Search for gravitational-wave bursts in the third Advanced LIGO-Virgo run with coherent WaveBurst enhanced by machine learning

Marek J. Szczepańczyk,[1] Francesco Salemi,[2,3,*] Sophie Bini,[2,3] Tanmaya Mishra,[1] Gabriele Vedovato,[4] V. Gayathri,[1] Imre Bartos,[1] Shubhagata Bhaumik,[1] Marco Drago,[5,6] Odysse Halim,[7,8] Claudia Lazzaro,[4,9] Andrea Miani,[2,3] Edoardo Milotti,[7,8] Giovanni A. Prodi,[10,3] Shubhanshu Tiwari,[11] and Sergey Klimenko[1]

[1]Department of Physics, University of Florida, PO Box 118440, Gainesville, Florida 32611-8440, USA
[2]Università di Trento, Dipartimento di Fisica, I-38123 Povo, Trento, Italy
[3]INFN, Trento Institute for Fundamental Physics and Applications, I-38123 Povo, Trento, Italy
[4]Università di Padova, Dipartimento di Fisica e Astronomia, I-35131 Padova, Italy
[5]Università di Roma La Sapienza, I-00185 Roma, Italy
[6]INFN, Sezione di Roma, I-00185 Roma, Italy
[7]Dipartimento di Fisica, Università di Trieste, I-34127 Trieste, Italy
[8]INFN, Sezione di Trieste, I-34127 Trieste, Italy
[9]INFN, Sezione di Padova, I-35131 Padova, Italy
[10]Università di Trento, Dipartimento di Matematica, I-38123 Povo, Trento, Italy
[11]Physik-Institut, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

This paper presents a search for generic short-duration gravitational-wave (GW) transients (or GW bursts) in the data from the third observing run of Advanced LIGO and Advanced Virgo. We use a coherent WaveBurst (cWB) pipeline enhanced with a decision-tree classification algorithm for more efficient separation of GW signals from noise transients. The machine-learning (ML) algorithm is trained on a representative set of noise events and a set of simulated stochastic signals that are not correlated with any known signal model. This training procedure preserves the model-independent nature of the search. We demonstrate that the ML-enhanced cWB pipeline can detect GW signals at a larger distance than previous model-independent searches. The sensitivity improvements are achieved across the broad spectrum of simulated signals, with the goal of testing the robustness of this model-agnostic search. At a false-alarm rate of one event per century, the detectable signal amplitudes are reduced up to almost an order of magnitude, most notably for the single-cycle signal morphologies. This ML-enhanced pipeline also improves the detection efficiency of compact binary mergers in a wide range of masses, from stellar mass to intermediate-mass black holes, both with circular and elliptical orbits. After excluding previously detected compact binaries, no new gravitational-wave signals are observed for the twofold Hanford-Livingston and the threefold Hanford-Livingston-Virgo detector networks. With the improved sensitivity of the all-sky search, we obtain the most stringent constraints on the isotropic emission of gravitational-wave energy from short-duration burst sources.

I. INTRODUCTION

The first detection of a gravitational wave (GW) signal, GW150914 [1], in 2015, during the earliest observing run of the Advanced LIGO detectors [2], opened a new observation window to study the Universe. In the following runs, the Advanced LIGO and Advanced Virgo [3] detectors have recorded around a hundred GW signals [4–7]. All these detected GW sources are interpreted as compact binary coalescences (CBC), with the majority being binary black hole (BBH) mergers. However, gravitational waves from other, less understood GW sources, still remain elusive [8,9].

Several astrophysical sources are predicted to produce short bursts of GW radiation with a duration of up to a few seconds. They could be roughly divided into two categories. The first category includes late inspiral and merger waves from CBCs comprised of black holes and/or neutron stars. Types of CBC signals include binary systems with circular orbits [10–14], eccentric orbits [15–21], head-on collisions [22,23], extreme mass-ratio binaries [24], primordial black holes [25–28], and hyperbolic encounters [29–31]. Gravitational waves may also be generated in the postmerger phase of the binary neutron star systems [32,33]. The second category includes all other anticipated burst progenitors, such as star explosions, neutron star

glitches, or unknown sources. Prime examples of the star explosions are core-collapse supernovae (CCSNe) [34–41], hypernovae [42], superluminous supernovae [43], supernovae type Ia [44], and pulsation pair-instability supernovae [45–47]. Also, gravitational waves may be produced during a black-hole formation from a collapsing star [48–50], quantum chromodynamics phase transition in a core-collapse supernova [51], or in a collapsar [52,53]. The resulting gravitational waves from these source types are typically predicted to be stochastic. Neutron stars or pulsar glitches are expected to generate ringdown GW signals [54–57]. Other potential burst sources are the soft gamma repeaters [58] and cosmic strings (CSs) [59,60].

For most of the GW sources mentioned above, the simulations of accurate and computationally efficient GW waveforms (templates) are not readily available, which limits the use of the matched filter methods. This is the reason why the burst algorithms [61–67], designed to detect a wide range of GW sources without templates, are actively used for GW searches in data collected by the LIGO [2], Virgo [3], KAGRA [68], and GEO600 [69] detectors.

Coherent WaveBurst (cWB) [61] has been used to search for burst signals [70–76] and BBH signals [4,5,7,77–81]. It has made major contributions in the discovery of the first GW signal GW150914 [1,82] and the first intermediate-mass black hole merger GW190521 [83,84]. Also, cWB contributed to the gravitational wave transient catalogs (GWTCs) [4,5,7], to population studies of IMBH sources [85,86], and the search for specific CBC properties, such as higher-order multipoles [87,88].

By using a few assumptions about GW signals, cWB identifies the excess power triggers in the data from multiple detectors. A reduction of the false-alarm rate is accomplished with a postproduction veto procedure making use of summary statistics calculated for each trigger. This paper presents a generic burst search using cWB enhanced with machine learning (ML) to improve the separation of GW signals from the noise transients. Similar to the standard veto procedure, supervised classification algorithms can achieve an efficient reduction of the cWB false-alarm rate. In [89] a neural network method analyzing the time-frequency patterns of reconstructed cWB triggers was suggested to improve the detection of BBH signals. In [90] a machine-learning method was used to improve the identification of CCSNe. In both cases, a strong model dependence was introduced at the cWB postproduction stage, limiting the pipeline detection to a specific source used for the ML training. A more practical weakly modeled classification based on the XGBoost decision trees was introduced in [91,92] to enhance the cWB detection of the BBH signals. Unlike template-based searches that find events based on a specific signal model, the XGBoost classification is designed to penalize events inconsistent with generic signal features. More recently, a different signal-noise classification based on the Gaussian

mixture model (GMM) methodology was developed for generic burst searches with cWB [93]. In this approach, the GMM likelihood model is tuned on a separate subset of the simulated events used to benchmark the performances, sparsely sampling the ranges of burst frequency and duration. Instead, in this paper, we use the XGBoost decision-tree algorithm, which is trained on a representative set of noise events and a set of simulated signals with stochastic waveforms that are not correlated with any known signal model. The training set is densely covering the frequency band and the range of burst duration selected for the analysis. We demonstrate the sensitivity improvement across a broad spectrum of simulated signals not used for the XGBoost training. With this more sensitive algorithm, we reanalyze the data from the third observing run (O3) of the LIGO and Virgo detectors, targeting bursts in the frequency band up to 1 kHz.

The paper is structured as follows. Section II describes the ML-enhanced cWB pipeline, the data analyzed, and simulated signals used for training the ML algorithm. Section III describes signal models used for testing the search sensitivity. Section IV provides results of the ML-enhanced cWB search compared with the LIGO-Virgo-KAGRA (LVK) all-sky search for short duration bursts [8].

II. METHOD

A. Coherent WaveBurst pipeline enhanced with machine learning

Coherent WaveBurst (cWB) is a burst search pipeline that can detect GW signals from astrophysical sources without templates [61–65]. The cWB algorithm identifies excess-power triggers in the time-frequency data obtained with the multiresolution Wilson-Daubechies-Meyer (WDM) wavelet transform [94] of the strain data from multiple detectors. By using the constrained likelihood method [61], it reconstructs the source location in the sky and the signal waveform: incoming gravitational waves should produce coherent responses over the detector network, while in general, the noise events are not correlated. The rate of triggers produced by the pipeline is controlled by model-independent veto thresholds on the excess-power statistics and the coherent amplitude η_0 that characterizes the coherence of a cWB trigger across the detector network. The veto thresholds are set to be sufficiently low to avoid costly reruns of the pipeline trigger production. Further reduction of the cWB false-alarm rate is achieved with a more stringent veto procedure by applying thresholds on the summary statistics calculated for each trigger, including the η_0, the cross-correlation coefficient between the detectors, the reconstructed signal-to-noise ratio in each detector, the number of WDM resolutions used for the reconstruction and the waveform shape parameters (see the Appendix for more details).

Although this veto procedure generally works quite efficiently, designing vetoes in the multidimensional space

of summary statistics is challenging and requires re-tuning the veto thresholds for each detector network configuration and observing run. In addition, due to distinctive noise sources, cWB triggers are typically split into different categories (so-called "search bins") and analyzed separately. To solve these issues, a boosted decision-tree algorithm, eXtreme-Gradient Boosting (XGBoost) [95], was adopted and implemented within the cWB framework to automate the signal-noise classification of cWB triggers [91]. Two types of input data are used: signal events from simulations and noise events from background estimations. For each event, a selected subset of cWB summary statistics (see the Appendix for more details) is used by XGBoost as input features to train a signal-noise model. This ML method simplifies the analysis approach by jointly analyzing all candidate events within a single search bin. As described in Refs. [91,92], the detection statistic for the ML-enhanced cWB algorithm is defined by:

$$\eta_r = \eta_0 \cdot W_{XGB}, \qquad (1)$$

where η_0 is the cWB ranking statistic and W_{XGB} is the XGBoost penalty factor ranging between 0 (noise) and 1 (signal).

