


Improving significance of binary black hole mergers in Advanced LIGO data using deep learning: Confirmation of GW151216

Shreejit Jadhav^{1,*}, Nikhil Mukund^{2,1,†}, Bhooshan Gadre^{3,1,‡}, Sanjit Mitra^{1,§} and Sheelu Abraham^{1,||}

¹*Inter-University Centre for Astronomy and Astrophysics (IUCAA), Post Bag 4, Ganeshkhind, Pune 411 007, India*

²*Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut) and Institut für Gravitationsphysik, Leibniz Universität Hannover, Callinstr. 38, 30167 Hannover, Germany*

³*Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany*

 (Received 23 October 2020; revised 24 May 2021; accepted 8 July 2021; published 22 September 2021)

We present a novel machine learning (ML) based strategy to search for binary black hole (BBH) mergers in data from ground-based gravitational wave (GW) observatories. This is the first ML-based search that not only recovers all the compact binary coalescences (CBCs) in the first GW transients catalog (GWTC-1), but also makes a clean detection of GW151216, which was not significant enough to be included in the catalog. Moreover, we achieve this by only adding a new coincident ranking statistic (MLStat) to a standard analysis that was used for GWTC-1. In CBC searches, reducing contamination by terrestrial and instrumental transients, which create a loud noise background by triggering numerous false alarms, is crucial to improving the sensitivity for detecting true events. The sheer volume of data and a large number of expected detections also prompts the use of ML techniques. We perform transfer learning to train “InceptionV3,” a pretrained deep neural network, along with curriculum learning to distinguish GW signals from noisy events by analyzing their continuous wavelet transform (CWT) maps. MLStat incorporates information from this ML classifier into the coincident search likelihood used by the standard `pyCBC` search. This leads to at least an order of magnitude improvement in the inverse false-alarm-rate (IFAR) for the previously “low significance” events GW151012, GW170729, and GW151216. We also perform the parameter estimation of GW151216 using `SEOBNRV4HM_ROM`. We carry out an injection study to show that MLStat brings substantial improvement to the detection sensitivity of Advanced LIGO for all compact binary coalescences. The average improvement in the sensitive volume is $\sim 10\%$ for low chirp masses ($0.8\text{--}5 M_{\odot}$) and $\sim 30\%$ for higher masses ($5\text{--}50 M_{\odot}$). Performance in the lower masses may become even better if the training set for the ML classifier, currently restricted to black hole binaries with component masses in the range $2\text{--}98 M_{\odot}$ only, is expanded to include binaries with neutron stars. Considering the impressive ability of the statistic to distinguish signals from glitches, the list of marginal events from MLStat could be quite reliable for astrophysical population studies and further follow-up. This work demonstrates the immense potential and readiness of MLStat for finding new sources in current data and the possibility of its adaptation in similar searches.

DOI: [10.1103/PhysRevD.104.064051](https://doi.org/10.1103/PhysRevD.104.064051)

I. INTRODUCTION

The Advanced LIGO [1,2] and the Advanced Virgo [3] observatories have reported the detection of fifty compact binary coalescences (CBCs) so far [4–10]. They consist of binary black hole (BBH) mergers, including one with asymmetric component masses [11], another one resulting in the production of an intermediate mass black hole

(IMBH) [12], a possible neutron star—black hole (NSBH) merger [13], and a couple of binary neutron star mergers [14,15] in their data spanning the first two observational runs (O1, O2) and the first half of the third observational run (O3a). As the LIGO and Virgo detectors advance their sensitivities and with KAGRA joining the search [16,17], the rate of GW detection is expected to increase multifold [18].

