signal-consistency* tests (which quantify how much a signal's amplitude and frequency evolution is consistent with theoretical waveforms [33]), and *data quality vetoes* (which identify time periods when the detector glitch rate is elevated). We describe these in detail later.

The *statistical significance* after the consistency tests have been applied is then quantified by computing the false alarm probability (FAP) or false alarm rate (FAR) of each candidate; we define both below. For this, the background of noise-induced candidates is estimated by performing *time shifts*, whereby the coincidence and consistency tests are run after imposing relative time offsets on the data from different detectors. Any consistent candidate found in this way must be due to noise; furthermore, if the noise of different detectors is uncorrelated, the resulting background rate is representative of the rate at zero shift.

The *sensitivity* of the search to CBC waves is estimated by adding simulated signals (*injections*) to the detector data, and verifying which are detected by the pipeline. With this diagnostic we can tune the search to a specific class of signals (e.g., a region in the mass plane), and we can give an astrophysical interpretation, such as an upper limit on CBC rates [34], to completed searches.

As discussed below, commissioning a GW search with the IHOPE pipeline requires a number of parameter tunings, which include the handling of coincidences, the signal-consistency tests, and the final ranking of triggers. To avoid biasing the results, IHOPE permits a blind analysis: the results of the non-time-shifted analysis can be sequestered, and tuning performed using only the injections and time-shifted results. Later, with the parameter tunings frozen, the non-time-shifted results can be unblinded to reveal the candidate GW events.

This paper is organized as follows. In Sec. II we provide a brief overview of the IHOPE pipeline, and describe its first few stages (data conditioning, template placement, filtering, coincidence), which would be sufficient to implement a search in Gaussian noise but not, as we show, in real detector data. In Sec. III we describe the various techniques that have been developed to eliminate the majority of background triggers due to non-Gaussian noise. In Sec. IV we describe how the IHOPE results are used to make astrophysical statements about the presence or absence of signals in the data, and to put constraints on CBC event rates. Last, in Sec. V we discuss ways in which the analysis can be enhanced to improve sensitivity, reduce latency, and find use in the advanced-detector era.

A basic matched filter based search of LIGO and Virgo data [35] would yield background noise events, due to non-Gaussian artefacts in the data, over one hundred times louder than expected in Gaussian noise. This would reduce the rate of observable signals over six orders of magnitude below expectations. In order to demonstrate how the various stages of the IHOPE analysis reduce or eliminate the effects of transients in the data, we show representative IHOPE output, taken from a search of one month of LIGO data from the S5 run (the third month in Ref. [12]), when all three LIGO detectors (but not Virgo) were operational. For comparison, we also run the same search on Gaussian noise generated at the design sensitivity of the Laser Interferometer Gravitational-wave Observatory (LIGO) detectors (using the same data times as the real data). The search focused on low-mass CBC signals with component masses $>1 M_\odot$ and total mass $<25 M_\odot$. Where we perform GW-signal injections (see Sec. IV C), we adopt a population of binary-neutron-star inspirals, uniformly distributed in distance, coalescence time, sky position, and orientation angles. By the end of the analysis, sensitive search volume is only 30% less than Gaussian-noise expectations.

## II. MATCHED FILTER SEARCH WITH MULTIPLE-DETECTOR COINCIDENCE

The waveform from a nonspinning CBC, as observed by a ground-based detector and neglecting higher-order amplitude corrections, can be written as[1]

$$h(\tau) = h_0(\tau) \cos\Phi_0 + h_{\pi/2}(\tau) \sin\Phi_0, \qquad (1)$$

with

$$\begin{pmatrix} h_0(\tau) \\ h_{\pi/2}(\tau) \end{pmatrix} = A f(\tau)^{2/3} \begin{pmatrix} \cos(\Phi(\tau)) \\ -\sin(\Phi(\tau)) \end{pmatrix}. \qquad (2)$$

Here, $\tau = t_c - t$ is a time variable relative to the coalescence time, $t_c$. The constant amplitude $A$ and phase $\Phi_0$, between them, depend on all the binary parameters: masses, sky location and distance, orientation, and (nominal) orbital phase at coalescence. The time-dependent frequency $f(\tau)$ and phase $\Phi(\tau)$, calculated through the post-Newtonian formalism [19], depend only on the component masses[2] and on the absolute time of coalescence. At leading order,

$$f(\tau) = \frac{c^3}{8\pi G M} \left( \frac{5GM}{c^3 \eta \tau} \right)^{3/8}, \qquad (3)$$

where $M$ is the total mass and $\eta$ is the symmetric mass ratio $\eta = m_1 m_2 / M^2$. This sets the timescale of the problem: a binary with $1 M_\odot$ components evolves from 40 Hz to merger in 45 s.

At its core, the IHOPE pipeline utilizes matched filtering to search for signals in the detector data. The squared signal-to-noise ratio (SNR) $\rho^2$ for the data $s$ and template $h$, analytically maximized over $A$ and $\Phi_0$, is given by

$$\rho^2 = \frac{(s|h_0)^2 + (s|h_{\pi/2})^2}{(h_0|h_0)}, \qquad (4)$$

where $(\cdot|\cdot)$ is the standard noise-weighted inner product

$$(a|b) = 4 \, \text{Re} \int_{f_{\text{low}}}^{f_{\text{high}}} \frac{\tilde{a}^*(f)\tilde{b}(f)}{S_n(f)} df, \qquad (5)$$

with $S_n(f)$ the one-sided detector-noise power spectral density (PSD). Here we assume that $\tilde{h}_{\pi/2}(f) = i\tilde{h}_0(f)$, which is identically true for waveforms defined in the frequency domain with the stationary-phase approximation [36], and approximately true for all slowly evolving CBC waveforms.

The challenge then is to search over the full space of masses and coalescence times in multiple detectors. The stages of the IHOPE pipeline are presented schematically in Fig. 1, and are described in detail in Secs. II, III, and IV of this paper. First, the *science data* to be analyzed is identified and split into 2048 s *blocks*, and the power spectral density is estimated for each block (see Sec. II A).[3] Next, a template bank is constructed independently for each detector and each block (Sec. II B). The data blocks are matched-filtered against each bank template, and the times when the SNR rises above a set threshold are recorded as *triggers* (Sec. II C). The triggers from each detector are then compared to identify *coincidences*—that is, triggers that occur in two or more detectors with similar masses and compatible times (Sec. II D).

If detector noise was Gaussian and stationary, we could proceed directly to the statistical interpretation of the triggers. Unfortunately, non-Gaussian noise glitches generate both an increase in the number of low-SNR triggers as well as high-SNR triggers that form long tails in the distribution of SNRs. The increase in low-SNR triggers will cause an small, but inevitable, reduction in the sensitivity of the search. It is, however, vital to distinguish the high-SNR background triggers from those caused by real GW signals. To achieve this, the coincident triggers are used to generate a reduced template bank for a second round of matched-filtering in each detector (see the beginning of Sec. III). This time, signal-consistency tests are performed on each trigger to help differentiate background from true signals (Secs. III A and III B). These tests are computationally expensive, so we reserve them for this second pass. Single-detector triggers are again compared for coincidence, and the final list is clustered and ranked (Sec. III E), taking into account signal consistency, amplitude consistency among detectors (Sec. III C), as well as the times in which the detectors were not operating optimally (Sec. III D). These steps leave coincident triggers that have a quasi-Gaussian distribution; they can now be evaluated for statistical significance, and used to derive event-rate upper limits in the absence of a detection.

To do this, the steps of the search that involve coincidence are repeated many times, artificially shifting the time stamps of triggers in different detectors, such that no true GW signal would actually be found in coincidence (Sec. IV A). The resulting *time-shift* triggers are used to calculate the FAR of the *in-time* (zero-shift) triggers. Those with FAR lower than some threshold are the candidate GW signals (Sec. IV B). Simulated GW signals are then injected into the data, and by observing which injections are recovered as triggers with FAR lower than some threshold, we can characterize detection efficiency as a function

---

[1]This is valid under the assumption that the detector response, determined by the position of the source relative to the detector, is constant over the duration of the signal. For compact binaries in ground based detectors, this is reasonable as the sources are in band for at most minutes (in the advanced detectors).

[2]Strictly, the waveforms depend upon the redshifted component masses $(1 + z)m_{1,2}$. Note, however, that this does not affect the search as one can simply replace the masses by their redshifted values.

[3]The choice of block length, and other analysis durations, is determined by the length of the templates which ranges from 45 s for a CBC with components of $1 M_\odot$ or greater down to a fraction of a second for higher mass signals.
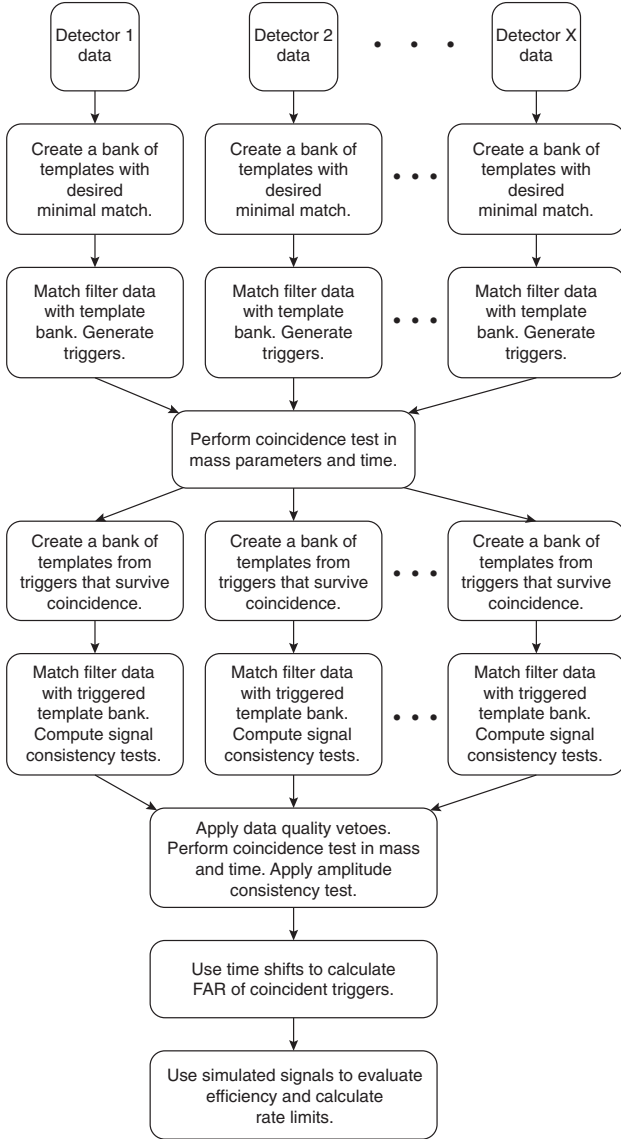
FIG. 1. Structure of the IHOPE pipeline. Template bank creation, matched filtering, and coincidence testing are described in Sec. II. To mitigate the effect of non-Gaussian noise, signal, and amplitude consistency tests are performed on triggers and data quality vetoes are applied; the calculation and applitcation of these tests are described in Sec. III. Surviving triggers are ranked by their false alarm rate FAR, and the efficiency of the pipeline is evaluated using simulated signals. These final steps are described in Sec. IV.

of distance and other parameters (Sec. IV C), providing an astrophysical interpretation for the search. Together with the FARs of the loudest triggers, the efficiency yields the upper limits (Sec. IV D).