This study extends the ML-enhanced cWB method [91,92] to generic burst searches in the frequency band [16, 1024] Hz. In order to preserve the model-independent nature of cWB and be sensitive to generic GW signals, when training the XGBoost, we select as input features the summary statistics that do not depend on the waveform morphology (see the Appendix for details). Moreover, we do not train XGBoost on anticipated GW signals that follow specific astrophysical distributions. Instead, we use a stochastic set of band-limited white-noise-burst (WNB) signals: i.e., white noise contained within an *ad-hoc* time-frequency range. The WNB signals populate a wide range of randomly chosen signal durations and frequency bands. To that end, we use two sets of WNB for training: (a) WNBs uniformly distributed in central frequency in the range [24,996] Hz, bandwidth [10,300] Hz, and duration is logarithmically distributed between 0.1 ms and 1 ms; and (b) WNBs logarithmically distributed in central frequency [24,450] Hz, bandwidth [10,250] Hz and duration [1,50] ms. The WNB signals are explicitly chosen to span a space of two parameters (Q_a, Q_p): Q_a is an estimator of how much energy occurs outside the largest oscillation of the event's waveform and Q_p is a parameter linked to the event number of cycles [91,92]. Within this simplified 2-D parameters space, low-Q noise transients can be better identified by the ML classifier and, therefore, penalized. Furthermore, this approach efficiently removes one of the most challenging noise sources—blip glitches [8,96]. The details of the changes introduced to the ML method for improving the sensitivity toward generic burst sources are described in the Appendix.

B. Data

The third observing run of the Advanced LIGO, Hanford (H) and Livingston (L), and Advanced Virgo (V) detectors consists of two epochs separated by a commissioning break: O3a (from April 1, 2019, to October 1, 2019) and O3b (from November 1, 2019, to March 27, 2020). The two epochs have significantly different rates of short-duration glitches.

In this paper, we analyze HL and HLV detector networks. The HV and LV networks are omitted due to their lower sensitivities and shorter additional observing time with respect to the HL network. The threefold network generally allows a more accurate characterization of the GW signal than a two-detector network. Previous burst searches on O3 data, in particular the LVK one [8], did not report analysis of the HLV detector network because its detection performances were found to be not as good as the ones achieved using the HL detector network. This is due to concurrent causes, including differences in spectral and directional sensitivities between the two LIGO detectors and Virgo. After removing periods affected by poor environmental conditions and detector hardware anomalies, the analyzed duration of coincident data between LIGO detectors is 206.57 days (104.94 and 101.63 days for O3a and O3b, respectively). In comparison, it is 143.3 days (75.19 and 68.11 days) for HLV. We calculated the background dataset of our search by using the time-shift analysis. We accumulated 980.7 and 1096.0 years of data for the HL network for O3a and O3b, respectively, while it was 572.9 and 395.8 years for the HLV network. Around 50% of the background data is used for training the XGBoost model.

C. Metric statistics for comparing performances

The sensitivity of searches for GW bursts is usually measured in terms of detection efficiency as a function of root-squared-sum (rss) strain amplitude of the signal, h_{rss}[1]:

$$h_{rss} = \sqrt{\int_{-\infty}^{\infty} [h_+^2(t) + h_\times^2(t)]dt}, \qquad (2)$$

where h_+ and $h_\times$ are polarization components. Detection efficiency $\epsilon(h_{rss})$ is estimated by adding (i.e., injecting) selected GW signals in the detector noise over a wide range of amplitudes. For all-sky searches, the modeled source location and orientation in the sky are uniformly random. The resulting $\epsilon(h_{rss})$ is the fraction of detected events over injected ones at h_{rss} amplitude with a search threshold that ensures a minimum statistical significance, such as a minimum inverse false alarm rate (iFAR).

---

[1]More precisely, for isotropic emission h_{rss} indicates signal strength at Earth, while for anisotropic emissions we define h_{rss} as the maximum amplitude emitted by the source, inversely scaled by the source distance, as in [8].

The h_{rss} amplitude corresponding to 50% detection efficiency, h_{rss50}, has been widely used as a benchmark of the typical search sensitivity. Here we report h_{rss50} at iFAR ≥ 100 years as in [8], i.e., for rather significant detection candidates.

Detection efficiency can also be expressed as a function of the luminosity distance r by assuming a reference amplitude value of the GW emission, $h_{rss}(r_0)$ at some nominal distance r_0:

$$\epsilon(r) \approx \epsilon(h_{rss}(r_0)) \times h_{rss}(r_0)/h_{rss}(r). \tag{3}$$

This approximation is appropriate whenever the emission process of interest is close to that of a standard siren or its model sets a specific energy scale (as for CCSNe models). Equation (3) also describes how detection efficiency $\epsilon(r)$ scales for different assumptions on the $(r_0, h_{rss}(r_0))$ parameters, provided that waveform morphology and spectral sensitivities of detectors are unchanged.

Detectable volume, or sensitivity volume $\mathcal{V}$, is a benchmark more directly related to the detection probability of a source population, whose spatial and GW amplitude distributions are assumed to be known. In case the source is modeled as a standard siren with a uniform spatial density and rate, the sensitive volume can be estimated by integrating the detection efficiency[2] (see [70]):

$$\mathcal{V} = 4\pi(r_0 h_{rss}(r_0))^3 \int_0^\infty \frac{dh_{rss}}{h_{rss}^4} \epsilon(h_{rss}) \tag{4}$$

and volume $\mathcal{V}$ can be rescaled for different amplitude-distance relations $h_{rss}(r_0)$.

The following expression provides the energy radiated in GWs in case the emission is narrowband and isotropic:

$$E_{GW} = \frac{\pi^2 c^3}{G} r_0^2 f_0^2 h_{rss}^2(r_0), \tag{5}$$

where r_0 is the distance to the source and f_0 is the central frequency of the GW signal. This approximation is good to within a few % for all the ad-hoc signals listed in Sec. III A, apart from the broad-band Gaussian pulses. By using h_{rss50} in Eq. (5), one can benchmark the typical sensitivity of the search in terms of E_{GW} at a reference distance (50% detection efficiency, for the selected iFAR threshold). Conversely, assuming a source emitting E_{GW} around a peak frequency f_0, Eq. (5) can be used to estimate the product $r_0 h_{rss}(r_0)$:

$$r_0 h_{rss}(r_0) = \frac{\sqrt{G}}{\pi \sqrt{c^3 f_0}} \sqrt{E_{GW}}. \tag{6}$$

and knowledge of the h_{rss50} value allows estimating the typical range of the search.

---

[2]This formula is valid for source distances at negligible cosmological redshift.

III. SIGNAL MODELS

The generic burst searches are designed to be sensitive to a wide range of GW morphologies, and the ML-enhanced cWB is tested with various possible burst signals. This section describes *ad-hoc* signals and the waveforms derived from several astrophysical models (namely CCSNe, ringdowns, and cosmic strings) for a total of 53 tested signal morphologies, listed in Table I. For robustness, sensitivity to different types of compact binary systems is also studied.

A. *ad-hoc*

The *ad-hoc* signals estimate the algorithm's sensitivity to generic GW morphologies. They include sine-Gaussians (SGs), Gaussian pulses (GAs), and WNBs. The SGs are fully characterized by the central frequency f_0 and quality factor Q determining the signal's duration. The GAs have only one parameter, i.e., the duration of one standard deviation τ_{GA}. Finally, the WNBs are described by a lower frequency bound f_{low}, a bandwidth Δf, and a duration τ_{WNB}. More details on these signal morphologies can be found in [8,70].

B. Astrophysical

One of the most interesting astrophysical burst sources is CCSNe. These exploding stars are very challenging to model since all forces of nature on a micro- and macroscale impact the explosion. GWs from CCSNe are stochastic, and their typical duration is of the order of 1 s. The typical energies range between $10^{-10} M_\odot c^2$ and $10^{-7} M_\odot c^2$ with peak frequencies at around 1 kHz. More details about the properties of GWs from CCSNe can be found in [41]. We study waveforms from various models that explore progenitor star rotation and masses, explosion phases, energies and GW signatures. The 10 analyzed neutrino-driven explosion models are: Andresen *et al.* [97] (And + 17) s11, Kuroda *et al.* [98] (Kur + 16) TM1, Müller *et al.* [99] (Mul + 12) L15, O'Connor and Couch [38] (Oco + 18) mesa20, Powell and Müller [100] (Pow + 19) he3.5 and s18, Radice *et al.* [36] (Rad + 19) s9, s13, s25. Additionally, a magnetorotationally driven explosion is analyzed: Abdikamalov *et al.* [101] (Abd + 14) A4O01.0.

We study signals modeled as ringdowns (RDs), representing the final stages of BBH mergers. RDs are also expected from the excitation of fundamental modes in neutron stars [33], but their typical frequencies are around 2–3 kHz which is beyond our search range. The RD model follows cf. Eq. (3.6) in Ref. [70]:

$$h_+(t) \propto \exp(-t/\tau) \sin(2\pi f_0 t)$$
$$h_\times(t) \propto \exp(-t/\tau) \cos(2\pi f_0 t), \tag{7}$$

where f_0 is the central frequency, and τ is the decay time. We use f_0 to be: 70 Hz, 235 Hz, and 849 Hz, and τ is chosen to generate signals with half, one, and two cycles.

TABLE I. The $h_{\rm rss}$ values (in units of 10^{-22} Hz$^{-1/2}$) for which 50% detection efficiency and sensitive volumes $\mathcal{V}$ (in units of $[(\frac{h_{\rm rss,0}}{1e-22} r_0)^3]$) is achieved with an iFAR of 100 years for each of the injected signal morphologies. We combine O3a and O3b data, STD and XGB stands for standard search and ML-enhanced search, respectively.