Currently, these detectors make use of both modeled and unmodeled searches to detect potential GW signals in their calibrated data [19–23]. Had the noise in the detectors been Gaussian and stationary, only the time-coincidence of GW signals in more than one detector would be a decisive

*shreejit@iucaa.in

†nikhil.mukund@aei.mpg.de

‡bhooshan.gadre@aei.mpg.de

§sanjit@iucaa.in

||sheelu@iucaa.in

criterion to find a candidate event. The quadrature sum of signal-to-noise ratios (SNRs) in individual detectors would suffice as the ranking statistic [24–27] and modeling the background analytically would be possible in this simplistic case [28,29]. However, the data are heavily contaminated by nonstationary and non-Gaussian transients of instrumental and terrestrial origin, also known as “glitches.” These noise transients mimic the astrophysical signals and result in a persistent problem of false alarms [30,31], which impact the search sensitivity of coalescing compact binaries and GW bursts. As the detectors become more sensitive, along with an improvement in detection rate of GW signals the rate of occurrence of such glitches may also increase, resulting in a considerable number of real GW candidates being reported with reduced significance.

Various methods of compact binary coalescence search [19,21,32] were used to detect gravitational wave sources in LIGO’s first observational run (O1) [33,34]. These searches resulted in two confident BBH detections viz., GW150914 and GW151226, each at a false-alarm rate (FAR) of $<6.0 \times 10^{-7} \text{ yr}^{-1}$ ($> 5.3\sigma$), and a much less confident third detection viz. GW151012. The PyCBC analysis detected GW151012 with a FAR of 0.37 yr^{-1} (1.7σ) [4,35,36]. Later, using an improved ranking statistic, this value could be reduced to 0.17 yr^{-1} [9]. Similarly, in the offline PyCBC analysis of the second observational run (O2), GW170729 was detected with a FAR of 1.36 yr^{-1} .¹ The statistic used in 2-OGC [37] reduced the FAR values of GW151012 and GW170729 to 0.0045 and 0.15, respectively. An investigation shows that the triggers with $\text{SNR} \approx 9$ mainly consist of false alarms caused by glitches [36], which increase the noise background and reduce the significance of marginal events, thereby obstructing the science we can do with them.

In this work, we demonstrate how application of a machine learning algorithm, designed to discern real events from spurious glitches, can be used to improve the standard matched filtering based analysis used by LIGO [9]. We use transfer learning with InceptionV3 [38] to classify CBCs and glitches in the LIGO data. We construct a new coincident ranking statistic (MLStat) by incorporating the output of the ML classifier into the original ranking statistic. Here, we extend the PyCBC search analysis used in GWTC-1 [9,39]; however, a similar methodology can also be implemented with other pipelines like cWB [21], GstLAL [20], and MBTA [22] to seek sensitivity improvements. We show significant improvement in the inverse-false-alarm rates (IFARs) of the low-significance events GW151012 and GW170729 with MLStat. We also confirm the event

GW151216 from O1 [40,41] and report the parameter estimation results for completeness.

In the past few years, machine learning methods have been applied in GW data analysis extensively [42–53]. Though, most of the works have relied on simulated data, there have been efforts in recent years to implement ML methods on real LIGO data [47,50–52]. However, careful considerations are required while working with real data. For example, the results discussed in [47] appear convincing, but the data considered in the work are only limited to 4096 seconds around the three events from O1 and thus, may not prove equally effective on longer durations of real LIGO data which contain a much wider variety of noise artifacts. Also, as discussed in [53], applying deep learning models directly on a continuous stream of whitened data may require additional postprocessing. The performance metrics of the ML models should be evaluated using a data handling scheme that is *a priori* blind to the locations of the injections made. As our analysis is based on triggers obtained after matched-filtering, which localizes the time of merger for true events, this issue is avoided. We are not aware of any ML based search that has been demonstrated to be able to catch all the GWTC-1 events.

The paper is organized as follows. Section II introduces the state-of-the-art deep learning model, viz., InceptionV3, describes its implementation on the LIGO data using transfer learning and gives details about training and validation. Section III describes the construction and operation of MLStat. Results of the injection study used to test MLStat are given in Sec. III A. In Sec. IV, we give details of the analysis of data from first and second observing runs. The Sec. IV A discusses the approximation method introduced by us to gauge the significance improvement of the full foreground of first and second observing runs. Section V concludes the paper.