## A. Data segmentation and conditioning, power-spectral-density generation

As a first step in the pipeline, IHOPE identifies the stretches of detector data that should be analyzed: for each detector, such *science segments* are those for which the detector was locked (i.e., interferometer laser light was resonant in Fabry-Perot cavities [1]), no other experimental work was being performed, and the detector's "science mode" was confirmed by a human "science monitor." IHOPE builds a list of science-segment times by querying a network-accessible database that contains this information for all detectors.

The LIGO and Virgo GW-strain data are sampled at 16 384 and 20 000 Hz, respectively, but both are down-sampled to 4096 Hz prior to analysis [15], since at frequencies above 1 to 2 kHz detector noise overwhelms any likely CBC signal. This sampling rate sets the Nyquist frequency at 2048 Hz; to prevent aliasing, the data are preconditioned with a time-domain digital filter with low-pass cutoff at the Nyquist frequency [15]. While CBC signals extend to arbitrarily low frequencies, detector sensitivity degrades rapidly, so very little GW power could be observed below 40 Hz. Therefore, we usually suppress signals below 30 Hz with two rounds of 8th-order Butterworth high-pass filters, and analyze data only above 40 Hz.

Both the low- and high-pass filters corrupt the data at the start and end of a science segment, so the first and last few seconds of data (typically 8 s) are discarded after applying the filters. Furthermore, SNRs are computed by correlating templates with the (noise-weighted) data stream, which is only possible if a stretch of data of at least the same length as the template is available. Altogether, the data are split into 256 s segments, and the first and last 64 s of each segment are not used in the search. Neighboring segments are overlapped by 128 s to ensure that all available data are analyzed.

The strain PSD is computed separately for every 2048 s *block* of data (consisting of 15 overlapping 256 s segments). The blocks themselves are overlapped by 128 s. The block PSD is estimated by taking the median [37] (in each frequency bin) of the segment PSDs, ensuring robustness against noise transients and GW signals (whether real or simulated). The PSD is used in the computation of SNRs, and to set the spacing of templates in the banks. Science segments shorter than 2064 s (2048 s block length and 16 s to account for the data corruption due to low- and high-pass filtering) are not used in the analysis, since they cannot provide an accurate PSD estimate.

### B. Template-bank generation

The template bank must be sufficiently dense in parameter space to ensure a minimal loss of matched-filtering SNR for any CBC signal within the mass range of interest; however, the computational cost of a search is proportional to the number of templates in a bank. The method used to place templates must balance these considerations. This problem is well explored for nonspinning CBC signals [38–44], for which templates need only be placed across

the two-dimensional *intrinsic*-parameter space spanned by the two component masses. The other *extrinsic* parameters enter only as amplitude scalings or phase offsets, and the SNR can be maximized analytically over these parameters after filtering by each template.

Templates are placed in parameter space so that the *match* between any GW signal and the best-fitting template is better than a *minimum match* MM (typically 97%). The match between signals $h$ with parameter vectors $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ is defined as

$$\max_{t_2, \Phi_2} \frac{(h(\boldsymbol{\xi}_1)|h(\boldsymbol{\xi}_2))}{\sqrt{(h(\boldsymbol{\xi}_1)|h(\boldsymbol{\xi}_1))}\sqrt{(h(\boldsymbol{\xi}_2)|h(\boldsymbol{\xi}_2))}}, \qquad (6)$$

where $\boldsymbol{\xi}_i$ encodes all of the parameters of the signal (specifically the masses $m_1$ and $m_2$; the coalescence time $t_c$ and phase $\Phi_c$ and the amplitude $A$), and $t_2$ and $\Phi_2$ are the time and phase of coalescence of the waveform $h_2$. The MM represents the worst-case reduction in matched-filtering SNR due to using a discrete bank of templates, and correspondingly the worst-case reduction in the maximum detection distance of a search. Thus, under the assumption of sources uniformly distributed in volume, the reduction in detection rate due to template-bank discreteness is bounded by $\mathrm{MM}^3$, or $\simeq 10\%$ for $\mathrm{MM} = 97\%$.

It is computationally expensive to obtain template mismatches for pairs of templates using Eq. (5), so an approximation based on a parameter-space *metric* is used instead:

$$1 - (h(\boldsymbol{\xi})|h(\boldsymbol{\xi} + \delta\boldsymbol{\xi})) \simeq \sum_{ij} g_{ij}(\boldsymbol{\xi})\delta\xi^i \delta\xi^j, \qquad (7)$$

where

$$g_{ij}(\boldsymbol{\xi}) = -\frac{1}{2}\frac{\partial^2(h(\boldsymbol{\xi})|h(\boldsymbol{\xi}))}{\partial\xi^i \partial\xi^j}. \qquad (8)$$

The approximation holds as long as the higher order terms in the expansion can be disregarded. When generating a template bank, we project out the dependence on the unwanted parameters [39], namely, $A$, $\Phi$, and $t_c$, to generate a metric on the subspace of masses. It is also convenient to choose parameters (i.e., coordinates $\boldsymbol{\xi}$) that make the metric almost constant, such as the "chirp times" $\tau_0$, $\tau_3$ given by [45]

$$\tau_0 = \frac{5}{256\pi f_{\mathrm{low}}\eta}\left(\frac{\pi G M f_{\mathrm{low}}}{c^3}\right)^{-5/3}, \qquad (9)$$

$$\tau_3 = \frac{5}{8 f_{\mathrm{low}}\eta}\left(\frac{\pi G M f_{\mathrm{low}}}{c^3}\right)^{-2/3}. \qquad (10)$$

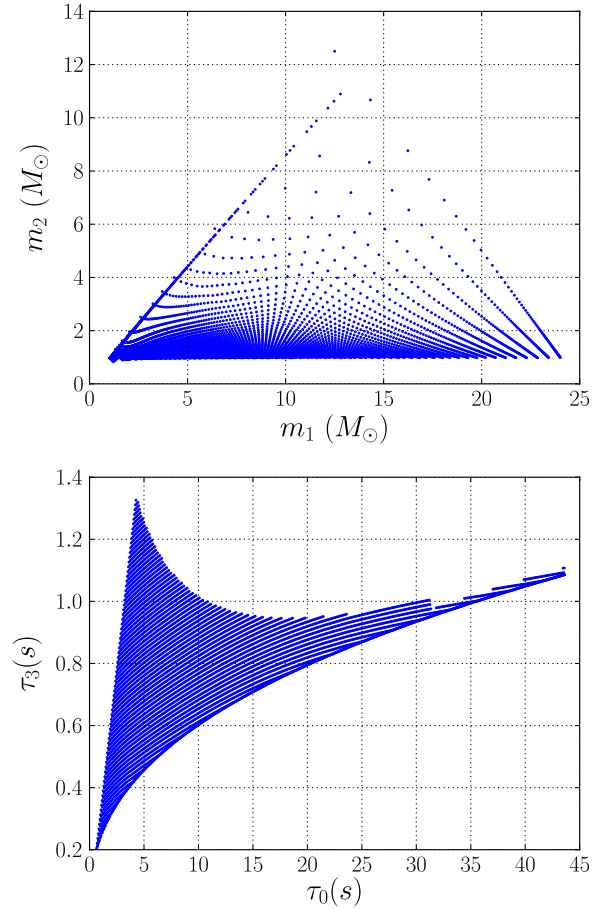Here $f_{\mathrm{low}}$ is the lower frequency cutoff used in the template generation, and $\tau_0$ is the template length, at leading order.



FIG. 2 (color online). A typical template bank for a low-mass CBC inspiral search, as plotted in $m_1$-$m_2$ space (top panel) and $\tau_0$-$\tau_3$ space (bottom panel). Templates are distributed more evenly over $\tau_0$ and $\tau_3$, since the parameter-space metric is approximately flat in those coordinates.

For the S5–S6 and VSR1–3 CBC searches, templates were placed on a regular hexagonal lattice in $\tau_0$-$\tau_3$ space [38], sized so that MM would be 97% [11,12,14]. The metric was computed using inspiral waveforms at the second post-Newtonian (2PN) order in phase. Higher-order templates are now used in searches (some including merger and ringdown), but not for template placement; work is ongoing to implement that. Figure 2 shows a typical template bank in both $m_1$-$m_2$ and $\tau_0$-$\tau_3$ space for the low-mass CBC search. For a typical data block, the bank contains around 6000 templates (Virgo, which has a flatter noise PSD, requires more).

As Eqs. (8) and (5) imply, the metric depends on both the detector-noise PSD and the frequency limits $f_{\mathrm{low}}$ and $f_{\mathrm{high}}$. We set $f_{\mathrm{low}}$ to 40 Hz, while $f_{\mathrm{high}}$ is chosen naturally as the frequency at which waveforms end[4] (200 Hz and 2 kHz for

---

[4]For post-Newtonian waveforms, this is commonly taken to be the frequency of the innermost stable circular orbit of a point particle around a black hole of the same total mass, see Ref. [28] for details.

the highest- and lowest-mass signals, respectively). The PSD changes between data blocks, but usually only slightly, so template banks stay roughly constant over time in a data set.

### C. Matched filtering

The central stage of the pipeline is the matched filtering of detector data with bank templates, resulting in a list of *triggers* that are further analyzed downstream. This stage was described in detail in Ref. [28]; here we sketch its key features.

The maximized statistic $\rho^2$ of Eq. (4) is a function only of the component masses and the time of coalescence $t_c$. Now, a time shift can be folded in the computation of inner products by noting that $g(\tau) = h(\tau - \Delta t_c)$ transforms to $\tilde{g}(f) = e^{i2\pi f \Delta t_c} \tilde{h}(f)$; therefore, the SNR can be computed as a function of $t_c$ by the inverse Fourier transform (a complex quantity)

$$(s|h)(\Delta t_c) = 4 \int_{f_{\text{low}}}^{f_{\text{high}}} \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)} e^{2\pi i f \Delta t_c} df. \qquad (11)$$

Furthermore, if $\tilde{h}_{\pi/2}(f) = i\tilde{h}_0(f)$ then Eq. (11), computed for $h = h_0$, yields $(s|h_0)(\Delta t_c) + i(s|h_{\pi/2})(\Delta t_c)$.

The IHOPE matched-filtering engine implements the discrete analogs of Eqs. (4) and (11) [28] using the efficient FFTW library [46]. The resulting SNRs are not stored for every template and every possible $t_c$; instead, we only retain triggers that exceed an empirically determined threshold (typically 5.5), and that corresponds to maxima of the SNR time series—that is, a trigger above the threshold is kept only if there are no triggers with higher SNR within a predefined time window, typically set to the length of the template (this is referred to as *time clustering*).