| | HL network | | HLV network | |
| | $h_{\rm rss50}$ [1/$\sqrt{\rm Hz}$] | $\mathcal{V}$ [$(\frac{h_{\rm rss,0}}{1e-22} r_0)^3$] | $h_{\rm rss50}$ [1/$\sqrt{\rm Hz}$] | $\mathcal{V}$ [$(\frac{h_{\rm rss,0}}{1e-22} r_0)^3$] |
Morphology	STD/XGB	STD/XGB	STD/XGB	STD/XGB
Gaussian pulse				
$\tau = 4.0$ ms[a]	27.0/5.5	$9.6 \times 10^{-2}/2.5 \times 10^{-1}$	94.5/13.8	$7.2 \times 10^{-2}/8.2 \times 10^{-2}$
$\tau = 2.5$ ms[a]	16.7/4.2	$1.1 \times 10^{-1}/5.1 \times 10^{-1}$	31.8/10.9	$1.7 \times 10^{-1}/1.2 \times 10^{-1}$
$\tau = 1.0$ ms[a]	11.1/3.7	$1.1 \times 10^{-1}/1.5$	13.9/8.4	$9.7 \times 10^{-2}/3.5 \times 10^{-1}$
$\tau = 0.1$ ms[a]	12.6/3.6	$2.1 \times 10^{-1}/4.9 \times 10^{-1}$	17.5/11.7	$1.7 \times 10^{-1}/9.9 \times 10^{-2}$
Sine-Gaussian linear				
$f = 70$ Hz, $Q = 9$	1.5/1.4	5.5/7.1	1.4/1.4	5.0/5.8
$f = 100$ Hz, $Q = 9$	1.2/1.1	$9.8/1.2 \times 10^1$	1.2/1.2	8.2/9.0
$f = 235$ Hz, $Q = 9$	1.0/0.9	$1.4 \times 10^1/1.9 \times 10^1$	1.1/1.0	$1.1 \times 10^1/1.4 \times 10^1$
$f = 361$ Hz, $Q = 9$	1.2/1.1	$1.0 \times 10^1/1.5 \times 10^1$	1.3/1.2	$7.9/1.0 \times 10^1$
Sine-Gaussian circular				
$f = 70$ Hz, $Q = 3$	1.1/1.0	4.1/5.6	1.2/1.2	3.4/4.3
$f = 70$ Hz, $Q = 100$	1.1/1.0	4.5/6.6	1.1/1.0	4.4/5.6
$f = 153$ Hz, $Q = 9$	0.8/0.8	$1.1 \times 10^1/1.6 \times 10^1$	0.9/0.9	$9.2/1.1 \times 10^1$
$f = 235$ Hz, $Q = 3$	0.9/0.8	$9.9/1.4 \times 10^1$	0.9/0.9	8.1/8.8
$f = 235$ Hz, $Q = 100$	0.8/0.7	$1.3 \times 10^1/2.6 \times 10^1$	0.8/0.7	$1.1 \times 10^1/1.6 \times 10^1$
$f = 554$ Hz, $Q = 9$	1.1/1.0	4.3/6.2	1.2/1.2	3.5/4.2
$f = 849$ Hz, $Q = 3$	1.6/1.5	1.5/1.9	1.8/1.7	1.1/1.2
$f = 849$ Hz, $Q = 9$	1.4/1.3	2.1/3.1	1.6/1.5	1.6/1.9
$f = 849$ Hz, $Q = 100$	1.4/1.2	2.2/3.8	1.5/1.4	1.9/2.7
Sine-Gaussian elliptical				
$f = 70$ Hz, $Q = 3$	2.1/2.0	1.3/1.8	2.3/2.2	1.0/1.2
$f = 70$ Hz, $Q = 100$	2.0/1.8	1.5/2.0	1.9/1.7	1.4/1.3
$f = 153$ Hz, $Q = 9$	1.5/1.4	3.4/4.8	1.6/1.5	2.9/3.3
$f = 235$ Hz, $Q = 3$	1.6/1.5	2.6/3.6	1.7/1.8	2.5/2.7
$f = 235$ Hz, $Q = 100$	1.4/1.2	3.4/6.0	1.4/1.3	3.4/4.5
$f = 554$ Hz, $Q = 9$	2.0/1.8	1.4/1.9	2.3/2.1	1.0/1.3
$f = 849$ Hz, $Q = 3$	2.9/2.7	$4.5 \times 10^{-1}/5.5 \times 10^{-1}$	3.5/3.2	$3.4 \times 10^{-1}/4.0 \times 10^{-1}$
$f = 849$ Hz, $Q = 9$	2.6/2.4	$5.9 \times 10^{-1}/8.8 \times 10^{-1}$	3.0/2.8	$4.8 \times 10^{-1}/5.5 \times 10^{-1}$
$f = 849$ Hz, $Q = 100$	2.6/2.3	$6.3 \times 10^{-1}/1.1$	2.8/2.6	$5.6 \times 10^{-1}/7.1 \times 10^{-1}$
White Noise Burst				
$f = 150$ Hz	1.0/0.9	7.8/9.2	1.1/1.0	6.1/6.6
$f = 300$ Hz	1.0/1.0	6.2/8.0	1.2/1.1	4.9/5.8
$f = 700$ Hz	1.5/1.4	2.2/3.3	1.8/1.5	1.6/2.0
Supernova				
Mul + 12 L15	1.1/1.0	⋯	1.2/1.1	⋯
Abd + 14 A4O01.0	2.5/2.4	⋯	3.0/13.1	⋯
Kur + 17 SFHx	1.2/1.1	⋯	1.4/1.2	⋯
Rad + 18 s9	2.3/1.9	⋯	3.3/2.3	⋯
Rad + 18 s13	2.4/2.0	⋯	3.1/2.2	⋯
And + 19 s11	2.2/1.8	⋯	2.9/2.2	⋯
Oco + 18 mesa20_pert	3.4/2.8	⋯	4.9/3.5	⋯
Oco + 18 mesa20	3.9/3.3	⋯	5.5/4.7	⋯
Pow + 19 he3.5	2.8/2.4	⋯	3.9/2.6	⋯
Rad + 18 s25	3.9/3.3	⋯	5.1/3.7	⋯
Pow + 19 s18	3.0/2.4	⋯	4.2/2.7	⋯

(Table continued)

TABLE I. (*Continued*)

Morphology	HL network		HLV network	
	$h_{\rm rss50}$ [$1/\sqrt{\rm Hz}$]	$\mathcal{V}$ [$(\frac{h_{\rm rss,0}}{1e-22} r_0)^3$]	$h_{\rm rss50}$ [$1/\sqrt{\rm Hz}$]	$\mathcal{V}$ [$(\frac{h_{\rm rss,0}}{1e-22} r_0)^3$]
	STD/XGB	STD/XGB	STD/XGB	STD/XGB
Ringdown				
0.5 cycle, $f = 70$ Hz	2.0/1.6	1.8/2.4	3.5/7.6	$4.1 \times 10^{-1}/3.1 \times 10^{-1}$
1 cycle, $f = 70$ Hz	1.3/1.2	3.2/4.2	2.6/3.8	$6.7 \times 10^{-1}/7.8 \times 10^{-1}$
2 cycles, $f = 70$ Hz	1.1/1.1	3.8/4.7	2.3/1.8	$9.9 \times 10^{-1}/1.3$
0.5 cycle, $f = 235$ Hz	4.7/1.4	1.9/5.1	7.3/4.0	1.1/1.8
1 cycle, $f = 235$ Hz	1.1/0.9	$6.0/1.0 \times 10^1$	1.3/1.1	5.2/6.9
2 cycles, $f = 235$ Hz	0.9/0.8	$8.6/1.2 \times 10^1$	1.0/0.9	7.5/8.9
0.5 cycle, $f = 849$ Hz	2.6/2.0	$5.5 \times 10^{-1}/7.8 \times 10^{-1}$	2.3/3.1	1.4/1.3
1 cycle, $f = 849$ Hz	2.2/1.6	$8.5 \times 10^{-1}/1.8$	1.4/1.3	2.6/3.1
2 cycles, $f = 849$ Hz	1.8/1.5	1.2/2.1	1.2/1.2	3.3/3.8
Cosmic String				
$f_{\rm cutoff} = 50$ Hz[a]	208.3/49.8	$1.6 \times 10^{-4}/3.8 \times 10^{-4}$	336.4/246.7	$1.8 \times 10^{-4}/9.2 \times 10^{-5}$
$f_{\rm cutoff} = 150$ Hz	133.5/48.8	$1.7 \times 10^{-5}/4.7 \times 10^{-4}$	180.2/117.9	$9.0 \times 10^{-6}/1.3 \times 10^{-4}$
$f_{\rm cutoff} = 500$ Hz	119.6/52.0	$2.3 \times 10^{-5}/6.0 \times 10^{-4}$	155.6/114.0	$1.1 \times 10^{-5}/1.7 \times 10^{-4}$
$f_{\rm cutoff} = 1500$ Hz	114.6/50.1	$2.3 \times 10^{-5}/6.0 \times 10^{-4}$	148.4/106.5	$1.1 \times 10^{-5}/1.7 \times 10^{-4}$

[a]For these signal morphologies, both $\mathcal{V}$ values for the standard cWB search were estimated directly using the data points, not from the detection efficiency fit of the data (for more details see Sec. IV B, sensitivity volume).

Other potential burst sources are cosmic strings (CS). CSs are one-dimensional topological defects possibly formed after a spontaneous symmetry phase transition in the early Universe and are usually described in grand unified theories. CS cusps and kinks propagating on string loops are expected to generate GW bursts. The CS models allow generating templates and matched filtering was used in the recent LVK search for CSs [102]. In our analysis, we adopt GW waveforms from CS cusps [60], that have been used in the search for CSs with O1 LIGO data [103] and are characterized by the amplitude, low-frequency cut-off of 1 Hz and high-frequency cut-off $f_{\rm cutoff}$ of 50 Hz, 150 Hz, 500 Hz, and 1500 Hz.[3] For the first time, we report the cWB sensitivity on these sources, both for the standard and the ML-enhanced search. These results cannot be directly compared to the Ref. [102] where the detection efficiencies are a function of signal amplitude, not $h_{\rm rss}$.

The first set of binaries is BBH systems with quasicircular orbits. The waveforms are calculated using effective-one-body SEOBNRv4 model [104] that includes only dominant $(l, m) = (2, 2)$ mode. The component masses span from $5M_\odot$ to $100M_\odot$, following the power law + peak [105] mass function. The mass ratio ranges between 1/4 to 1. Signals have a spin aligned with the

orbital angular momentum with a magnitude randomly distributed between −0.99 and 0.99.

The second set of binaries is intermediate-mass black hole systems. Similarly to [106], we use 46 IMBH binaries with a total mass between $120M_\odot$ and $800M_\odot$, and a mass ratio from 1 to 1/10. The systems are precessing, with aligned and antialigned spins. These waveforms are derived from numerical relativity simulations ([107–109]). Each IMBH binary is uniformly distributed in sky location, inclination angle, and within a comoving volume, optimized on the signal strength: for each signal waveform, we fix the maximum redshift by calculating conservative upper bounds on the optimal three-detector network signal-to-noise ratio (SNR) to avoid generating injections well outside any possible detection range.

The third set contains binaries with eccentric orbits. The eBBH systems could be formed through gravitational capture in close encounters, which causes small orbital separation between the component black holes and high initial orbital eccentricity. We consider 35 mass bins which range between $100M_\odot$ and $250M_\odot$, mass ratio 1, and eccentricities between 0.19–0.96. The simulated eBBH signals used in this analysis are obtained from numerical relativity simulations from [108].