II. INCORPORATING MACHINE LEARNING

Success for any ML classifier depends on its ability to learn and optimally separate the relevant features from the dataset. Within astronomy community, features are traditionally chosen by a domain expert as in the case of several stellar [54–56], gamma-ray burst (GRB) [57], galaxy [58–60], quasar [61–63], and GW classification schemes [64,65]. The advent of graphics processing units (GPUs) and the availability of larger training sets have resulted in techniques based on deep learning to gain more prominence in recent years. Convolutional neural networks (CNNs), in particular, have shown remarkable success in tasks pertaining to image classification and are consistently outperforming feature-based classical ML techniques when applied to standard bench-marking datasets [66,67]. Large architectures of CNNs, like Google’s Inception model and Microsoft’s Resnet model, have been trained using cutting-edge GPU technology on large datasets containing images of objects in day-to-day life [38,68]. Transfer learning

¹Note that GstLAL detected GW151012 and GW170729 with better significance than PyCBC. In cWB, GW151012 could not be detected, while GW170729 was detected with a significance higher than PyCBC and GstLAL [9,36].

makes use of the rich feature extraction capability of such pretrained models and allows repurposing them for a different classification task. The final few layers of a pretrained network are replaced with trainable new layers while the rest of the network is retained as before. The training of such a network on a new image dataset essentially maps the extracted features to new classes, thus making the training faster while still maintaining the accuracy. Another advantage of transfer learning is that the amount of training data required to reach the prescribed accuracy is hugely reduced.

Inception networks [69] put forward a strategy for making effective deeper networks through tactics different from merely increasing either the layers or neurons per layer. For example, the InceptionV3 network [38] differs from the traditional monolithic CNN architecture through the usage of factorizing convolutions, parallel structures, and extensive dimensionality reduction techniques. These schemes make the architecture well suited for applications with stringent memory and computational constraints but still provide state-of-the-art performance without overfitting the data.

The transients of astrophysical or terrestrial origin seen in LIGO data have a peculiar evolution in the time-frequency domain. This provides a 2D representation that can be analyzed by the image-based classifiers mentioned above. During the early inspiral phase of a CBC waveform, the evolution of frequency with time can be approximated as $f \propto (t_c - t)^{-3/8}$, where t_c is the time of coalescence [70]. This explains the peculiar chirplike shape of CBC signals in time-frequency maps, which can be exploited to differentiate them from glitches. We get much more accurate description of the full time-frequency evolution of CBCs using the inspiral-merger-ringdown (IMR) waveforms. We use transfer learning along with curriculum learning to retrain InceptionV3 for the classification of continuous wavelet transform (CWT) maps of transients in the LIGO data. For each evaluated image, the network outputs a posterior probability across the 17 transient classes used for training. However, only the probability corresponding to CBC class (P_{CBC}) will be used in our analysis. The probability corresponding to other glitch classes may also be incorporated in the future analyses provided the leakage between CBC and other glitch classes is carefully evaluated.

Previously, omega scans have been used to map the LIGO data to the time-frequency domain, where the chirplike evolution of CBCs can be distinguished from other noise transients more effectively [64,65]. We use the continuous wavelet transform (CWT) with an analytic Morlet (amor) wavelet to construct time-frequency scalograms of the whitened strain data to be analyzed. Wavelets, in general, provide a much better time and frequency resolution compared to short-time Fourier transform (STFT) [71]. Discriminator based on relative wavelet

energy has previously been demonstrated to be effective in separating various transient classes and was successfully applied to Advanced LIGO's first observational run data [64]. We create CWT scalograms by whitening and bandpassing the data around a GPS trigger between 16 Hz and 512 Hz. We consider a data slice of 1 second duration with the trigger time kept at center, convert it into a scalogram and save it as a grayscale image with pixels denoting absolute values of CWT coefficients. As InceptionV3 is trained on natural images, the features extracted from different channels are most likely to differ from each other based on the biases in the images of natural objects. The choice of using grayscale colormap ensures the complete glitch morphology is saved in each of the three channels rather than getting divided based on the colormap. The network's convolution filters thus observe the full evolution of a transient in each channel, and the channel-based biases are marginalized.