For a single template and time and for detector data consisting of Gaussian noise, $\rho^2$ follows a $\chi^2$ distribution with two degrees of freedom, which makes a threshold of 5.5 seem rather large: $p(\rho > 5.5) = 2.7 \times 10^{-7}$. However, we must account for the fact that we consider a full template bank and maximize over time of coalescence: the bank makes for, conservatively, a thousand independent trials at any point in time, while the rapid falloff of the template autocorrelation (as demonstrated in Fig. 7) renders trials separated by 0.1 s in time essentially independent. Therefore, we expect to see a few triggers above this threshold already in a few hundred seconds of Gaussian noise, and a large number in a year of observing time. Furthermore, since the data contain many non-Gaussian noise transients, the trigger rate will be even higher. In Fig. 3 we show the distribution of triggers as a function of SNR in a month of simulated Gaussian noise in blue (dark gray) and real data in red (light gray) from LIGO's fifth science run (S5). The difference between the two is clearly noticeable, with a tail of high SNR triggers extending to SNRs well over 1000 in real data.
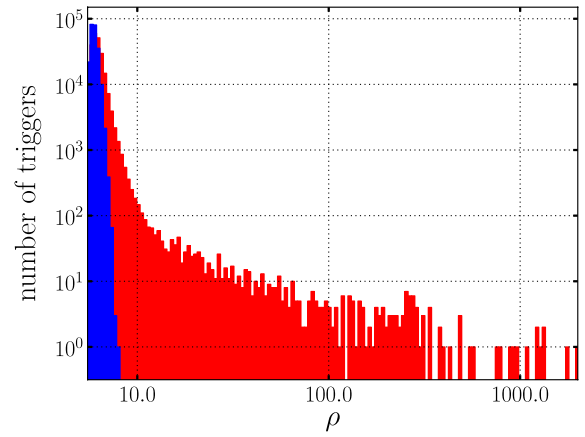


FIG. 3 (color online). Distribution of single detector trigger SNRs in a month of simulated Gaussian noise in blue (dark gray) and real S5 LIGO data in red (light gray) from the Hanford interferometer H1.

It is useful to not just cluster in time, but also across the template bank. When the SNR for a template is above threshold, it is probable that it will be above threshold also for many neighboring templates, which encode very similar waveforms. The IHOPE pipeline selects only one (or a few) triggers for each event (be it a GW or a noise transient), using one of two algorithms. In *time-window* clustering, the time series of triggers from all templates is split into windows of fixed duration; within each window, only the trigger with the largest SNR is kept. This method has the advantage of simplicity, and it guarantees an upper limit on the trigger rate. However, a glitch that creates triggers in one region of parameter space can mask a true signal that creates triggers elsewhere. This problem is remedied in *TrigScan* clustering [47], whereby triggers are grouped by both time and recovered (template) masses, using the parameter-space metric to define their proximity (for a detailed description see Ref. [48]). However, when the data are particularly glitchy *TrigScan* can output a number of triggers that can overwhelm subsequent data processing such as coincident trigger finding.

### D. Multidetector coincidence

The next stage of the pipeline compares the triggers generated for each of the detectors, and retains only those that are seen in *coincidence*. Loosely speaking, triggers are considered coincident if they occurred at roughly the same time, with similar masses; see Ref. [49] for an exact definition of coincidence as used in recent CBC searches. To wit, the "distance" between triggers is measured with the parameter-space metric of Eq. (8), maximized over the signal amplitude $A$ and phase $\Phi_0$. Since different detectors at different times have different noise PSDs and therefore metrics, we construct a constant-metric-radius ellipsoid in $\tau_0$-$\tau_3$-$t_c$ space, using the appropriate metric for every trigger in every detector, and we deem pairs of triggers to
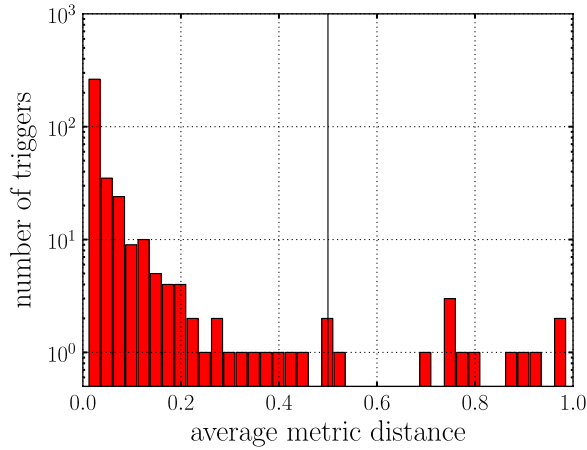
FIG. 4 (color online). Distribution of average parameter-space distance between coincident triggers associated with simulated GW signals in a month of representative S5 data, as recovered by the LIGO H1 and L1 detectors. The solid, vertical black line represents the threshold for triggers to be considered coincident; triggers to the right of this line were not considered to be coincident. The small number of simulated signals which failed coincidence were ones where the signal did not have sufficient SNR in at least one of the detectors to be a convincing detection. This cut does not cause signals to be missed that might otherwise be claimed as detections.

be coincident if their ellipsoids intersect. The radius of the ellipsoids is a tunable parameter. Computationally, the operation of finding all coincidences is vastly sped up by noticing that only triggers that are close in time could possibly have intersecting ellipsoids; therefore, the triggers are first sorted by time, and only those that share a small time window are compared.

When the detectors are not co-located, the coincidence test must also take into account the light travel time between detectors. This is done by minimizing the metric distance between the triggers over the possible range of time delays due to light travel time between the two detectors. In practice, this is achieved by iteratively adding a small value $\delta t_c$ to the end time of *one* of the detectors and calculating the metric distance. The lowest value of the metric distance is then used to determine if the triggers are coincident or not.

In Fig. 4 we show the distribution of metric distances (the minimum value for which the ellipsoids centred on the triggers overlap) for coincident triggers associated with simulated GW signals (see Sec. IV C). The number of coincidences falls off rapidly with increasing metric distances, whereas it would remain approximately constant for *background* coincident triggers generated by noise. However, since accuracy of parameter recovery diminishes at lower SNR, it is the quieter triggers from farther GW sources (which are statistically more likely) that are recovered with the largest metric distances. Therefore, larger coincidence ellipsoids can improve the overall sensitivity of a search.
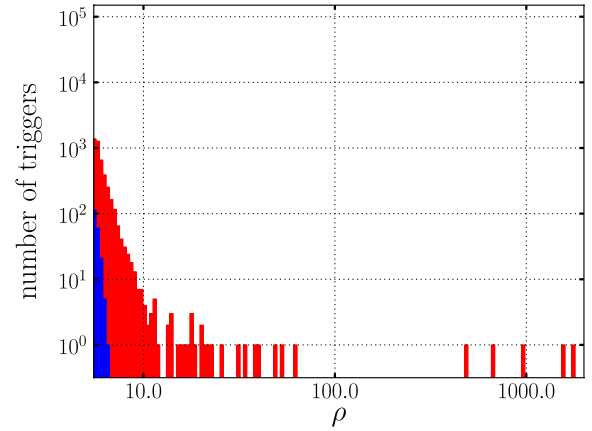


FIG. 5 (color online). Distribution of single detector SNRs for H1 coincident triggers in a month of simulated Gaussian noise in blue (dark gray) and representative S5 data in red (light gray). Coincidence was evaluated after time-shifting the SNR time series, so that only background coincidences caused by noise would be included. Comparison with Fig. 3 shows that the coincidence requirement reduces the high-SNR tail, but by no means eliminates it.

The result of the coincidence process is a list of all triggers that have SNR above threshold in two or more detectors and consistent parameters (masses and coalescence times) across detectors. When more than two detectors are operational, different combinations and higher-multiplicity coincidences are possible (e.g., three detectors yield triple coincidences and three types of double coincidences), and all coincidence types are stored.

In Fig. 5 we show the distribution of coincident H1 triggers as a function of SNR in a month of simulated Gaussian noise in blue (dark gray) and real S5 LIGO data in red (light gray). The largest single-detector SNRs for Gaussian noise are ∼7–8, comparable with (although somewhat larger than) early theoretical expectations [50,51]. However, the distribution in real data is significantly worse, with SNRs of hundreds and even thousands. If we were to end our analysis here, a GW search in real data would be a hundred times less sensitive (in distance) than a search in Gaussian, stationary noise with the same PSD.

## III. MITIGATING THE EFFECTS OF NON-GAUSSIAN NOISE

To further reduce the tail of high-SNR triggers caused by the non-Gaussianity and nonstationarity of noise, the IHOPE pipeline includes a number of *signal-consistency* tests, which compare the properties of the data around the time of a trigger with those expected for a real GW signal. After removing duplicates, the coincident triggers in each 2048 s block are used to create a *triggered template bank*. Any template in a given detector that forms at least one coincident trigger in each 2048 s block will enter the triggered

template bank for that detector and chunk. The new bank is again used to filter the data as described in Sec. II C, but this time signal-consistency tests are also performed. These include the $\chi^2$ (Sec. III A) and $r^2$ (Sec. III B) tests. Coincident triggers are selected as described in Sec. II D, and they are also tested for the consistency of relative signal amplitudes (Sec. III C); at this stage, *data-quality vetoes* are applied (Sec. III D) to sort triggers into categories according to the quality of data at their times.

The computational cost of the entire pipeline is reduced greatly by applying the expensive signal-consistency checks only in this second stage; the triggered template bank is, on average, a factor of $\sim 10$ smaller than the original template bank in the analysis described in Ref. [12]. However, the drawback is greater complexity of the analysis, and the fact that the coincident triggers found at the end of the two stages may not be identical.

## A. The $\chi^2$ signal-consistency test

The basis of the $\chi^2$ test [33] is the consideration that although a detector glitch may generate triggers with the same SNR as a GW signal, the manner in which the SNR is accumulated over time and frequency is likely to be different. For example, a glitch that resembles a delta function corresponds to a burst of signal power concentrated in a small time-domain window, but smeared out across all frequencies. A CBC waveform, on the other hand, will accumulate SNR across the duration of the template, consistently with the *chirp*-like morphology of the waveform.

To test whether this is the case, the template is broken into $p$ orthogonal subtemplates with support in adjacent frequency intervals, in such a way that each subtemplate would generate the same SNR on average over Gaussian noise realizations. The actual SNR achieved by each subtemplate filtered against the data is compared to its expected value, and the squared residuals are summed. Thus, the $\chi^2$ test requires $p$ inverse Fourier transforms per template. For the low-mass CBC search, we found that setting $p = 16$ provides a powerful discriminator without incurring an excessive computational cost [52].