IV. RESULTS

This section discusses the O3 results of the ML-enhanced cWB search and the sensitivity of the twofold and threefold detector networks on the broadest set of

---

[3]Since CS signals have a red spectrum, most of their signal energy is cut by the lower frequency limit of our search. In our operating conditions, only 10% of the total $h_{\rm rss}$ amplitude of the CS waveforms falls inside the analyzed frequency band.
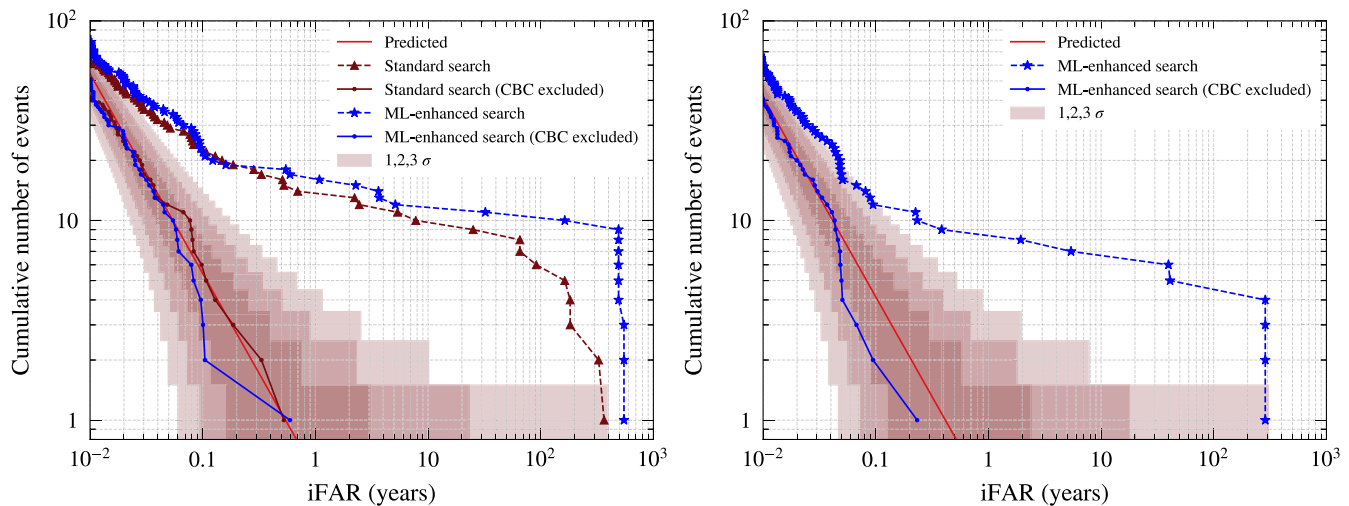
FIG. 1. Cumulative number of events versus iFAR found by the standard (brown dashed line) and ML-enhanced searches (blue dashed line) in O3. The red solid line shows the expected mean value of the background and the shaded regions are the 1σ, 2σ, and 3σ Poisson uncertainty intervals. Left: results for the HL network. At iFAR ≥ 1 year, the ML-enhanced search detected 16 CBC events compared to 14 events for the standard search. Right: results for the HLV network, the standard search was not performed, while 10 events are detected with iFAR ≥ 1 year. In both panels, the loudest events' significance saturates due to the limited amount of background available for testing ($1/2$ was already used for training), i.e., roughly 500 (300) years for the HL (HLV) network. After removing the known CBC events (continuous lines), the ML-enhanced search reports a null result for both networks.

signal morphologies performed to date. We compare our results with the standard generic burst search performed by the LVK collaboration [8]. More specifically, we used the same standard cWB methods, including the same three search bins to rank separately volumes of signal parameter space dominated by very different false alarm distributions. Due to the different noise floors and noise transients during O3, the analysis performed by LVK generic burst search considered O3a and O3b separately. In this work, we train two separate XGBoost models for O3a and O3b, respectively. For brevity, all results are reported for the full O3.

A. GW detections

All detected GW transients are already established CBC emissions [7], and no evidence is found for other source classes.

Figure 1 presents our search results compared to Ref. [8] for the HL detector network. At iFAR larger than one year, the present analysis detects more known CBC sources compared to the standard (i.e., 16 with respect to 14), and the significance of 12 out of the 14 common detections is increased. The two CBC detections, which were previously missed by the standard search, are GW190602 with iFAR = 3.7 years, SNR = 11.4 and 48 solar masses, and GW191230 with iFAR = 1.08 years, SNR = 10.4, 36 solar masses. In addition, the ML-enhanced search identified 14 known CBC sources as subthreshold events (iFAR smaller than one year) compared to 7 in the standard analysis. After excluding the known CBC sources, the remaining on-source events are compatible with the

estimated background, and the loudest event shows an iFAR of 0.59 years, comparable to the net observing time (62% false alarm probability).

This work reports for the first time the O3 detections by a search for GW bursts using the three-detector network, HLV, see right panel of Fig. 1. The HLV search detected 8 CBC sources with iFAR larger than 1 year, with comparable or lower significance with respect to the HL detections (which is consistent with the discussion at the end of Sec. IV C). After excluding the known CBC sources, the remaining on-source events are consistent with the estimated background.

B. Sensitivity of two-detector network HL

Sensitivity studies show that the ML-enhanced search is systematically improving performances across the tested signal models, therefore preserving the model-agnostic character of the search. Here we focus on the sensitivity achieved on the LIGO Hanford-Livingston network.

Amplitude sensitivity. Table I and Fig. 2 report the comparison between the standard and ML-enhanced searches in terms of typically detectable amplitude, h_{rss50} at iFAR ≥ 100 years, for *ad-hoc* and astrophysical waveforms. The XGBoost postproduction improves the detection efficiency for all the 53 signal morphologies considered. The greatest improvement is achieved for the waveforms with a few cycles similar to low-Q noise transients (see Sec. II), GAs and CSs. The reduction of h_{rss50} with respect to the standard search, $h_{\mathrm{rss50,XGB}}/h_{\mathrm{rss50,STD}}$ is around 0.20–0.33 for GAs and 0.24–0.44 for CSs. The linear SGs show an
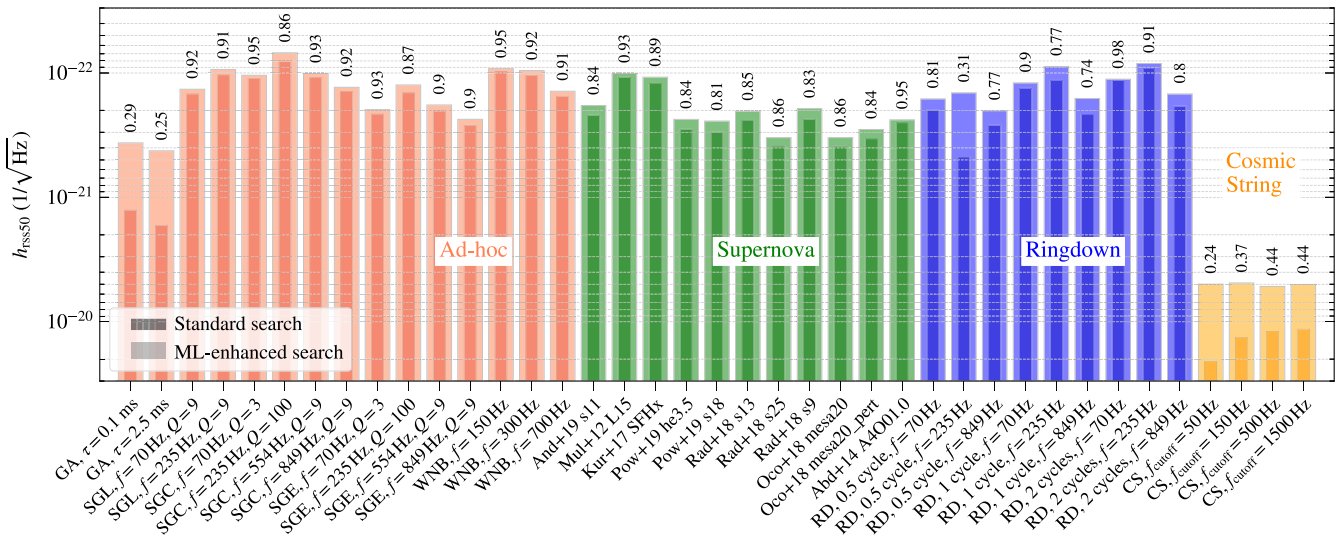
FIG. 2. Resulting h_{rss50} achieved with cWB with standard postproduction veto procedure (darker colors) and with ML-enhanced cWB (lighter colors) for the HL network on full O3 and at iFAR $\geq$ 100 years. The waveforms reported are a subset of those listed in Table I: ad-hoc signals ordered according to central frequency (red), core-collapse supernovae (green), ringdown waveforms (blue), and cosmic strings (yellow). The values on the top show the reduction factor on h_{rss50} with respect to the standard search; h_{rss50} ordinate scale decreases going upwards.

improvement of 0.90–0.94, circular SGs 0.84–0.95, elliptical SGs 0.87–0.93, and WNBs 0.91–0.95. Typically the detectable amplitude improves between 0.30 and 0.98 for ringdowns and 0.81–0.94 for CCSN waveforms. Overall, the ML-enhanced search achieves more homogeneous values of h_{rss50}, having decreased the gap of sensitivity between the least-performing and best-performing waveform families. In particular, taking into account that for the CSs only 10% of the injected h_{rss} falls inside the analyzed frequency band, amplitude sensitivities to CSs are now comparable to the GA cases.
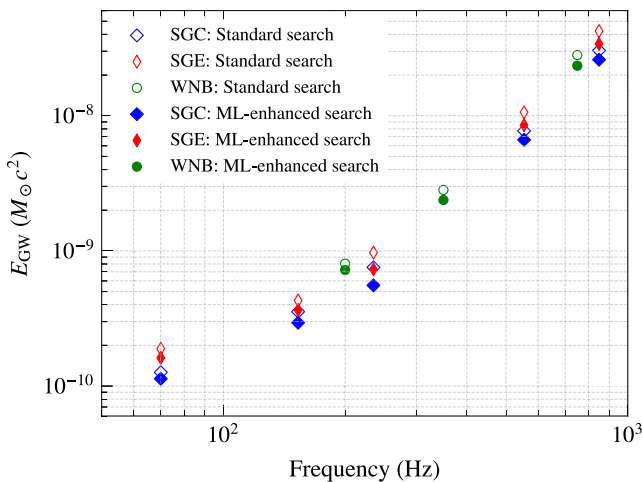
Figure 3 reports the typical detectable energy radiated in GWs for *ad-hoc* waveforms, evaluated according to the Eq. (5) by assuming a source distance of 10 kpc, 50% detection efficiency and iFAR $\geq$ 100 years. Due to the null detection results for non-CBC GW transients, this can be interpreted as constraints on the product of luminosity distance and amplitude for burst sources. The ML-enhanced cWB improves significantly the constraints across the frequency spectrum for all tested morphologies.