The CWT image data were generated by using a subset of GravitySpy data [65,72]. The data were manually curated to remove images that lacked clear features or where the features could not be covered in the 1 second long CWT scalograms. The data were then split into training (70%) and validation (30%) sets. As transfer learning involves training a significantly fewer number of neurons, typically, a very small amount of training data (up to a thousand images per class) suffices. This size also makes manual curation practically possible. Due to bandpassing, classes like 1080lines, 1400ripples, and violin mode were rendered redundant and were thus excluded. Also, the power-line class was divided into power line and power line2 to separate shorter and longer transients. Data for CBC class were generated by considering real LIGO strain data and by injecting CBC signals having chirp distances² between 5 Mpc and 300 Mpc with the SEOBNRv4 waveform approximant [74]. The component masses were sampled between ($2 M_{\odot}, 98 M_{\odot}$) with a constraint on total mass, $M \leq 100 M_{\odot}$. Though some of the low mass binaries would merge at frequencies higher than 512 Hz, a major part of these signals contributing most of the power would still be covered by the allowed frequency band. Besides, the upper cutoff of 512 Hz eliminates the relatively louder noise transients at higher frequencies and allows the classifier to focus on the part of signal lying in the most sensitive frequency band of LIGO. A separate class named Gaussian noise was created, which consisted of plain whitened strain data that did not contain any transients. The classes koi fish, no glitch, paired doves and none of the above were omitted from the data due to similarity with at least one of the other classes or lack of feature distinction or coverage in the 1 second long CWT

²Chirp distance accounts for the leading order chirp mass dependence in the amplitude of a CBC signal; see Appendix F in [73].

True Class	AC	BL	EL	GN	HX	CBC	LM	LFB	LFL	PL1	PL2	RBL	SL	SC	TM	WL	WH
AC	93.2	0.0	0.0	0.0	0.0	0.0	2.3	0.0	0.0	0.0	4.5	0.0	0.0	0.0	0.0	0.0	0.0
BL	0.0	87.3	0.0	0.0	0.5	0.0	1.6	0.0	0.0	0.0	0.0	0.0	7.9	0.0	0.0	0.0	2.6
EL	0.0	0.0	98.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.0	0.0
GN	0.0	0.0	0.0	95.3	0.0	4.6	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HX	0.0	1.2	0.0	0.0	98.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CBC	0.0	0.0	0.0	1.9	0.0	98.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LM	0.0	0.7	0.0	0.0	0.0	0.0	97.8	0.7	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0
LFB	0.0	0.0	0.0	0.0	1.6	0.0	1.6	91.9	3.2	0.0	0.0	0.0	0.0	0.0	1.6	0.0	0.0
LFL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.9	91.4	0.0	1.4	0.0	1.4	0.0	0.0	0.0	2.9
PL1	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	98.1	0.9	0.0	0.0	0.0	0.0	0.0	0.0
PL2	5.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.7	91.9	0.0	0.0	0.0	0.0	0.0	0.0
RBL	0.0	2.8	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	94.3	0.0	0.0	0.0	0.0	0.0	1.4
SL	1.8	0.0	0.0	0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	97.3	0.0	0.0	0.0	0.0
SC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.6	0.0	98.8	0.0	0.0
TM	0.6	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	1.3	0.0	0.0	97.4	0.0	0.0	0.0
WL	0.0	0.0	0.0	0.0	0.0	10.0	0.0	0.0	10.0	0.0	0.0	0.0	0.0	0.0	80.0	0.0	0.0
WH	0.0	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.9	0.0	1.0	0.0	0.0	95.2	0.0
	AC	BL	EL	GN	HX	CBC	LM	LFB	LFL	PL1	PL2	RBL	SL	SC	TM	WL	WH