For a GW signal that matches the template waveform exactly, the sum of squared residuals follows the $\chi^2$ distribution with $2p - 2$ degrees of freedom. For a glitch, or a signal that does not match the template, the expected value of the $\chi^2$-test is increased by a factor proportional to the total SNR$^2$, with a proportionality constant that depends on the mismatch between the signal and the template. For signals, we may write the expected $\chi^2$ value as

$$\langle \chi^2 \rangle = (2p - 2) + \epsilon^2 \rho^2, \tag{12}$$

where $\epsilon$ is a measure of signal-template mismatch. Even if CBC signals do not match template waveforms perfectly, due to template-bank discreteness, theoretical waveform inaccuracies [53], spin effects [27], calibration uncertainties [54], and so on, they will still yield significantly
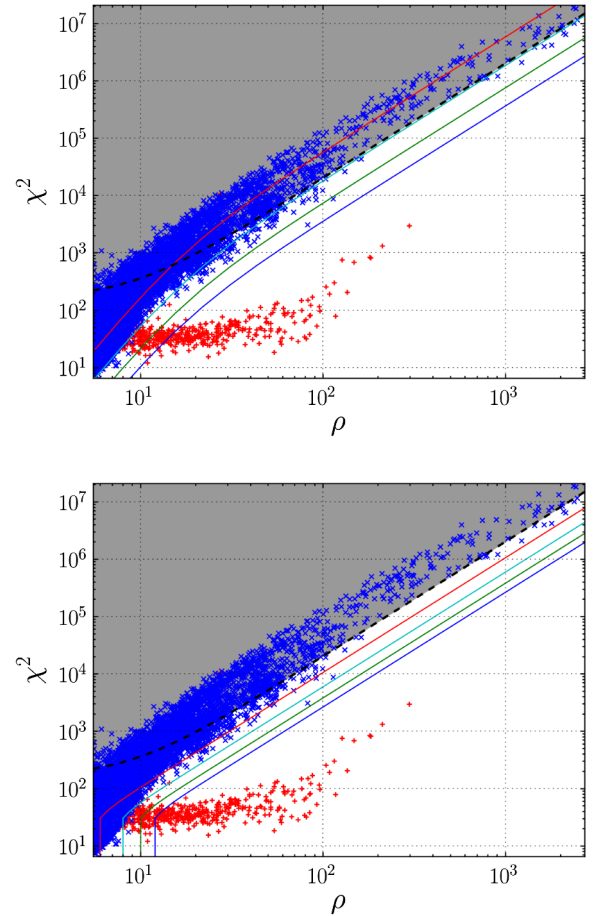


FIG. 6 (color online). The $\chi^2$ test plotted against SNR for triggers in a month of representative S5 data after the $\chi^2$ test has been applied, and the $r^2$ cut has been applied for triggers with $\rho < 12$. The blue crosses mark time shifted background triggers, the red pluses mark simulated-GW triggers. The grey shaded region bounded by a black dashed line denotes the area in which triggers are vetoed by the chi-squared cut (Sec. III A). The solid, colored lines on the plots indicate lines of constant effective SNR (top panel) and new SNR (bottom panel), which are described in Sec. III E. Larger values of effective/new SNR are at the bottom and right end of the plots.

smaller $\chi^2$ than most glitches. It was found empirically that a good fraction of glitches are removed (with minimal effect on simulated signals) by imposing a SNR-dependent $\chi^2$ threshold of the form

$$\chi^2 \leq \xi^2(p + \delta\rho^2), \tag{13}$$

with $\xi^2 = 10$ and $\delta = 0.2$.

In Fig. 6 we show the distribution of $\chi^2$ as a function of SNR. Even following the cut, a clear separation between noise background and simulated signals can easily be observed. This will be used later in formulating a detection statistic that combines the values of both $\rho$ and $\chi^2$. We note that this chi-squared cut is of little scientific benefit, as triggers that are removed by this cut would not be
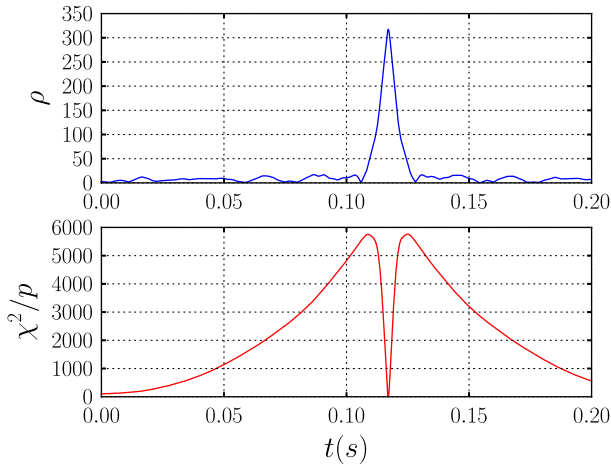
FIG. 7 (color online). Value of SNR and $\chi^2$ as a function of time, for a simulated CBC signal with SNR $= 300$ in a stretch of S5 data from the H1 detector. The SNR shows a characteristic rise and fall around the signal. The $\chi^2$ value is small at the time of the signal, but increases steeply to either side as the template waveform is offset from the signal in the data.



FIG. 8 (color online). The time $\Delta T$ that the $\chi^2$ is above the threshold $pr^2$ as a function of SNR, for all second-stage H1 triggers in a month of representative S5 data. The blue (dark gray) crosses mark all background triggers that fail the cut; blue (dark gray) circles indicate background triggers that pass it. Red (light gray) symbols mark simulated-GW triggers, with only a single signal removed by the cut.

considered detection candidates by the statistic that we derive in Sec. III E. However, this cut does significantly reduce the number of triggers that are being handled, without which computational memory restrictions become a concern.

### B. The $r^2$ signal-consistency test

We can also test the consistency of the data with a postulated signal by examining the time series of SNRs and $\chi^2$s. For a true GW signal, this would show a single sharp peak at the time of the signal, with the width of the falloff determined by the autocorrelation function of the template [55,56]. Thus, counting the number of time samples around a trigger for which the SNR is above a set threshold provides a useful consistency test [57]. Examining the behavior of the $\chi^2$ time series provides a more powerful diagnostic [58]. To wit, the $r^2$ test sets an upper threshold on the amount of time $\Delta T$ (in a window $T$ prior to the trigger[5]) for which

$$\chi^2 \geq pr^2, \qquad (14)$$

where $p$ is the number of subtemplates used to compute the $\chi^2$. We found empirically that setting $T = 6$ s and $r^2 = 15$ produces a powerful test [58]. Figure 7 shows the characteristic shape of the $\chi^2$ time series for CBC signals: close to zero when the template is aligned with the signal, then increasing as the two are offset in time, before falling off again with larger time offsets.

An effective $\Delta T$ threshold must be a function of SNR; the $\Delta T$ commonly used for IHOPE searches is

$$\Delta T < \begin{cases} 2 \times 10^{-4} \text{ s} & \text{for } \rho < 12, \\ \rho^{9/8} \times 7.5 \times 10^{-3} \text{ s} & \text{for } \rho \geq 12. \end{cases} \qquad (15)$$

The threshold for $\rho < 12$ eliminates triggers for which *any* sample is above the threshold from Eq. (14).

In Fig. 8 we show the effect of such an SNR test. For $\rho < 12$, the value of $\Delta T$ is smaller than the sample rate, therefore, triggers are discarded if there are any time samples in the 6 s prior to the trigger for which Eq. (14) is satisfied. [Since the 6 s window includes the trigger, for some SNRs this imposes a more stringent requirement than the $\chi^2$ test (13).] For $\rho \geq 12$, the threshold is SNR dependent. The $r^2$ test is powerful at removing a large number of high-SNR background triggers [the blue (dark gray) crosses], without affecting the triggers produced by simulated GW signals [the red (light gray) circles]. The cut is chosen to be conservative to allow for any imperfect matching between CBC signals and template waveforms.

### C. Amplitude-consistency tests

The two LIGO Hanford detectors H1 and H2 share the same vacuum tubes, and therefore expose the same sensitive axes to any incoming GW. Thus, the ratio of the H1 and H2 SNRs for true GW signals should equal the ratio of detector sensitivities. We can formulate a formal test of H1–H2 *amplitude consistency*[6] in terms of a GW source's

---

[5]The nonsymmetric window was chosen because the merger-ringdown phase of CBC signals, which is not modeled in inspiral-only searches, may cause an elevation in the $\chi^2$ time series after the trigger.
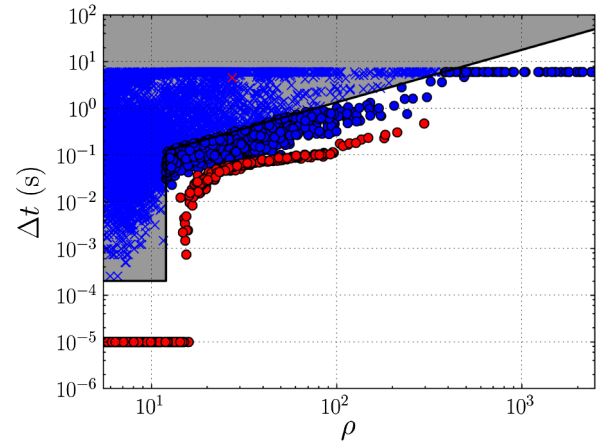
[6]The detector H2 was not operational during LIGO run S6, so the H1–H2 amplitude-consistency tests were not applied; they were, however, used in searches over data from previous runs.
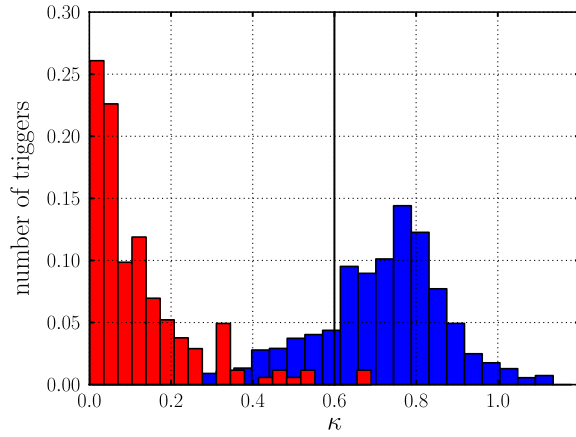
FIG. 9 (color online). Distribution of $\kappa$ [Eq. (16)], the fractional difference in the effective distances measured by H1 and H2 for coincident triggers in those detectors in a month of representative S5 data. Background triggers plotted in blue (dark gray) tend to have larger than simulated-GW triggers shown in red (light gray).

*effective distance $D_{\mathrm{eff},A}$*—the distance at which an optimally located and oriented source would give the SNR observed with detector $A$. Namely, we require that

$$\kappa = 2\frac{|D_{\mathrm{eff,H1}} - D_{\mathrm{eff,H2}}|}{D_{\mathrm{eff,H1}} + D_{\mathrm{eff,H2}}} \leq \kappa^*; \qquad (16)$$

setting a threshold $\kappa^*$ provides discrimination against noise triggers while allowing for some measurement uncertainty. In Fig. 9 we show the distribution of $\kappa$ for simulated-GW triggers and background triggers in a month of representative S5 data. We found empirically that setting $\kappa^* = 0.6$ produces a powerful test.

An amplitude-consistency test can be defined also for triggers that are seen in only one of H1 and H2. We do this by removing any triggers from H1 which are loud enough that we would have expected to observe a trigger in H2 (and vice-versa). We proceed by calculating $\sigma_A$, the distance at which an optimally located and oriented source yields an SNR of 1 in detector $A$, and noting that $D_{\mathrm{eff},A} = \sigma_A/\rho_A$. Then, by rearranging (16), we are led to require that a trigger that is seen only in H1 satisfy

$$\rho_{\mathrm{H1}} < \frac{\sigma_{\mathrm{H1}}}{\sigma_{\mathrm{H2}}}\left(\frac{2 + \kappa^*}{2 - \kappa^*}\right)\rho^*_{\mathrm{H2}}, \qquad (17)$$

where $\rho^*_{\mathrm{H2}}$ is the SNR threshold used for H2. The effective distance cut removes essentially all H2 triggers for which there is no H1 coincidence: since H2 typically had around half the distance sensitivity of H1, a value of $\kappa^* = 0.6$ imposes $\rho_{\mathrm{H2}} < \rho^*_{\mathrm{H1}}$.