Detection range. As mentioned in Sec. II C, in the case of CCSNe models the detection efficiency can be reported as a function of luminosity distance according to Eq. (3). Figure 4 shows examples for three CCSN models.
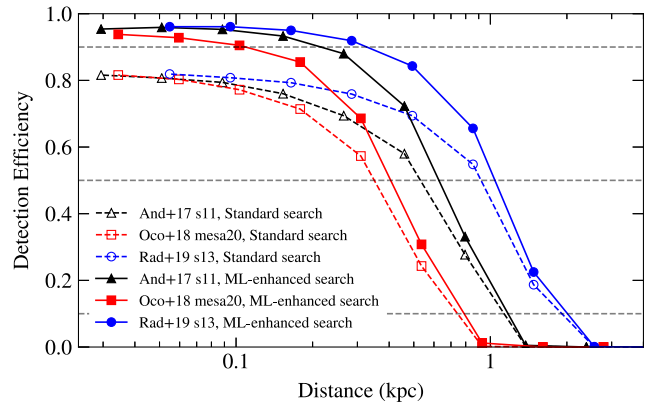

FIG. 3. Radiated energy in GWs at 50% detection efficiency and iFAR $\geq$ 100 years for a source distance of 10 kpc. The ML-enhanced cWB improves the constraints across the frequency spectrum for all tested morphologies.

FIG. 4. Detection efficiency vs distance for sample supernova waveforms, for HL network at iFAR $\geq$ 100 years. The ML-enhanced search improves the detection distance at 50% detection efficiency; the probability of detections at a closer distance increases significantly.
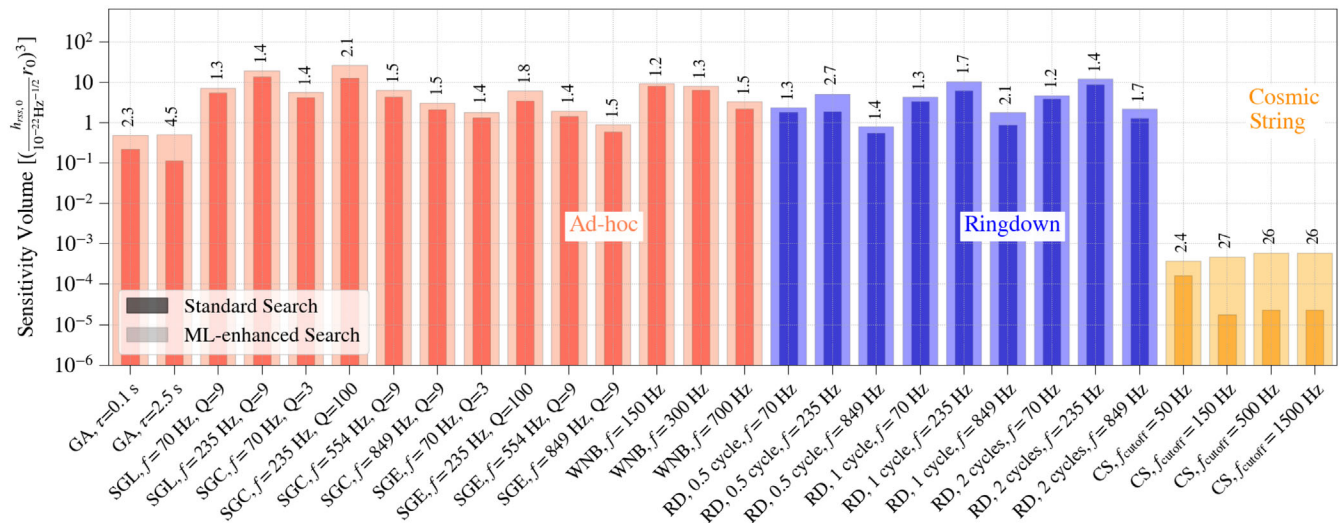
FIG. 5. Sensitivity volume obtained with cWB standard postproduction veto procedure (darker colors) and with ML-enhanced cWB (lighter colors) for HL network on full O3 data, at iFAR ≥ 100 years. The ordinate reports relative sensitive volumes normalized by $4\pi r_0^3$, where r_0 is the distance at which the source emits the reference amplitude $h_{rss} = 10^{-22}$ Hz$^{-1/2}$. We use this standard siren value across all reported signal models to highlight dependencies on signal morphology. The values on the top show the gain in the space volume $\mathcal{V}_{XGB}/\mathcal{V}_{STD}$. From left to right, the waveforms reported are ad-hoc signals ordered according to frequency (red), ringdown waveforms (blue), and cosmic strings (yellow).

The distances at 50% detection efficiencies are typically around 1 kpc. Similarly to the improvements in the detectable h_{rss50}, ML-enhanced cWB allows for an average increase in the detection distance. Moreover, the enhanced algorithm significantly increases the number of detected signals at closer distances.

Sensitivity volume. Figure 5 presents a volumetric benchmark for *ad-hoc* and astrophysical waveforms, based on Eq. (4). This benchmark is a proxy for the expected detection rate in case of sources uniformly distributed in volume and with a detection range within small cosmological redshifts. The sensitivity volume is mainly determined by the detection efficiency $\epsilon(h_{rss})$ in the lower h_{rss} range, typically well below h_{rss50}. This benchmark is estimated using Eq. (4) by fitting $\epsilon(h_{rss})$ over a measured grid of h_{rss} values. In a few cases, namely the GA waveforms and the CS $f_{cutoff} = 50$ Hz, the fits fail and Eq. (4) is estimated directly from the data points. Results in Fig. 5 assume a standard siren model with equal strength across all tested signal models to highlight dependencies on the signal morphology. We do not report sensitivity volumes for CCSNs since each model requires a specific emission strength and current detection ranges are not compatible with the assumption of uniform distribution in volume.

The sensitivity volume shows a systematic improvement with respect to the standard search, with the volume ratio $\mathcal{V}_{XGB}/\mathcal{V}_{STD}$ ranging between 1.2–1.5 for WNBs, 1.3–2.1 for SGs, and 1.2–2.7 for RDs. Here too, the most substantial improvements are achieved for GAs, 2.3–4.5, and CSs, 2.4–27, confirming that the ML-enhanced search

greatly improves its discriminating power against blip glitches and allows to achieve more homogeneous performances across tested signals. The depletion of CSs' results compared to other signals is mostly due to the fact that only a fraction of their injected amplitudes (10%) falls inside the analyzed frequency band.

Receiver operating characteristics curves for compact binaries. We extend the robustness test to show that the ML-enhanced cWB for generic transients is also providing a better sensitivity to CBC sources than the standard
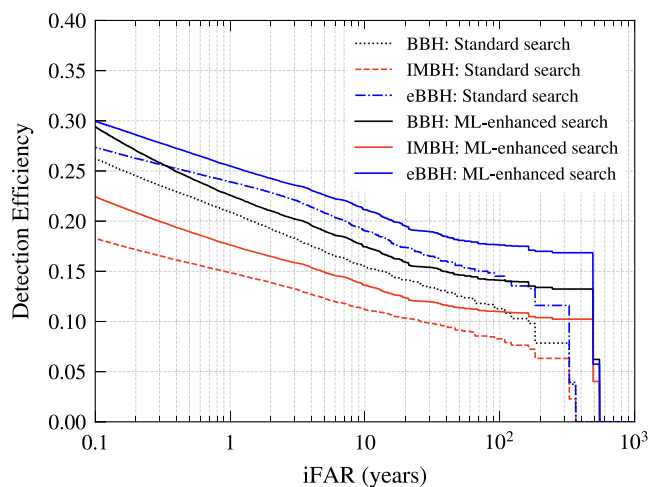

FIG. 6. Detection efficiency versus iFAR for BBH (black), IMBH (red), and eBBH (blue). ML-enhanced cWB (solid lines) shows an increase in detection efficiency with respect to the standard search (dashed lines).
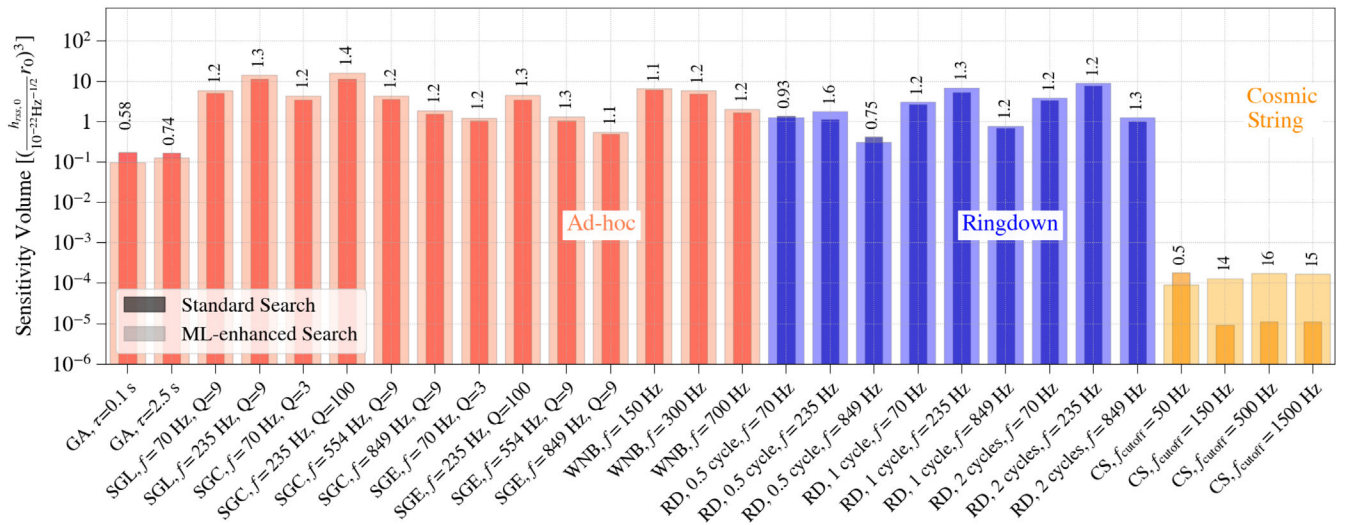
FIG. 7. Sensitivity volume obtained with cWB standard postproduction veto procedure (darker colors) and with cWB ML-enhanced (lighter colors) for HLV network during full O3 run, at iFAR=100 years. The values above each column show the space volume gain $\mathcal{V}_{XGB}/\mathcal{V}_{STD}$. From left to right, the waveforms reported are: *ad-hoc* signals ordered according to frequency (red), ringdown waveforms (blue), and cosmic strings (yellow).

method for bursts. This test is performed on three CBC simulation sets, stellar mass BBHs, IMBHBs, and eBBHs, each one spanning different regions in parameter spaces and fiducial volumes as described in Sec. III B. The receiver operating characteristics curves for each set are shown in Fig. 6. Detection efficiencies improve across the entire iFAR threshold range, detecting around 8%–25% more from our BBH set, 18%–32% more from our IMBH set, and 8%–21% more from our eBBH set.[4] The results of this test are found to be consistent with the increased number of CBC detections presented in Sec. IV A.