FIG. 1. Confusion matrix (in percent) for the combined validation data from all the levels of curriculum learning for 17 transient classes. The performance of CBC class is of particular importance to MLStat.

maps. A total of 17 transient classes have been used, viz., air compressor (AC), blip (BL), extremely loud (EL), Gaussian noise (GN), helix (HX), injected chirp (CBC), light modulation (LM), low frequency burst (LFB), low frequency lines (LFL), power line (PL1), power line2 (PL2), repeating blips (RBL), scattered light (SL), scratchy (SC), tomte (TM), wandering line (WL) and whistle (WH).

To perform transfer learning, we load the 316 layers deep directed acyclic graph (DAG) network of InceptionV3 and freeze the weights of the first 250 layers. We also add a 40% dropout layer and an additional fully connected layer of size 1024 with leaky-relu activation function before the final replaced fully connected layer which now maps to the 17 transient classes with softmax activation function. We adopt a curriculum learning strategy specifically in CBC class to train the network step-by-step, going from higher SNR values to lower ones. We observed that this method remarkably increased the network accuracy as compared to a single training session with full data. The network was trained using stochastic gradient descent with momentum (SGD-M) [75]. After three levels of training, the trained network achieved a training accuracy of 100% and a validation accuracy of 95.8% (note that as the data is being curated manually, very high accuracies are expected). In the last level of learning, the training and validation accuracies plateau after the 12th epoch and show only statistical fluctuations. We analyzed all the checkpoints of the network recorded at the end of every epoch and chose the best checkpoint based on overall validation accuracy. In Fig. 1, we show the confusion matrix for the entire validation data. A high percentage of predictions lie on the diagonal indicating satisfactory performance of the classifier. Figure 2 shows the receiver operating characteristic (ROC) curve for our network, demonstrating its efficiency of detecting injections against other triggers (mostly originating from noise). Though we do not use the ROC curve to

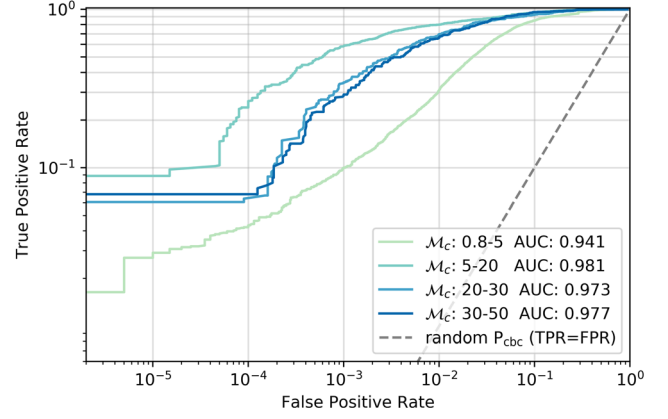


FIG. 2. The receiver operating characteristic curves for different chirp mass bins obtained by comparing simulated injections against randomly chosen noise triggers from Hanford and Livingston data. Log scales are used to visualize the differences better. Masses for the noise triggers are taken from the respective templates triggering them. As some injections lie out of the training range, the performance in the 0.8–5 M_{\odot} chirp mass bin is suboptimal as compared to other bins.

set a threshold on P_{CBC} , the high area under the curve (AUC) scores are indicative of the network’s ability to distinguish signals across all chirp mass bins from glitches.