Neither test was used between any other pair of detectors because, in principle, any ratio of effective distances is possible for a real signal seen in two nonaligned detectors. However, large values of $\kappa$ are rather unlikely, especially

for the Hanford and Livingston LIGO detectors, which are *almost* aligned. Therefore, amplitude-consistency tests should still be applicable.

### D. Data-quality vetoes

Environmental factors can cause periods of elevated detector glitch rate. In the very worst (but very rare) cases, this makes the data essentially unusable. More commonly, if these glitchy periods were analyzed together with periods of relatively clean data, they could produce a large number of high-SNR triggers, and possibly mask GW candidates in clean data. It is therefore necessary to remove or separate the glitchy periods.

This is accomplished using *data quality* (DQ) flags [59–61]. All detectors are equipped with environmental and instrumental monitors; their output is recorded in the detector's *auxiliary* channels. Periods of heightened activity in these channels (e.g., as caused by elevated seismic noise [62]) are automatically marked with DQ flags [63]. DQ flags can also be added manually if the detector operators observe poor instrumental behavior.

If a DQ flag is found to be strongly correlated with CBC triggers, and if the flag is *safe* (i.e., not triggered by real GWs), then it can be used a DQ *veto*. Veto safety is assessed by comparing the fraction of hardware GW injections that are vetoed with the total fraction of data that is vetoed. During the S6 and VSR2–3 runs, a simplified form of IHOPE was run daily on the preceding 24 hours of data from each detector individually, specifically looking for non-Gaussian features that could be correlated with instrumental or environmental effects [62,64]. The results of these daily runs were used to help identify common glitch mechanisms and to mitigate the effects of non-Gaussian noise by suggesting data quality vetoes.

Vetoes are assigned to categories based on the severity of instrumental problems and on how well the couplings between the GW and auxiliary channels are understood [59–61]. Correspondingly, CBC searches assign data to four DQ categories:

Category 1. Seriously compromised or missing data. The data are entirely unusable, to the extent that they would corrupt noise PSD estimates. These times are excluded from the analysis, as if the detector was not in science mode (introduced in Sec. II A).

Category 2. Instrumental problems with known couplings to the GW channel. Although the data are compromised, these times can still be used for PSD estimation. Data flagged as category-2 are analyzed in the pipeline, but any triggers occurring during these times are discarded. This reduces the fragmentation of science segments, maximizing the amount of data that can be analyzed.

Category 3.    Likely instrumental problems, casting doubt on triggers found during these times. Data flagged as category-3 are analyzed and triggers are processed. However, the excess noise in such times may obscure signals in clean data. Consequently, the analysis is also performed excluding time flagged as category-3, allowing weaker signals in clean data to be extracted. These data are excluded from the estimation of upper limits on GW-event rates.

Good data.    Data without any active environmental or instrumental source of noise transients. These data are analyzed in full.

Poor quality data are effectively removed from the analysis, reducing the total amount of analyzed time. For instance, in the third month of the S5 analysis reported in Ref. [12], removing category-1 times left $1.2 \times 10^6$ s of data when at least two detectors were operational; removing category-2 and -3 times left $1.0 \times 10^6$ s, although the majority of lost time was category-3, and was therefore analyzed for loud signals.

### E. Ranking statistics

The application of signal-consistency and amplitude-consistency tests, as well as data-quality vetoes, is very effective in reducing the non-Gaussian tail of high-SNR triggers. In Fig. 10 we show the distribution of H1 triggers that are coincident with triggers in the L1 detector (in time shifts) and that pass all cuts. For consistency, identical cuts have been applied to the simulated, Gaussian data, including vetoing times of poor data quality in the real data.
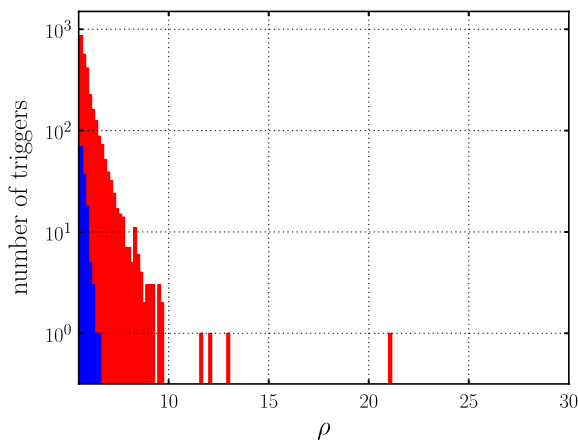


FIG. 10 (color online).   Distribution of single detector SNRs for H1 triggers found in coincidence with L1 triggers (in time shifts) in a month of simulated Gaussian noise in blue (dark gray) and representative S5 data in red (light gray). These triggers have survived $\chi^2$, $r^2$, and H1–H2 amplitude-consistency tests, as well as DQ vetoes.

The majority of these cuts have minimal impact, although the data quality vetoes will remove a (random) fraction of the triggers arising in the simulated data analysis.

Remarkably, in the real data, almost no triggers are left that have SNR $> 10$. Nevertheless, a small number of coincident noise triggers with large SNR remain. These triggers have passed all cuts, but they generally have significantly worse $\chi^2$ values than expected for true signals, as we showed in Fig. 6.

It is therefore useful to rank triggers using a *combination* of SNR and $\chi^2$, by introducing a *reweighted SNR*. Over the course of the LIGO-Virgo analyses, several distinct reweighted SNRs have been used. For the LIGO S5 run and Virgo's first science run (VSR1), we adopted the *effective SNR* $\rho_{\mathrm{eff}}$, defined as [11]

$$\rho_{\mathrm{eff}}^2 = \frac{\rho^2}{\sqrt{\left(\frac{\chi^2}{n_{\mathrm{dof}}}\right)\left(1 + \frac{\rho^2}{250}\right)}}, \qquad (18)$$

where $n_{\mathrm{dof}} = 2p - 2$ is the number of $\chi^2$ degrees of freedom, and the factor 250 was tuned empirically to provide separation between background triggers and simulated GW signals. The normalization of $\rho_{\mathrm{eff}}$ ensures that a "quiet" signal with $\rho \simeq 8$ and $\chi^2 \simeq n_{\mathrm{dof}}$ will have $\rho_{\mathrm{eff}} \simeq \rho$.

Figure 6 shows contours of constant $\rho_{\mathrm{eff}}$ in the $\rho$-$\chi^2$ plane. While $\rho_{\mathrm{eff}}$ successfully separates background triggers from simulated-GW triggers, it can artificially elevate the SNR of triggers with unusually small $\chi^2$. As discussed in Ref. [65], these can sometimes become the most significant triggers in a search. Thus, a different statistic was adopted for the LIGO S6 run and (VSR23). This *new SNR* $\rho_{\mathrm{new}}$ [14] was defined as

$$\rho_{\mathrm{new}} = \begin{cases} \rho & \text{for } \chi^2 \le n_{\mathrm{dof}}, \\ \rho\left[\frac{1}{2}\left(1 + \left(\frac{\chi^2}{n_{\mathrm{dof}}}\right)^3\right)\right]^{-1/6} & \text{for } \chi^2 > n_{\mathrm{dof}}. \end{cases} \qquad (19)$$

Figure 6 also shows contours of constant $\rho_{\mathrm{new}}$ in the $\rho$-$\chi^2$ plane. The new SNR was found to provide even better background-signal separation, especially for low-mass nonspinning inspirals [14], and it has the desirable feature that $\rho_{\mathrm{new}}$ does not take larger values than $\rho$ when the $\chi^2$ is less than the expected value. Other ways of defining a detection statistic as a function of $\rho$ and $\chi^2$ can be defined and optimized for analyses covering different regions of parameter space and different data sets.

For coincident triggers, the reweighted SNRs measured in the coincident detectors are added in quadrature to give a *combined*, reweighted SNR, which is used to rank the triggers and evaluate their statistical significance. Using this ranking statistic, we find that the distribution of background triggers in real data is remarkably close to their distribution in simulated Gaussian noise, as shown in Figs. 11 and 12. Thus, our consistency tests and DQ vetoes have successfully eliminated the vast majority of high SNR
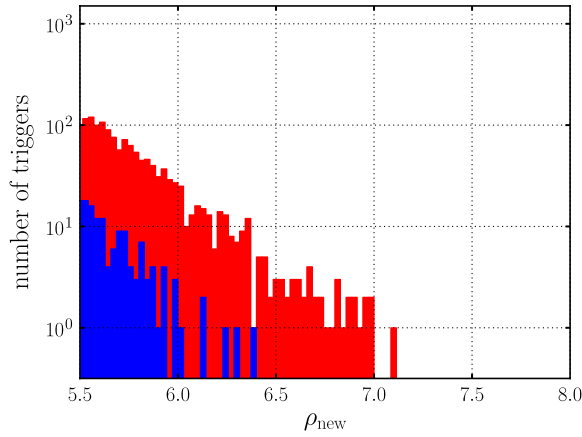
FIG. 11 (color online). Distribution of single detector new SNR, $\rho_{\text{new}}$, for H1 triggers found in coincidence with L1 triggers (in time shifts) in a month of simulated Gaussian noise in blue (dark gray) and representative S5 data in red (light gray). The tail of high (SNR) triggers due to non-Gaussian noise has been virtually eliminated—a remarkable achievement given that the first stage of the pipeline generated single-detector triggers with SNR $> 1,000$.
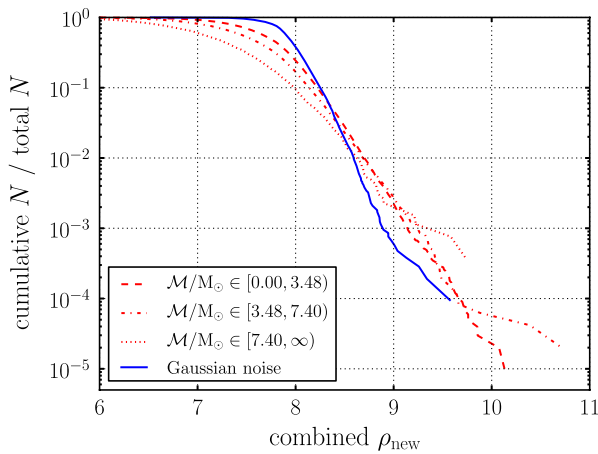


FIG. 12 (color online). Fraction of time-shift coincident triggers between H1 and L1 in a month of representative S5 data that have combined new SNR greater than or equal to the *x*-axis value, for three chirp-mass bins. The distribution from a month of Gaussian noise is also shown for comparison. The tails of the distributions become more shallow for larger chirp masses $\mathcal{M}$, so triggers with higher $\mathcal{M}$ are more likely to have higher SNRs.

triggers due to non-Gaussian noise from the search. While this comes at the inevitable cost of missing potential detections at times of poor data quality, it significantly improves the detection capability of a search.