C. Sensitivity of three-detector network HLV

We test the ML-enhanced cWB on HLV data using the same set of simulated signals as for the HL network. For brevity, we report the results only in terms of h_{rss50} and sensitivity volume (Table I, Fig. 7). Compared to the standard HLV analysis, the ML-enhanced cWB improves the h_{rss50} for 48 out of the 53 considered waveforms.[5] Also,

the sensitive volume $\mathcal{V}$ improves for 36 out of the 42 considered waveforms.

Analyzing the HLV results for O3a and O3b separately, we observe that ML-enhanced cWB search demonstrates better performance in O3a. The O3b analysis is more challenging due to a significantly higher rate of short-duration noise transients. We remark that in order to take full advantage of the addition of Virgo data, the analysis should search for both GW polarization components. In order to preserve the generality of the search, those polarization components cannot be assumed to be correlated. This is not the case for the HL network, where both instruments sense approximately the same GW polarization component from most sky directions. Therefore, the HLV analysis provides a less efficient discrimination of glitches, resulting in a lower detection efficiency of GW signals at a given statistical confidence.

V. CONCLUSIONS

This paper reports a search for generic GW burst sources on the data from the third observing run of the Advanced LIGO and Virgo detectors. We use coherent WaveBurst pipeline enhanced with machine learning to discriminate noise from signal events. The ML algorithm is trained using generic signals with stochastic morphology that do not match any known signal model. This procedure preserves the model-independent nature of the search.

The ML enhancement substantially improves the all-sky search for burst sources. We achieved a systematic improvement of sensitivity across the very broad set of tested signal morphologies, including *ad-hoc* signal classes and waveforms from astrophysical sources (core-collapse supernovae, cosmic strings and ringdown waveforms).

---

[4] Our receiver operating characteristic curves cannot be compared across different sets, nor are meant to be representative of a sensitive volume for each set. This is so because each set is built under very different assumptions on rate densities and distributions of intrinsic parameters of the sources.

[5] While a light degradation of performances for a few waveforms can be considered acceptable when using ML-enhanced cWB, in the case of the Abdikamalov supernova waveform [101], the h_{rss50} increases by a factor of 4. The simulated events for such waveform have a time-frequency evolution similar to short-duration transient noise and belong to a Qa-Qp region not entirely covered by our signal training set. This resulted in a poor separation from blip glitches on O3b data in the HLV configuration, see the Appendix.

The most significant improvement in sensitivity compared to the LVK O3 search [8] is achieved for low-Q factor signals, such as cosmic strings and Gaussian pulses. In fact, the ML-enhanced algorithm is more successful in discriminating the dominant class of noise events such as blip glitches. Detection performances are measured in terms of the detectable signal amplitudes, distance ranges and space volumes. In particular, detectable amplitudes improve from a few percent to almost an order of magnitude. The ML-enhanced cWB is also more efficient in reconstructing compact binary coalescences with stellar mass and intermediate-mass black holes. These tests demonstrate that the algorithm robustly detects a wide range of short-duration GW transients.

This search detects more compact binaries than the previous cWB burst search, improving their significance. The results of the threefold LIGO-Virgo detector network with O3 data are reported here for the first time. After the compact binaries are excised, we report no evidence for new GW transients for both twofold LIGO and threefold LIGO-Virgo detector networks.

APPENDIX: XGBOOST MODEL

The ML-enhanced cWB search was introduced in Ref. [91] and updated for the O3 BBH search in Ref. [92]. We use a decision tree-based ensemble learning classifying algorithm called eXtreme-Gradient Boosting (XGBoost). The cWB pipeline is utilized to reconstruct interesting events and generate summary statistics. A carefully selected subset of summary statistics is used as the list of input features for the ML model.

In this subsection, we discuss the necessary updates to the BBH/IMBH ML-enhanced cWB search [92], in order to tune it for the generic burst search:

While the cWB search described in Ref. [92] looks for compact binary coalescence signals that do not depend on accurate waveform modeling, the generic searches should be designed for an even wider range of burst morphologies. The definitions of the summary statistics used for the O3 BBH burst search can be found in Ref. [91]. We use the following subset of the summary statistics to train XGBoost: norm (n_f), cross-correlation coefficient (c_c), quality of event reconstruction (χ^2), square of SNR over likelihood (SNR_i/L; where i = 0 for the HL network, and $i = \{0, 1\}$ for the HLV network), effective correlated SNR (η_0). Additionally, we use shape parameters (Q_a, Q_p) which are developed to identify short-duration low-Q glitches. We exclude model-dependent summary statistics: central frequency (f_0), duration (ΔT_s), and bandwidth (ΔF_s) as they are strongly correlated with the signal parameters like total/component masses. We also exclude the chirp parameters (chirp mass $\mathcal{M}$, chirp ellipticity e_M) since they depend on the time-frequency evolution of the signal.

We use white-noise-burst (WNB) signals to train XGBoost, as WNBs allow probing of different regions in the time-frequency map where we expect to find GW signals. Training XGBoost on WNBs ensures the model independence of the ML-enhanced cWB generic burst search as the training set is completely independent of anticipated GW signal models. We generate two sets of WNB for training: (a) WNBs uniformly distributed in central frequency in the range [24,996] Hz, bandwidth [10,300] Hz, and duration is logarithmically distributed between 0.1 ms and 1 ms. (b) WNBs logarithmically distributed in central frequency [24,450] Hz, bandwidth [10,250] Hz and duration [1,50] ms. The choice of the WNB signals is largely motivated by looking at the Q_a-Q_p parameters and ensuring that there are no gaps left in the parameter space where we expect to find GW events. Using WNB signals instead of *ad-hoc* GW waveform signals helps in not biasing the ML algorithm to specific signal models.

We use 50% of the accumulated background for each run and generate the above-mentioned WNB signals for training. Testing is done on various GW signal simulations and the remaining background for O3a and O3b, respectively.

We expect to encounter GW signals with very different waveform morphology than BBH/IMBH. We start by extending the cap for the η_0 statistic to $\eta_0 = 20$ for O3a and 25 for O3b (which was earlier set to 11 for BBH/IMBH searches in the O1-O2 reanalysis paper [91], and to 20 in the O3 reanalysis paper [92]). Since the number of summary statistics used as input features was halved, we change the XGBoost hyper-parameter `max_depth` from 13 to 6, and `min_child_weight` was optimized from 10 to 5. All the other XGBoost hyper-parameters are the same as reported in the O3 reanalysis paper for the BBH/IMBH search [92]. With these changes, short grid searches were performed to determine the optimal value for the user-defined weight options (introduced in Ref. [91]) for O3a ($q = 5$, $A = 40$) and O3b ($q = 6$, $A = 80$). The list of

TABLE II.   Entries for XGBoost hyperparameters.

XGBoost hyper-parameter	Entry
objective	binary:logistic
tree_method	hist
grow_policy	lossguide
n_estimators	20,000[a]
max_depth	6
learning_rate	0.03
min_child_weight	5.0
colsample_bytree	1.0
subsample	0.6
gamma	2.0

[a]n_estimators is optimized using early stopping, where the training stops when the validation score stops improving.

final XGBoost hyper-parameter values can be found in Table II. The same XGBoost hyperparameters are used for both the twofold (HL) and threefold (HLV) networks.

A further correction is applied to the ML penalty factor P_{XGB} to reduce the high SNR background outliers. P_{XGB} is ranking criteria obtained from XGBoost (see Ref. [91]). These high SNR background outliers usually correspond to a low value of Q_a and Q_p (especially blip glitches), and we can suppress them in the affected parameter space by applying the following correction:

$$P'_{\mathrm{XGB}} = \begin{cases} P_{\mathrm{XGB}} - \alpha(0.15 + \Delta Q_{a,p}) & \text{if } \Delta Q_{a,p} \le 0.15 \\ P_{\mathrm{XGB}} & \text{if } \Delta Q_{a,p} > 0.15 \end{cases}, \tag{A1}$$

where α is usually equal to 1, except for HLV network on O3b data where $\alpha = 3$, due to the higher transient noise rate and $\Delta Q_{a,p} = Q_a(Q_p - 0.6)$ defines the penalization curve in the $Q_a - Q_p$ plane. More details can be found in Ref. [91]. The factor W_{XGB} in Eq. (1) is a monotonic function of P'_{XGB} to increase the dynamic range.

[1] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Observation of Gravitational Waves from a Binary Black Hole Merger, Phys. Rev. Lett. **116,** 061102 (2016).

[2] J. Aasi *et al.* (LIGO Scientific Collaboration), Advanced LIGO, Classical Quantum Gravity **32,** 074001 (2015).

[3] F. Acernese *et al.* (Virgo Collaboration), Advanced Virgo: A second-generation interferometric gravitational wave detector, Classical Quantum Gravity **32,** 024001 (2015).

[4] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs, Phys. Rev. X **9,** 031040 (2019).

[5] R. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run, Phys. Rev. X **11,** 021053 (2021).

[6] R. Abbott *et al.* (LIGO Scientific and VIRGO Collaborations), GWTC-2.1: Deep extended catalog of compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run, arXiv:2108.01045.

[7] R. Abbott *et al.* (LIGO Scientific, VIRGO, and KAGRA Collaborations), GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run, arXiv:2111.03606.

[8] R. Abbott *et al.* (KAGRA, VIRGO, and LIGO Scientific Collaborations), All-sky search for short gravitational-wave bursts in the third Advanced LIGO and Advanced Virgo run, Phys. Rev. D **104,** 122004 (2021).