III. CONSTRUCTION OF MLSTAT

We build our ML tool as an augmentation of one of the standard pipelines used by the LIGO-VIRGO Collaboration, viz. PyCBC. The PyCBC workflow performs matched filtering on the data with a bank of templates. Triggers are then collected by thresholding and clustering the SNR time series. For each trigger, the SNR is reweighted with two types of noise suppressing vetoes [76,77], and a single detector ranking statistic is calculated ensuring approximately constant trigger rate across the search parameter space [39]. The pipeline then finds triggers coincident in Hanford and Livingston detectors with an allowed time window of 15 ms, which accounts for the light travel time and uncertainty in recorded times of coalescence. A semicoherent ranking statistic (we call it base statistic hereafter) is evaluated for these foreground triggers [39]. To construct the background, the same procedure is repeated between the two detectors with time shifted triggers. A false-alarm-rate, which decides the significance of each foreground trigger, is then evaluated based on this background.

We analyze the triggers collected by the PyCBC workflow with our classifier. For each trigger, our classifier gives the probability (P_{CBC}) of it belonging to the CBC class. The likelihood ratio for the detection of a GW signal in the given data from the i^{th} detector is

$$\Lambda(\mathcal{H}_\lambda | s_i) = \frac{p(s_i | \mathcal{H}_\lambda)}{p(s_i | \mathcal{H}_0)}, \quad (1)$$

TABLE I. List of candidate events from the extended analysis of O1 and O2 sorted by FAR_{ml} [78]. We report one to 2 orders of magnitude improvements in IFAR for GW151012, GW170729, and GW151216, obtained by analyzing the respective data chunks which makes them definite confident detections. The FAR_{ml} values quoted for events lying outside these analyzed chunks are inferred by assuming similar significance improvements across respective observational runs (see Sec. IV A). The table is split into two parts to show these events separately. All the catalog events [9] are successfully found by the combined analysis (exception is GW170818 as it was not found by PyCBC). GW170817 has very low P_{CBC} values in both the detectors as it lies outside the training range of the network. Nevertheless, it is still recovered due to its loudness (details in Sec. III). Three marginal events with $\text{FAR}_{\text{ml}} < 1/\text{month}$ are included, which may be worth following-up and may prove to be useful for estimating astrophysical distributions. Mass and spin values (detector frame) are those of the best matching search template giving highest coincident MLStat and need not match the full Bayesian parameter estimation results. FAR_{base} values are also given for comparison. The newly detected event GW151216 is shown in bold.

UTC	MLStat	FAR_{base} [yr^{-1}]	FAR_{ml} [yr^{-1}]	m_1^t [M_{\odot}]	m_2^t [M_{\odot}]	s_{2z}^t	s_{1z}^t	$(\text{P}_{\text{CBC}})_{\text{H}}$	$(\text{P}_{\text{CBC}})_{\text{L}}$
2015-10-12T09:54:43 ^{a, d}	9.0	0.17	4.58×10^{-3}	25.75	17.60	-0.36	0.73	0.889	0.996
2017-07-29T18:56:29 ^{a, d}	8.7	1.36	1.79×10^{-2}	67.52	32.53	0.27	-0.09	0.836	1
2015-12-16T09:24:16^{b, d}	8.3	57.74	0.69	41.78	34.35	0.85	0.98	0.977	0.997
2017-07-28T04:54:39 ^d	7.9	410.49	3.08	33.1	2.50	-0.85	-0.45	0.793	0.990
2015-10-16T13:57:41 ^{c, d}	8.0	...	4.80	364.0	5.35	0.98	-0.24	0.998	0.885
2017-08-17T12:41:04 ^a	28.3	$< 1.3 \times 10^{-5}$	$< 1.3 \times 10^{-5}$	1.46	1.30	-0.02	0.01	1.7×10^{-5}	1.44×10^{-3}
2017-08-14T10:30:43 ^a	12.7	$< 1.3 \times 10^{-5}$	$< 1.3 \times 10^{-5}$	33.14	25.38	0.68	-0.95	1	1
2017-01-04T10:11:58 ^a	10.8	$< 1.4 \times 10^{-5}$	$< 1.4 \times 10^{-5}$	40.87	13.91	-0.70	0.80	0.999	1
2015-09-14T09:50:45 ^a	15.5	$< 1.5 \times 10^{-5}$	$< 1.5 \times 10^{-5}$	44.21	32.16	0.78	-0.86	1	1
2015-12-26T03:38:53 ^a	11.8	$< 1.7 \times 10^{-5}$	$< 1.7 \times 10^{-5}$	14.83	8.50	-0.09	0.81	0.849	1
2017-08-23T13:13:58 ^a	10.2	$< 3.3 \times 10^{-5}$	$< 3.3 \times 10^{-5}$	47.94	16.23	-0.92	0.64	0.999	1
2017-08-09T08:28:21 ^a	10.5	$< 1.45 \times 10^{-4}$	$< 1.45 \times 10^{-4}$	47.62	16.21	-0.57	0.91	0.983	1
2017-06-08T02:01:16 ^a	12.7	$< 3.1 \times 10^{-4}$	$< 3.1 \times 10^{-4}$	16.82	6.10	0.11	0.88	0.999	1
2016-12-14T16:26:40	8.0	847.26	6.02	15.09	11.77	0.79	0.54	0.996	0.973