## IV. INTERPRETATION OF THE RESULTS

At the end of the data processing described above, the IHOPE pipeline produces a set of coincident triggers ranked by their combined reweighted SNR; these triggers have

passed the various signal-consistency and data-quality tests outlined above. While at this stage the majority of loud background triggers identified in real data have been eliminated or downweighted, the distribution of triggers is still different from the case of Gaussian noise, and it depends on the quality of the detector data and the signal parameter space being searched over. Therefore, it is not possible to derive an analytical mapping from combined reweighted SNR to event significance, as characterized by the FAR. Instead, the FAR is evaluated empirically by performing numerous *time-shift* analyses, in which artificial time shifts are introduced between the data from different detectors. (These are discussed in Sec. IV A.) Furthermore, the rate of triggers as a function of combined reweighted SNR varies over parameter space; to improve the FAR accuracy, we divide triggers into groups with similar combined reweighted SNR distributions (see Sec. IV B). The sensitivity of a search is evaluated by measuring the rate of recovery of a large number of simulated signals, with parameters drawn from astrophysically motivated distributions (see Sec. IV C). The sensitivity is then used to estimate the CBC event rates or upper limits as a function of signal parameters (see Sec. IV D).

### A. Background event rate from time shifts

The rate of coincident triggers as a function of combined reweighted SNR is estimated by performing numerous time-shift analyses: in each we artificially introduce different relative time shifts in the data from each detector [66]. The time shifts that are introduced must be large enough such that each time-shift analysis is statistically independent.

To perform the time-shift analysis in practice, we simply shift the triggers generated at the first matched-filtering stage of the analysis (Sec. II C), and repeat all subsequent stages from multidetector coincidence (Sec. II D) onwards. Shifts are performed on a ring: for each time-coincidence period (i.e., data segment where a certain set of detectors is operational), triggers that are shifted past the end are re-inserted at the beginning. Since the time-coincidence periods are determined *before* applying Category-2 and -3 DQ flags, there is some variation in analyzed time among time-shift analyses. To ensure statistical independence, time shifts are performed in multiples of 5 s; this ensures that they are significantly larger than the light travel time between the detectors, the autocorrelation time of the templates, and the duration of most non-transient glitches seen in the data. Therefore, any coincidences seen in the time shifts cannot be due to a single GW source, and are most likely due to noise-background triggers. It is possible, however, for a GW-induced trigger in one detector to arise in time-shift coincidence with noise in another detector. Indeed, this issue arose in Ref. [14], where a "blind injection" was added to the data to test the analysis procedure.

The H1 and H2 detectors share the Hanford beam tubes and are affected by the same environmental disturbances; furthermore, noise transients in the two detectors have been observed to be correlated. Thus, time-shift analysis is ineffective at estimating the coincident background between these co-located detectors, and it is not used. Coincident triggers from H1 and H2 when no other detectors are operational are excluded from the analysis. When detectors at additional sites are operational, we do perform time shifts, keeping H1 and H2 "in time" but shifting both relative to the other detectors.

Our normal practice is to begin by performing 100 time-shift analyses to provide an estimate of the noise background. If any coincident in-time triggers are still more significant (i.e., have larger combined reweighted SNR) than all the time-shifted triggers, additional time shifts are performed to provide an estimate of the FAR. A very significant candidate would have a very low FAR, and an accurate determination of its FAR requires a large number of time slides: in Ref. [14] over a million were performed. However, there is a limit to the number of statistically independent time shifts that are possible to perform, as explored in Ref. [67]. Additionally, as the number of time shifts grows, the computational savings of our two-stage search are diminished, because a greater fraction of the templates survive to the second filtering stage where the computationally costly signal-consistency tests are performed (see Sec. III A). We are currently investigating whether it is computationally feasible to run IHOPE as a single-stage pipeline and compute $\chi^2$ and $r^2$ for every trigger.

## B. Calculation of false-alarm rates

The FAR for a coincident trigger is given by the rate at which background triggers with the same or greater SNR occur due to detector noise. This rate is computed from the time-shift analyses; for a fixed combined reweighted SNR, it varies across the template mass space, and it depends on which detectors were operational and how glitchy they were. To accurately account for this, coincident triggers are split into *categories*, and FARs are calculated within each, relative to a background of comparable triggers. The triggers from each category are then recombined into a single list and ranked by their FARs.

Typically, signal-consistency tests are more powerful for longer-duration templates than for shorter ones, so the non-Gaussian background is suppressed better for low-mass templates, while high-mass templates are more likely to result in triggers with larger combined reweighted SNRs. In recent searches, triggers have been separated into three bins in *chirp mass* $\mathcal{M} = M\eta^{3/5}$, as the chirp mass determines the duration of the signal [as seen from Eq. (9)]. We use three chirp mass bins: $\mathcal{M} \leq 3.48 M_\odot$, $3.48 M_\odot < \mathcal{M} \leq 7.4 M_\odot$, and $\mathcal{M} > 7.4 M_\odot$. Figure 12 shows the cumulative distribution of coincident triggers between H1

and L1 as a function of combined $\rho_{\text{new}}$ for the triggers in each of these mass bins. As expected, the high-$\mathcal{M}$ bin has a greater fraction of high-SNR triggers.

The combined reweighted SNR is calculated as the quadrature sum of the SNRs in the individual detectors. However, different detectors can have different rates of non-stationary transients as well as different sensitivities, so the combined SNR is not necessarily the best measure of the significance of a trigger. Additionally, background triggers found in three-detector coincidence will have a different distribution of combined reweighted SNRs than two-detector coincident triggers [11]. Therefore, we separate coincident triggers by their *type*, which is determined by the coincidence itself (e.g., H1H2, or H1H2L1) and by the availability of data from each detector, known as "coincident time." Thus, the trigger types would include H1L1 coincidences in H1L1 double-coincident time; H1L1, H1V1, L1V1, and H1L1V1 coincidences in H1L1V1 triple-coincident time; and so on. When H1 and H2 are both operational, we have fewer coincidence types than might be expected as H1H2 triggers are excluded due to our inability to estimate their background distribution, and the effective distance cut removes H2L1 or H2V1 coincidences. The product of mass bins and trigger types yields all the trigger categories.

For simplicity, we treat times when different networks of detectors were operational as entirely separate experiments; this is straightforward to do, as there is no overlap in time between them. Furthermore, the data from a long science run is typically broken down into a number of distinct stretches, often based upon varying detector sensitivity or glitchiness, and each is handled independently.

For each category of coincident triggers within an experiment, an additional clustering stage is applied to ensure that only a single trigger is associated with each "event" in the data (either signal or noise transient). If there is another coincident trigger with a larger combined reweighted SNR within 10 s of a given trigger's end time, the trigger is removed. The same clustering is applied to in-time and time-shift coincidences. We then compute the FAR as a function of combined reweighted SNR as the rate (number over the total coincident, time-shifted search time) of time-shift coincidences observed with higher combined reweighted SNR within each category. These results must then be combined to estimate the overall significance of triggers: we calculate a *combined FAR* across categories by ranking all triggers by their FAR, counting the number of more significant time-shift triggers, and dividing by the total time-shift time. The resulting combined FAR is essentially the same as the uncombined FAR, multiplied by the number of categories that were combined. We often quote the inverse FAR (IFAR) as the ranking statistic, so that more significant triggers correspond to larger values. A loud GW may produce triggers in more than one mass bin, and consequently more than one candidate trigger might be

due to a single event. This is resolved by reporting only the coincident trigger with the largest IFAR associated with a given event. Figure 13 shows the expected mean (the dashed line) and variation (the shaded areas) of the cumulative number of triggers as a function of IFAR for the analysis of three-detector H1H2L1 time in a representative month of S5 data. The variations among time shifts (the thin lines) match the expected distribution. The duration of the time-shift analysis is $\sim 10^8$ s, but taking into account the six categories of triggers (three mass bins and two coincidence types), this yields a minimum FAR of $\sim 1$ yr$^{-1}$.

Clearly a FAR of $\sim 1$ yr$^{-1}$ is insufficient to confidently identify GW events. The challenge of extending background estimation to the level where a loud trigger can become a detection candidate was met in the S6–VSR2/3 search [14,68]. Remarkably, even for FARs of one in tens of thousands of years, no tail of triggers with large combined reweighted SNRs was observed. Evidently, the cuts, tests, and thresholds discussed in Sec. III are effective at eliminating any evidence of a non-Gaussian background, at least for low chirp masses.

In calculating the FAR, we treat all trigger categories identically, so we implicitly assign the same weight to each. However, this is not appropriate when the detectors have significantly different sensitivities, since a GW is more likely to be observed in the most sensitive detectors. In the search of LIGO S5 and Virgo VSR1 data [13], this approach was refined by weighting the categories on the basis of the search sensitivity for each trigger type.
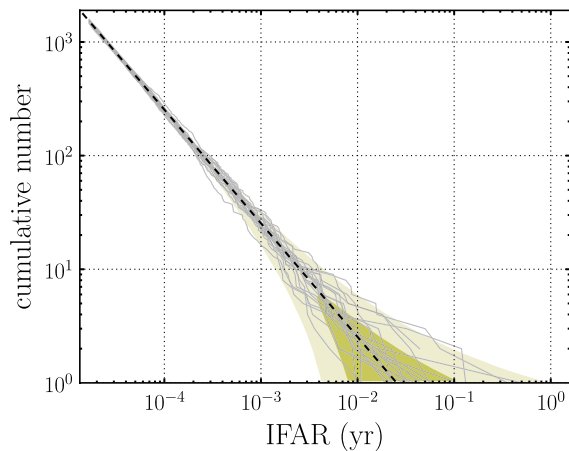


FIG. 13 (color online). Cumulative histogram of triggers above and IFAR threshold for all time-shift triggers in H1H2L1 triple-coincident time from a representative month of S5 data. The black dashed line marks the expected cumulative number of background triggers in the in-time data, as estimated from the time-shift analyses. The shaded regions mark its 1- and 2-$\sigma$ variation and the thin grey lines show the cumulative number for 20 of the time shifts, providing an additional indication of the expected deviation from the mean.

Additionally, if there were an accurate astrophysical model of CBC merger rates for different binary masses, the weighting could easily be extended to the mass bins.

## C. Evaluating search sensitivity

The sensitivity of a search is measured by adding simulated GW signals to the data and verifying their recovery by the pipeline, which also helps tune the pipeline's performance against expected sources. The simulated signals can be added as *hardware injections* [14,69], by actuating the end mirrors of the interferometers to reproduce the response of the interferometer to GWs, or as *software injections*, by modifying the data after it has been read into the pipeline. Hardware injections provide a better end-to-end test of the analysis, but only a limited number can be performed, since the data containing hardware injections cannot be used to search for real GW signals. Consequently, large-scale injection campaigns are performed in software.

Software injections are performed into all operational detectors *coherently* (i.e., with relative time delays, phases and amplitudes appropriate for the relative location and orientation of the source and the detectors). Simulated GW sources are generally placed uniformly over the celestial sphere, with uniformly distributed orientations. The mass and spin parameters are generally chosen to uniformly cover the search parameter space, since they are not well constrained by astrophysical observations, particularly so for binaries containing black holes [70]. Although sources are expected to be roughly uniform in volume, we do not follow that distribution for simulations, but instead attempt to place a greater fraction of injections at distances where they would be marginally detectable by the pipeline. The techniques used to reduce the dimensionality of parameter space, such as analytically maximizing the detection statistic, cannot be applied to the injections, which must cover the entire space. This necessitates large simulation campaigns.

The IHOPE pipeline is run on the data containing simulated signals using the same configuration as for the rest of the search. Injected signals are considered to be found if there is a coincident trigger within 1 s of their injection time. The loudest coincident trigger within the 1 s window is associated with the injection, and it may be louder than any trigger in the time-shift analyses. Using a 1 s time window to associate triggers and injections and no requirement on mass consistency may lead to some of these being found spuriously, in coincidence with background triggers. However, this effect has negligible consequences on the estimated search sensitivity near the combined reweighted SNR of the most significant trigger.