[9] R. Abbott *et al.* (KAGRA, Virgo, and LIGO Scientific Collaboration), All-sky search for long-duration gravitational-wave bursts in the third Advanced LIGO and Advanced Virgo run, Phys. Rev. D **104,** 102001 (2021).

[10] R. P. van der Marel, Intermediate-mass black holes in the universe: A review of formation theories and observational constraints, Coevolution of Black Holes and Galaxies (2004), arXiv:astro-ph/0302101.

[11] M. C. Miller and E. J. M. Colbert, Intermediate-mass black holes, Int. J. Mod. Phys. D **13,** 1 (2004).

[12] T. Ebisuzaki, J. Makino, T. G. Tsuru, Y. Funato, S. F. Portegies Zwart, P. Hut, S. McMillan, S. Matsushita, H. Matsumoto, and R. Kawabe, Missing link found? The runaway path to supermassive black holes, Astrophys. J. **562,** L19 (2001).

[13] F. Koliopanos, Intermediate mass black holes: A review, in *Proceedings of the XII Multifrequency Behaviour of High Energy Cosmic Sources Workshop* (2017), p. 51, arXiv:1801.01095.

[14] K. Inayoshi, E. Visbal, and Z. Haiman, The assembly of the first massive black holes, Annu. Rev. Astron. Astrophys. **58,** 27 (2020).

[15] K. Inayoshi, R. Hirai, T. Kinugawa, and K. Hotokezaka, Formation pathway of Population III coalescing binary black holes through stable mass transfer, Mon. Not. R. Astron. Soc. **468,** 5020 (2017).

[16] P. Marchant, N. Langer, P. Podsiadlowski, T. M. Tauris, and T. J. Moriya, A new route towards merging massive black holes, Astron. Astrophys. **588,** A50 (2016).

[17] S. E. de Mink and I. Mandel, The chemically homogeneous evolutionary channel for binary black hole

mergers: Rates and properties of gravitational-wave events detectable by Advanced LIGO, Mon. Not. R. Astron. Soc. **460,** 3545 (2016).

[18] S. F. Portegies Zwart and S. L. W. McMillan, Black hole mergers in the universe, Astrophys. J. Lett. **528,** L17 (2000).

[19] I. Bartos, B. Kocsis, Z. Haiman, and S. Márka, Rapid and bright stellar-mass binary black hole mergers in active galactic nuclei, Astrophys. J. **835,** 165 (2017).

[20] V. Gayathri, J. Healy, J. Lange, B. O'Brien, M. Szczepańczyk, I. Bartos, M. Campanelli, S. Klimenko, C. O. Lousto, and R. O'Shaughnessy, Eccentricity estimate for black hole mergers with numerical relativity simulations, Nat. Astron. **6,** 344 (2022).

[21] J. Samsing, I. Bartos, D. J. D'Orazio, Z. Haiman, B. Kocsis, N. W. C. Leigh, B. Liu, M. E. Pessah, and H. Tagawa, AGN as potential factories for eccentric black hole mergers, Nature (London) **603,** 237 (2022).

[22] J. C. Bustillo, N. Sanchis-Gual, A. Torres-Forné, and J. A. Font, Confusing Head-On Collisions with Precessing Intermediate-Mass Binary Black Hole Mergers, Phys. Rev. Lett. **126,** 201101 (2021).

[23] M. Ebersold, S. Tiwari, L. Smith, Y.-B. Bae, G. Kang, D. Williams, A. Gopakumar, I. S. Heng, and M. Haney, Observational limits on the rate of radiation-driven binary black hole capture events, Phys. Rev. D **106,** 104014 (2022).

[24] W. E. East, Gravitational waves from the collision of tidally disrupted stars with massive black holes, Astrophys. J. **795,** 135 (2014).

[25] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, Primordial black holes—Perspectives in gravitational wave astronomy, Classical Quantum Gravity **35,** 063001 (2018).

[26] A. M. Green and B. J. Kavanagh, Primordial black holes as a dark matter candidate, J. Phys. G **48,** 043001 (2021).

[27] G. Franciolini, Primordial black holes: From theory to gravitational wave observations, Ph.D. thesis, Geneva University, Department of Theoretical Physics, 2021, arXiv:2110.06815.

[28] M. Ebersold and S. Tiwari, Search for nonlinear memory from subsolar mass compact binary mergers, Phys. Rev. D **101,** 104041 (2020).

[29] L. De Vittori, P. Jetzer, and A. Klein, Gravitational wave energy spectrum of hyperbolic encounters, Phys. Rev. D **86,** 044017 (2012).

[30] G. Cho, A. Gopakumar, M. Haney, and H. M. Lee, Gravitational waves from compact binaries in post-Newtonian accurate hyperbolic orbits, Phys. Rev. D **98,** 024039 (2018).

[31] E. Codazzo, M. Di Giovanni, J. Harms, M. Dall'Amico, and M. Mapelli, Study on the detectability of gravitational radiation from single-binary encounters between black holes in nuclear star cluster: The case of hyperbolic flybys, Phys. Rev. D **107,** 023023 (2023).

[32] B. P. Abbott, R. Abbott, T. D. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Search for post-merger gravitational waves from the remnant of the binary neutron star merger GW170817, Astrophys. J. Lett. **851,** L16 (2017).

[33] O. Benhar, V. Ferrari, and L. Gualtieri, Gravitational wave asteroseismology revisited, Phys. Rev. D **70,** 124015 (2004).

[34] A. Mezzacappa *et al.*, Gravitational-wave signal of a core-collapse supernova explosion of a $15 M_\odot$ star, Phys. Rev. D **102,** 023027 (2020).

[35] J. Powell and B. Müller, Three-dimensional core-collapse supernova simulations of massive and rotating progenitors, Mon. Not. R. Astron. Soc. **494,** 4665 (2020).

[36] D. Radice, V. Morozova, A. Burrows, D. Vartanyan, and H. Nagakura, Characterizing the gravitational wave signal from core-collapse supernovae, Astrophys. J. **876,** L9 (2019).

[37] H. Andresen, E. Müller, H. T. Janka, A. Summa, K. Gill, and M. Zanolin, Gravitational waves from 3D core-collapse supernova models: The impact of moderate progenitor rotation, Mon. Not. R. Astron. Soc. **486,** 2238 (2019).

[38] E. P. O'Connor and S. M. Couch, Exploring fundamentally three-dimensional phenomena in high-fidelity simulations of core-collapse supernovae, Astrophys. J. **865,** 81 (2018).

[39] T. Kuroda, K. Kotake, K. Hayama, and T. Takiwaki, Correlated signatures of gravitational-wave and neutrino emission in three-dimensional general-relativistic core-collapse supernova simulations, Astrophys. J. **851,** 62 (2017).

[40] M. Obergaulinger and M.-A. Aloy, Magnetorotational core collapse of possible GRB progenitors. III. Three-dimensional models, Mon. Not. R. Astron. Soc. **503,** 4942 (2021).

[41] M. Szczepanczyk *et al.*, Detecting and reconstructing gravitational waves from the next galactic core-collapse supernova in the advanced detector era, Phys. Rev. D **104,** 102002 (2021).

[42] Y. Suwa and K. Murase, Probing the central engine of long gamma-ray bursts and hypernovae with gravitational waves and neutrinos, Phys. Rev. D **80,** 123008 (2009).

[43] Q. Cheng, S.-N. Zhang, Y.-W. Yu, and X.-P. Zheng, Probing the physics of newly born magnetars through observation of superluminous supernovae, Phys. Rev. D **97,** 103012 (2018).

[44] A. P. Stephan, S. Naoz, A. M. Ghez, M. R. Morris, A. Ciurlo, T. Do, K. Breivik, S. Coughlin, and C. L. Rodriguez, The fate of binaries in the galactic center: The mundane and the exotic, Astrophys. J. **878,** 58 (2019).

[45] C. L. Fryer, S. E. Woosley, and A. Heger, Pair instability supernovae, gravity waves, and gamma-ray transients, Astrophys. J. **550,** 372 (2001).

[46] J. Powell, B. Müller, and A. Heger, The final core collapse of pulsational pair instability supernovae, Mon. Not. R. Astron. Soc. **503,** 2108 (2021).

[47] N. Rahman, H.-T. Janka, G. Stockinger, and S. Woosley, Pulsational pair-instability supernovae: Gravitational collapse, black hole formation, and beyond, Mon. Not. R. Astron. Soc. **512,** 4503 (2022).

[48] P. Cerdá-Durán, N. DeBrye, M. A. Aloy, J. A. Font, and M. Obergaulinger, Gravitational wave signatures in black hole forming core collapse, Astrophys. J. Lett. **779,** L18 (2013).

[49] T. Kuroda, K. Kotake, T. Takiwaki, and F.-K. Thielemann, A full general relativistic neutrino radiation-hydrodynamics simulation of a collapsing very massive star and the

formation of a black hole, Mon. Not. R. Astron. Soc. Lett. **477,** L80 (2018).

[50] K.-C. Pan, M. Liebendörfer, S. M. Couch, and F.-K. Thielemann, Stellar mass black hole formation and multimessenger signals from three-dimensional rotating core-collapse supernova simulations, Astrophys. J. **914,** 140 (2021).

[51] T. Kuroda, T. Fischer, T. Takiwaki, and K. Kotake, Core-collapse supernova simulations and the formation of neutron stars, hybrid stars, and black holes, Astrophys. J. **924,** 38 (2022).

[52] Y.-F. Wei and T. Liu, Black hole hyperaccretion in collapsars. II. Gravitational waves, Astrophys. J. **889,** 73 (2020).

[53] P. Cerdá-Durán, N. DeBrye, M. A. Aloy, J. A. Font, and M. Obergaulinger, Gravitational wave signatures in black-hole-forming core collapse, Astrophys. J. Lett. **779,** L18 (2013).

[54] B. Haskell and A. Melatos, Models of pulsar glitches, Int. J. Mod. Phys. D **24,** 1530008 (2015).

[55] N. Andersson, G. L. Comer, and R. Prix, The superfluid two-stream instability, Mon. Not. R. Astron. Soc. **354,** 101 (2004).

[56] L. Warszawski and A. Melatos, Gravitational-wave bursts and stochastic background from superfluid vortex avalanches during pulsar glitches, Mon. Not. R. Astron. Soc. **423,** 2058 (2012).

[57] D. Lopez, S. Tiwari, M. Drago, D. Keitel, C. Lazzaro, and G. A. Prodi, Prospects for detecting and localizing short-duration transient gravitational waves from glitching neutron stars without electromagnetic counterparts, Phys. Rev. D **106,** 103037 (2022).