Figure 14 shows the results of a large number of software injections performed in one month of S5 data. For each injection, we indicate whether the signal was missed (red crosses) or found (circles, and stars for triggers louder
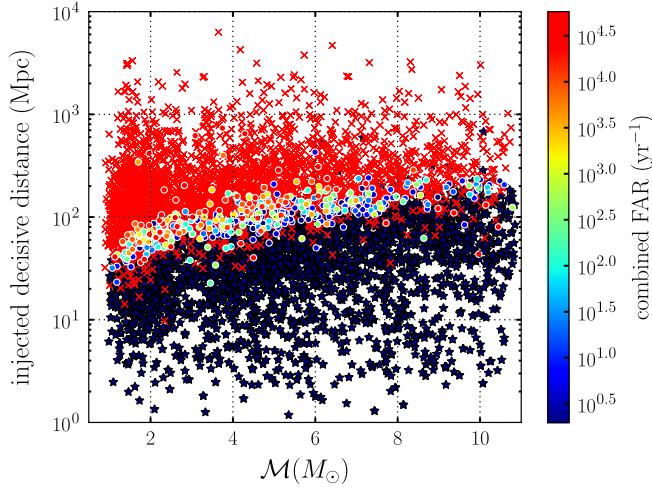
FIG. 14 (color online).  *Decisive distance* versus chirp mass, $\mathcal{M}$, (see main text for definition) of found and missed injections in one month of S5 data. Red crosses are missed injections; colored circles are injections found with nonzero combined FAR, which can be read off the color map on the right; dark-blue stars are injections found with triggers louder than any in the background from 100 time shifts. Nearby injections that are missed or found with high FARs are followed up to check for problems in the pipeline, and to improve data quality.

than any background). The recovery of simulated signals can be compared with the theoretically expected sensitivity of the search, taking into account variations over parameter space: the expected SNR of a signal is proportional to $\mathcal{M}^{5/6}$ (for low-mass binaries), inversely proportional to effective distance (see Sec. III C), and a function of the detectors' noise PSD. In a coincidence search, it is the second most sensitive detector to a given source that determines whether that source should be detected. To capture this notion, in Fig. 14, we show the distribution of found and missed simulations in chirp mass $\mathcal{M}$ and *decisive distance*—the second largest effective distance for the detectors that were operating at the time of the injection. One expects a relatively clean separation between found and missed signals, with the transition point corresponding to the sensitivity of the detectors to optimally oriented sources. Our empirical results are in good agreement with the stated sensitivity of the detectors [71,72]. A small number of signals are missed at low distances: these are virtually always found to lie close to loud non-Gaussian glitches in the detector data.

### D. Bounding the binary coalescence rate

The results of a search can be used to estimate (if positive detections are reported) or bound the rate of binary coalescences. An upper limit on the merger rate is calculated by evaluating the sensitivity of the search at the loudest observed trigger [34,73–75]. Heuristically, the 90% rate upper limit corresponds to a few (order 2–3)

signals occurring over the search time within a small enough distance to generate a trigger with IFAR larger than the loudest observed trigger.

More specifically, we assume that CBC events occur randomly and independently, and that the event rate is proportional to the star-formation rate, which is itself assumed proportional to blue-light galaxy luminosity [76]. For searches sensitive out to tens or hundreds of megaparsecs, it is reasonable to approximate the blue-light luminosity as uniform in volume, and quote rates per unit volume and time [18]. Since the search sensitivity varies with mass (as shown in Fig. 14), we typically divide the mass space into regions of roughly constant sensitivity and calculate an upper limit for each mass region. We follow Refs. [11,74] and infer the probability density for the merger rate $R$, in one mass region, given that in an observation time $T$ no other trigger was seen with IFAR larger than its loudest-event value, $\alpha_m$:

$$p(R|\alpha_m, T) \propto p(R)e^{-RV(\alpha_m)T}(1 + \Lambda(\alpha_m)RTV(\alpha_m));$$
(20)

here $p(R)$ is the prior probability density for $R$, usually taken as the result of previous searches or as a uniform distribution for the first search of a kind; $V(\alpha)$ is the volume of space in which the search could have seen a signal with IFAR $\geq \alpha$; and the quantity $\Lambda(\alpha)$ is the relative probability that the loudest trigger was due to a GWs rather than noise,

$$\Lambda(\alpha) = \frac{|V'(\alpha)|}{V(\alpha)} \frac{P_B(\alpha)}{P'_B(\alpha)},$$
(21)

where

$$P_B(\alpha) = e^{-T/\alpha}$$
(22)

denotes the probability of obtaining a background event with an IFAR greater than $\alpha$ and prime denotes differentiation with respect to $\alpha$. For a chosen confidence level $\gamma$ (typically $0.9 = 90\%$), the upper limit $R_*$ on the rate is then given by

$$\gamma = \int_0^{R_*} p(R|\alpha_m, T)dR.$$
(23)

It is clear from Eq. (20) that the decay of $p(R|\alpha_m, T)$ and the resulting $R_*$ depend critically on the *sensitive volume* $V(\alpha_m)$. In previous sections we have shown how IHOPE is highly effective at filtering out triggers due to non-Gaussian noise, thus improving sensitivity, and in the context of computing upper limits, we can quantify the residual effects of non-Gaussian features on $V(\alpha_m)$. In Fig. 15 we show the search *efficiency* for efficiency for binary neutron star (BNS) signals, i.e., the fraction of BNS injections found with IFAR above a fiducial value, here set to the IFAR of the loudest in-time noise trigger, as a
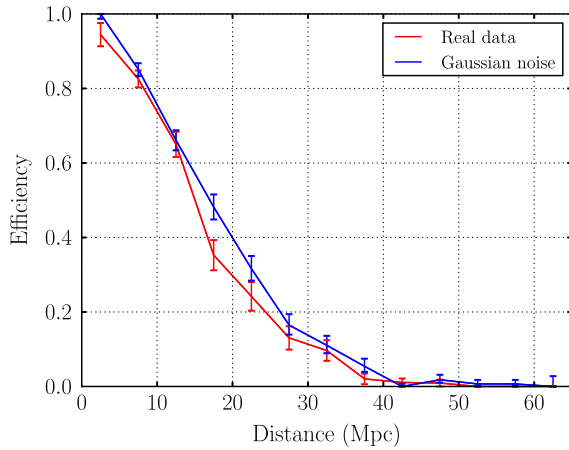
FIG. 15 (color online).    Search efficiency for BNS injections in a month of representative S5 data and in Gaussian noise, for a false-alarm rate equal to the FAR of the loudest foreground trigger in each analysis.

function of distance, for one month of S5 data and for a month of Gaussian noise with the same PSDs.[7] Despite the significant non-Gaussianity of real data, the distance at which efficiency is 50% is reduced by $\sim$10% compared to Gaussian-noise expectations. The sensitive search volume (the volume weighted integral of efficiency) is within $\sim$30% of what would be achieved in Gaussian noise.

## V. DISCUSSION AND FUTURE DEVELOPMENTS

In this paper we have given a detailed description of the IHOPE software pipeline, developed to search for GWs from CBC events in LIGO and Virgo data, and we have provided several examples of its performance on a sample stretch of data from the LIGO S5 run. The pipeline is based on a matched-filtering engine augmented by a substantial number of additional modules that implement coincidence, signal-consistency tests, data-quality cuts, tunable ranking statistics, background estimation by time shifts, and sensitivity evaluation by injections. Indeed, with the IHOPE pipeline we can run analyses that go all the way from detector strain data to event significance and upper limits on CBC rates.

The pipeline was developed over a number of years, from the early versions used in LIGO's S2 BNS [5] search to its mature incarnation used in the analysis of S6 and VSR3 data [14]. One of the major successes of the IHOPE pipeline was the mitigation of spurious triggers from non-Gaussian noise transients, to such an extent that the overall volume sensitivity is reduced by less than 30% compared to what would be possible if noise was Gaussian. Nevertheless, there are still significant improvements that

---

[7]For Gaussian noise, we do not actually run injections through the pipeline, but compute the expected (SNR), given the sensitivity of the detectors at that time, and compare with the largest SNR among Gaussian-noise in-time triggers.

can and must be made to CBC searches if we are to meet the challenges posed by analyzing the data of advanced detectors. In the following paragraphs, we briefly discuss some of these improvements and challenges.

*Coherent analysis.*—As discussed above, the IHOPE pipeline comes close to the sensitivity that would be achieved if noise was Gaussian, with the same PSD. Therefore, while some improvement could be obtained by implementing more sophisticated signal-consistency tests and data-quality cuts, it will not be significant. If three or more detectors are active, sensitivity *would* be improved in a *coherent* [35,56,77] (rather than coincident) analysis that filters the data from all operating detectors simultaneously, requiring consistency between the times of arrival, relative amplitudes and phases of GW signals, as observed in each data stream. Such a search is challenging to implement because the data from the detectors must be combined differently for each sky position, significantly increasing computational cost.

Coherent searches *have* already been run for unmodeled burst-like transients [78], and for CBC signals in coincidence with gamma-ray-burst observations [79], but a full all-sky, all-time pipeline like IHOPE would require significantly more computation. A promising compromise may be a hierarchical search consisting of a first coincidence stage followed by the coherent analysis of candidates, although the estimation of background trigger rates would prove challenging as time shifts in a coherent analysis cannot be performed using only the recorded single detector triggers but require the full SNR time series.

*Background estimation.*—Confident GW detection requires that we assign a very low false-alarm probability to a candidate trigger [14]. In the IHOPE pipeline, this would necessitate a large number of time shifts, thus negating the computational savings of splitting matched filtering between two stages, or a different method of background estimation [68,80]. Whichever the solution, it will need to be automated to identify signal candidates rapidly for possible astronomical follow up.

*Event-rate estimation.*—After the first detections, we will begin to quote event-rate estimates rather than upper limits. The loudest-event method can be used for this Ref. [74], provided that the data are broken up so that much less than one gravitational wave signal is expected in each analyzed stretch. There are, however, other approaches [81] that should be considered for implementation.

*Template length.*—The sensitive band of advanced detectors will extend to lower frequencies ($\sim$10 Hz) than their first-generation counterparts, greatly increasing the length and number of templates required in a matched-filtering search. Increasing computational resources may not be sufficient, so we are investigating alternative approaches to filtering [82–86] and possibly the use of graphical processing units (GPUs).

*Latency.*—The latency of CBC searches (i.e., the ''wall-clock'' time necessary for search results to become available) has decreased over the course of successive science runs, but further progress is needed to perform prompt follow-up observations of GW candidate with conventional (electromagnetic) telescopes [82,87]. The target should be posting candidate triggers within minutes to hours of data taking, which was in fact achieved in the S6–VSR3 analysis with the MBTA pipeline [82].

*Template accuracy.*—While the templates currently used in IHOPE are very accurate approximations to BNS signals, they could still be improved for the purpose of neutron star-black hole (NSBH) and binary black hole (BBH) searches [53]. It is straightforward to extend IHOPE to include the effects of spin on the progress of inspiral (i.e., its phasing), but it is harder to include the orbital precession caused by spins and the resulting waveform modulations. The first extension would already improve sensitivity to BBH signals [88,89], but precessional effects are expected to be more significant for NSBH systems [32,90].