[58] B. Abbott *et al.* (LIGO Scientific Collaboration), Search for Gravitational Wave Bursts from Soft Gamma Repeaters, Phys. Rev. Lett. **101,** 211102 (2008).

[59] T. W. B. Kibble, Topology of cosmic domains and strings, J. Phys. A **9,** 1387 (1976).

[60] T. Damour and A. Vilenkin, Gravitational radiation from cosmic (super)strings: Bursts, stochastic background, and observational windows, Phys. Rev. D **71,** 063510 (2005).

[61] S. Klimenko *et al.*, Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors, Phys. Rev. D **93,** 042004 (2016).

[62] S. Klimenko and G. Mitselmakher, A wavelet method for detection of gravitational wave bursts, Classical Quantum Gravity **21,** S1819 (2004).

[63] S. Klimenko, S. Mohanty, M. Rakhmanov, and G. Mitselmakher, Constraint likelihood analysis for a network of gravitational wave detectors, Phys. Rev. D **72,** 122002 (2005).

[64] M. Drago, S. Klimenko, C. Lazzaro, E. Milotti, G. Mitselmakher, V. Necula, B. O'Brian, G. A. Prodi, F. Salemi, M. Szczepanczyk, S. Tiwari, V. Tiwari, V. Gayathri, G. Vedovato, and I. Yakushin, Coherent Wave-Burst, a pipeline for unmodeled gravitational-wave data analysis, SoftwareX **14,** 100678 (2021).

[65] S. Klimenko, G. Vedovato, V. Necula, F. Salemi, M. Drago, R. Poulton, E. Chassande-Mottin, V. Tiwari, C. Lazzaro, B. O'Brian, M. Szczepanczyk, S. Tiwari, and V. Gayathri, cWB pipeline library: 6.4.1 (2021), 10.5281/zenodo.5798976.

[66] R. Lynch, S. Vitale, R. Essick, E. Katsavounidis, and F. Robinet, Information-theoretic approach to the gravitational-wave burst detection problem, Phys. Rev. D **95,** 104046 (2017).

[67] N. J. Cornish and T. B. Littenberg, BayesWave: Bayesian inference for gravitational wave bursts and instrument glitches, Classical Quantum Gravity **32,** 135012 (2015).

[68] B. P. Abbott *et al.* (KAGRA Collaboration, LIGO Scientific Collaboration, and Virgo Collaboration), Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA, Living Rev. Relativity **23,** 3 (2020).

[69] H. Grote (LIGO Scientific Collaboration), The GEO 600 status, Classical Quantum Gravity **27,** 084003 (2010).

[70] J. Abadie *et al.* (LIGO Scientific and VIRGO Collaborations), All-sky search for gravitational-wave bursts in the second joint LIGO-Virgo run, Phys. Rev. D **85,** 122007 (2012).

[71] J. Abadie *et al.* (LIGO Scientific and Virgo Collaborations), All-sky search for gravitational-wave bursts in the first joint LIGO-GEO-Virgo run, Phys. Rev. D **81,** 102001 (2010).

[72] B. Abbott *et al.* (LIGO Scientific Collaboration), Search for gravitational-wave bursts in the first year of the fifth LIGO science run, Phys. Rev. D **80,** 102001 (2009).

[73] B. Abbott *et al.* (LIGO Scientific Collaboration), Search for gravitational-wave bursts in LIGO's third science run, Classical Quantum Gravity **23,** S29 (2006).

[74] B. Abbott *et al.* (TAMA Collaboration), Upper limits from the LIGO and TAMA detectors on the rate of gravitational-wave bursts, Phys. Rev. D **72,** 122004 (2005).

[75] B. Abbott *et al.* (LIGO Scientific Collaboration), Upper limits on gravitational wave bursts in LIGO's second science run, Phys. Rev. D **72,** 062001 (2005).

[76] B. Abbott *et al.* (LIGO Scientific Collaboration), First upper limits from LIGO on gravitational wave bursts, Phys. Rev. D **69,** 102001 (2004).

[77] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Search for intermediate mass black hole binaries in the first observing run of Advanced LIGO, Phys. Rev. D **96,** 022001 (2017).

[78] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Search for intermediate mass black hole binaries in the first and second observing runs of the Advanced LIGO and Virgo network, Phys. Rev. D **100,** 064064 (2019).

[79] J. Abadie *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Search for gravitational waves from intermediate mass binary black holes, Phys. Rev. D **85,** 102004 (2012).

[80] J. Aasi *et al.*, Search for gravitational radiation from intermediate mass black hole binaries in data from the second LIGO-Virgo joint science run, Phys. Rev. D **89,** 122003 (2014).

[81] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Search for eccentric binary black hole mergers with Advanced LIGO and Advanced Virgo during their first and second observing runs, Astrophys. J. **883,** 149 (2019).

[82] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Observing gravitational-wave transient GW150914

with minimal assumptions, Phys. Rev. D **93**, 122004 (2016); **94**, 069903(A) (2016).

[83] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW190521: A Binary Black Hole Merger with a Total Mass of $150M_\odot$, Phys. Rev. Lett. **125**, 101102 (2020).

[84] M. Szczepańczyk *et al.*, Observing an intermediate-mass black hole GW190521 with minimal assumptions, Phys. Rev. D **103**, 082002 (2021).

[85] R. Abbott *et al.*, Population properties of compact objects from the second LIGO-Virgo gravitational-wave transient catalog, Astrophys. J. Lett. **913**, L7 (2021).

[86] B. O'Brien, M. Szczepanczyk, V. Gayathri, I. Bartos, G. Vedovato, G. Prodi, G. Mitselmakher, and S. Klimenko, Detection of LIGO-Virgo binary black holes in the pair-instability mass gap, Phys. Rev. D **104**, 082003 (2021).

[87] R. Abbott *et al.*, GW190814: Gravitational waves from the coalescence of a 23 solar mass black hole with a 2.6 solar mass compact object, Astrophys. J. Lett. **896**, L44 (2020).

[88] G. Vedovato, E. Milotti, G. Prodi, S. Bini, M. Drago, V. Gayathri, O. Halim, C. Lazzaro, D. Lopez, A. Miani *et al.*, Minimally-modeled search of higher multipole gravitational-wave radiation in compact binary coalescences, Classical Quantum Gravity **39**, 045001 (2022).

[89] S. Vinciguerra, M. Drago, G. A. Prodi, S. Klimenko, C. Lazzaro, V. Necula, F. Salemi, V. Tiwari, M. C. Tringali, and G. Vedovato, Enhancing the significance of gravitational wave bursts through signal classification, Classical Quantum Gravity **34**, 094003 (2017).

[90] M. Cavaglià, S. Gaudio, T. Hansen, K. Staats, M. Szczepańczyk, and M. Zanolin, Improving the background of gravitational-wave searches for core collapse supernovae: A machine learning approach, Mach. Learn. **1**, 015005 (2020).

[91] T. Mishra, B. O'Brien, V. Gayathri, M. Szczepanczyk, S. Bhaumik, I. Bartos, and S. Klimenko, Optimization of model independent gravitational wave search for binary black hole mergers using machine learning, Phys. Rev. D **104**, 023014 (2021).

[92] T. Mishra *et al.*, Search for binary black hole mergers in the third observing run of Advanced LIGO-Virgo using coherent WaveBurst enhanced with machine learning, Phys. Rev. D **105**, 083018 (2022).

[93] D. Lopez, V. Gayathri, A. Pai, I. S. Heng, C. Messenger, and S. K. Gupta, Utilizing Gaussian mixture models in all-sky searches for short-duration gravitational wave bursts, Phys. Rev. D **105**, 063024 (2022).

[94] V. Necula, S. Klimenko, and G. Mitselmakher, Transient analysis with fast Wilson-Daubechies time-frequency transform, J. Phys. Conf. Ser. **363**, 012032 (2012).

[95] T. Chen and C. Guestrin, XGBoost: A scalable tree boosting system, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16 (ACM, New York, NY, USA, 2016), pp. 785–794.

[96] M. Cabero *et al.*, Blip glitches in Advanced LIGO data, Classical Quantum Gravity **36**, 155010 (2019).

[97] H. Andresen, B. Müller, E. Müller, and H.-T. Janka, Gravitational wave signals from 3D neutrino hydrodynamics simulations of core-collapse supernovae, Mon. Not. R. Astron. Soc. **468**, 2032 (2017).

[98] T. Kuroda, K. Kotake, and T. Takiwaki, A new gravitational-wave signature from standing accretion shock instability in supernovae, Astrophys. J. **829**, L14 (2016).

[99] E. Muller, H. T. Janka, and A. Wongwathanarat, Parametrized 3D models of neutrino-driven supernova explosions: Neutrino emission asymmetries and gravitational-wave signals, Astron. Astrophys. **537**, A63 (2012).

[100] J. Powell and B. Müller, Gravitational wave emission from 3D explosion models of core-collapse supernovae with low and normal explosion energies, Mon. Not. R. Astron. Soc. **487**, 1178 (2019).

[101] E. Abdikamalov, S. Gossan, A. M. DeMaio, and C. D. Ott, Measuring the angular momentum distribution in core-collapse supernova progenitors with gravitational waves, Phys. Rev. D **90**, 044001 (2014).

[102] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), Constraints on Cosmic Strings Using Data from the Third Advanced LIGO-Virgo Observing Run, Phys. Rev. Lett. **126**, 241102 (2021).

[103] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Constraints on cosmic strings using data from the first Advanced LIGO observing run, Phys. Rev. D **97**, 102002 (2018).

[104] A. Bohé *et al.*, Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors, Phys. Rev. D **95**, 044028 (2017).

[105] C. Talbot and E. Thrane, Measuring the binary black hole mass spectrum with an astrophysically motivated parameterization, Astrophys. J. **856**, 173 (2018).

[106] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Search for intermediate mass black hole binaries in the first and second observing runs of the Advanced LIGO and Virgo network, Phys. Rev. D **100**, 064064 (2019).

[107] A. H. Mroue *et al.*, Catalog of 174 Binary Black Hole Simulations for Gravitational Wave Astronomy, Phys. Rev. Lett. **111**, 241104 (2013).

[108] J. Healy, C. O. Lousto, Y. Zlochower, and M. Campanelli, The RIT binary black hole simulations catalog, Classical Quantum Gravity **34**, 224001 (2017).

[109] K. Jani, J. Healy, J. A. Clark, L. London, P. Laguna, and D. Shoemaker, Georgia tech catalog of gravitational wave-forms, Classical Quantum Gravity **33**, 204001 (2016).