*Parameter estimation.*—Last, while IHOPE effectively searches the entire template parameter space to identify candidate triggers, at the end of the pipeline the only information available about these are the estimated binary masses, arrival time, and effective distance. Dedicated follow-up analyses can provide much more detailed and reliable estimates of all parameters [91–94], but IHOPE itself could be modified to provide rough first-cut estimates.

## ACKNOWLEDGMENTS

[1] B. Abbott *et al.* (LIGO Scientific Collaboration), Rep. Prog. Phys. **72,** 076901 (2009).

[2] T. Accadia *et al.* (Virgo Collaboration), J. Instrum. **7,** P03012 (2012).

[3] H. Grote (LIGO Scientific Collaboration), Classical Quantum Gravity **25,** 114043 (2008).

[4] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **69,** 122001 (2004).

[5] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **72,** 082001 (2005).

[6] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **73,** 062001 (2006).

[7] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **72,** 082002 (2005).

[8] B. Abbott *et al.* (LIGO Scientific Collaboration, TAMA Collaboration), Phys. Rev. D **73,** 102002 (2006).

[9] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **78,** 042002 (2008).

[10] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **77,** 062002 (2008).

[11] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **79,** 122001 (2009).

[12] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **80,** 047101 (2009).

[13] J. Abadie *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. D **82,** 102001 (2010).

[14] J. Abadie *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. D **85,** 082002 (2012).

[15] D. A. Brown, Classical Quantum Gravity **22,** S1097 (2005).

[16] LIGO Scientific Collaboration, Report No. LIGO-T0900288-v3, 2010, https://dcc.ligo.org/cgi-bin/DocDB/ShowDocument?docid=2974.

[17] Virgo Collaboration, Advanced Virgo Sensitivity, 2010, https://wwwcascina.virgo.infn.it/advirgo/.

[18] J. Abadie *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Classical Quantum Gravity **27,** 173001 (2010).

[19] L. Blanchet, Living Rev. Relativity **5,** 3 (2002).

[20] J. M. Centrella, J. G. Baker, B. J. Kelly, and J. R. van Meter, Rev. Mod. Phys. **82,** 3069 (2010).

[21] M. Hannam, Classical Quantum Gravity **26,** 114001 (2009).

[22] U. Sperhake, E. Berti, and V. Cardoso, arXiv:1107.2819.

[23] E. Berti, V. Cardoso, J. A. Gonzalez, U. Sperhake, M. Hannam, S. Husa, and B. Brügmann, Phys. Rev. D **76,** 064034 (2007).

[24] L. A. Wainstein and V. D. Zubakov, *Extraction of Signals from Noise* (Prentice-Hall, Englewood Cliffs, NJ, 1962).

[25] T. Cokelaer and D. Pathak, Classical Quantum Gravity **26,** 045013 (2009).

[26] D. A. Brown and P. J. Zimmerman, Phys. Rev. D **81,** 024007 (2010).

[27] C. Van Den Broeck, D. Brown, T. Cokelaer, I. Harry, G. Jones, B. Sathyaprakash, H. Tagoshi, and H. Takahashi, Phys. Rev. D **80,** 024009 (2009).

[28] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, Phys. Rev. D **85,** 122006 (2012).

[29] T. A. Apostolatos, C. Cutler, G. Sussman, and K. Thorne, Phys. Rev. D **49,** 6274 (1994).

[30] T. A. Apostolatos, Phys. Rev. D **52**, 605 (1995).

[31] A. Buonanno, Y. Chen, and M. Vallisneri, Phys. Rev. D **67**, 104025 (2003); **74**, 029904(E) (2006).

[32] Y. Pan, A. Buonanno, Y. Chen, and M. Vallisneri, Phys. Rev. D **69**, 104017 (2004); **74**, 029905(E) (2006).

[33] B. Allen, Phys. Rev. D **71**, 062001 (2005).

[34] P. R. Brady and S. Fairhurst, Classical Quantum Gravity **25**, 105002 (2008).

[35] L. S. Finn and D. F. Chernoff, Phys. Rev. D **47**, 2198 (1993).

[36] S. Droz, D. J. Knapp, E. Poisson, and B. J. Owen, Phys. Rev. D **59**, 124016 (1999).

[37] D. A. Brown, Ph.D. thesis, University of Wisconsin-Milwaukee, 2004, arXiv:0705.1514.

[38] T. Cokelaer, Phys. Rev. D **76**, 102004 (2007).

[39] S. Babak, R. Balasubramanian, D. Churches, T. Cokelaer, and B. S. Sathyaprakash, Classical Quantum Gravity **23**, 5477 (2006).

[40] B. J. Owen and B. S. Sathyaprakash, Phys. Rev. D **60**, 022002 (1999).

[41] B. J. Owen, Phys. Rev. D **53**, 6749 (1996).

[42] R. Balasubramanian, B. S. Sathyaprakash, and S. V. Dhurandhar, Phys. Rev. D **53**, 3033 (1996).

[43] S. V. Dhurandhar and B. S. Sathyaprakash, Phys. Rev. D **49**, 1707 (1994).

[44] B. S. Sathyaprakash and S. V. Dhurandhar, Phys. Rev. D **44**, 3819 (1991).

[45] B. Sathyaprakash, Phys. Rev. D **50**, R7111 (1994).

[46] FFTW-Fastest Fourier Transform in the West, http://www.fftw.org/.

[47] A. S. Sengupta, J. A. Gupchup, and C. A. K. Robinson, in *A Clustering Algorithm for Inspiral Search Pipelines* (Cardiff University, Cardiff, Wales, 2006).

[48] C. Capano, Ph.D. thesis, Syracuse University, 2012.

[49] C. A. K. Robinson, B. S. Sathyaprakash, and A. S. Sengupta, Phys. Rev. D **78**, 062002 (2008).

[50] B. F. Schutz, NATO Adv. Study Inst. Ser. C, Math. Phys. Sci. **253**, 1 (1989).

[51] C. Cutler *et al.*, Phys. Rev. Lett. **70**, 2984 (1993).

[52] S. Babak, H. Grote, M. Hewitson, H. Lück, and K. Strain, Phys. Rev. D **72**, 022002 (2005).

[53] A. Buonanno, B. R. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, Phys. Rev. D **80**, 084043 (2009).

[54] J. Abadie *et al.* (LIGO Scientific Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **624**, 223 (2010).

[55] C. Hanna, Ph.D. thesis, Louisiana State University, 2008.

[56] I. W. Harry and S. Fairhurst, Phys. Rev. D **83**, 084002 (2011).

[57] P. Shawhan and E. Ochsner, Classical Quantum Gravity **21**, S1757 (2004).

[58] A. Rodríguez, M.S. thesis, Louisiana State University, 2007, arXiv:0802.1376.

[59] J. Slutsky *et al.*, Classical Quantum Gravity **27**, 165023 (2010).

[60] N. Christensen (for the LIGO Scientific Collaboration and the Virgo Collaboration), Classical Quantum Gravity **27**, 194010 (2010).

[61] J. Aasi *et al.* (Virgo Collaboration), Classical Quantum Gravity **29**, 155002 (2012).

[62] D. MacLeod, S. Fairhurst, B. Hughey, A. P. Lundgren, L. Pekowsky, J. Rollins, and J. R. Smith, Classical Quantum Gravity **29**, 055006 (2012).

[63] M. Ito, "Glitchmon: A DMT monitor to look for transient signals in selected channels," developed using the LIGO Data Monitoring Tool (DMT) library.

[64] L. Pekowsky, Ph.D. thesis, Syracuse University, 2012.

[65] J. Abadie *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. D **83**, 122005 (2011); and **86**, 069903(E) (2012).

[66] E. Amaldi *et al.*, Astron. Astrophys. **216**, 325 (1989).

[67] M. Was, M.-A. Bizouard, V. Brisson, F. Cavalier, M. Davier, P. Hello, N. Leroy, F. Robinet, and M. Vavoulidis, Classical Quantum Gravity **27**, 015005 (2010).

[68] T. Dent *et al.* (unpublished).

[69] D. A. Brown (for the LIGO Scientific Collaboration), Classical Quantum Gravity **21**, S1625 (2004).

[70] I. Mandel and R. O'Shaughnessy, Classical Quantum Gravity **27**, 114007 (2010).

[71] J. Abadie *et al.* (LIGO Scientific Collaboration and the Virgo Collaboration), arXiv:1003.2481.

[72] J. Abadie *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), arXiv:1203.2674.

[73] P. R. Brady, J. D. E. Creighton, and A. G. Wiseman, Classical Quantum Gravity **21**, S1775 (2004).

[74] R. Biswas, P. R. Brady, J. D. E. Creighton, and S. Fairhurst, Classical Quantum Gravity **26**, 175009 (2009).

[75] D. Keppel, Ph.D. thesis, Caltech, Pasadena, CA, 2009.

[76] E. S. Phinney, Astrophys. J. **380**, L17 (1991).

[77] A. Pai, S. Dhurandhar, and S. Bose, Phys. Rev. D **64**, 042004 (2001).

[78] J. Abadie *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. D **81**, 102001 (2010).

[79] M. Briggs *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Astrophys. J. **760**, 12 (2012).

[80] K. Cannon, C. Hanna, and D. Keppel, arXiv:1209.0718.

[81] C. Messenger and J. Veitch, arXiv:1206.3461.

[82] J. Abadie *et al.* (LIGO Scientific Collaboration), Astron. Astrophys. **541**, A155 (2012).

[83] K. Cannon, A. Chapman, C. Hanna, D. Keppel, A. C. Searle, and A. J. Weinstein, Phys. Rev. D **82**, 044025 (2010).

[84] K. Cannon, C. Hanna, D. Keppel, and A. C. Searle, Phys. Rev. D **83**, 084053 (2011).

[85] K. Cannon, C. Hanna, and D. Keppel, Phys. Rev. D **84**, 084003 (2011).

[86] K. Cannon, R. Cariou, A. Chapman, M. Crispin-Ortuzar, N. Fotopoulos *et al.*, Astrophys. J. **748**, 136 (2012).

[87] B. Metzger and E. Berger, Astrophys. J. **746**, 48 (2012).

[88] P. Ajith *et al.*, Phys. Rev. Lett. **106**, 241101 (2011).

[89] L. Santamaria *et al.*, Phys. Rev. D **82**, 064016 (2010).

[90] P. Ajith, Phys. Rev. D **84**, 084037 (2011).

[91] M. van der Sluys, V. Raymond, I. Mandel, C. Röver, N. Christensen, V. Kalogera, R. Meyer, and A. Vecchio, Classical Quantum Gravity **25**, 184011 (2008).

[92] M. V. van der Sluys, C. Röver, A. Stroeer, V. Raymond, I. Mandel, N. Christensen, V. Kalogera, R. Meyer, and A. Vecchio, Astrophys. J. **688**, L61 (2008).

[93] J. Veitch and A. Vecchio, Phys. Rev. D **81**, 062003 (2010).

[94] F. Feroz, M. P. Hobson, and M. Bridges, Mon. Not. R. Astron. Soc. **398**, 1601 (2009).