^aGWTC-1 events;

^bNot present in GWTC-1 but $\text{FAR}_{\text{ml}} < 1 \text{ yr}^{-1}$;

^cNot present in the PyCBC analysis used in GWTC-1;

^dDefinite FAR_{ml} (obtained from analysed chunks)

where, $s_i(t)$ is the time series strain data, \mathcal{H}_λ is the hypothesis stating that a signal $h_i(t, \lambda)$ is present in $s_i(t)$, and \mathcal{H}_0 is the null hypothesis. We update the above likelihood ratio to include P_{CBC} as follows³:

$$\Lambda'(\mathcal{H}_\lambda | s_i) = \frac{p(s_i | \mathcal{H}_\lambda) * (\text{P}_{\text{CBC}})_i}{p(s_i | \mathcal{H}_0)}. \quad (2)$$

Thus, the combined likelihood ratio manifests a new coincident ranking statistic (MLStat) \tilde{q}_{ml} , a simple extension to the coincident base statistic $\tilde{\rho}_c$, as,

$$\tilde{q}_{\text{ml}}^2 = \tilde{\rho}_c^2 + 2 * \log[(\text{P}_{\text{CBC}})_c], \quad (3)$$

where, the combined probability,

$$(\text{P}_{\text{CBC}})_c = (\text{P}_{\text{CBC}})_{\text{H}} * (\text{P}_{\text{CBC}})_{\text{L}}. \quad (4)$$

Here, H and L in the subscripts denote Hanford and Livingston observatories, respectively. Thus, for a detection

³The factor $(\text{P}_{\text{NOISE}})_i$ was not included in the denominator of the likelihood ratio to maintain the constraint $\tilde{q}_{\text{ml}} \leq \tilde{\rho}_c$.

with MLStat, the signal should look CBC-like in both the detectors within the time window required for coincidence.⁴

The astrophysical CBC signals should have P_{CBC} values close to 1, thus making the magnitude of second term on the right-hand side of Eq. (3) very small. Thus, the MLStat values will be very close to the base statistic for such events. On the other hand, noise triggers in the background which do not have a CBC-like composition in CWT maps will be pushed to lower values. Thus, discriminating real events from the noise contamination results in a decreased background, which effectively improves the significance of true CBC events. However, for louder signals, even the low P_{CBC} values in both the detectors will not result in any significant difference between the base statistic and MLStat. This works in our favor while recovering events like GW170817 (shown in Table I), which lie outside our present training range of the network but are recovered with

⁴Note that we do not use P_{CBC} for fitting the single detector event rate as done in the base statistic which may give us further improvements, especially in the regions of the parameter space, which can suffer due to high false-alarm rates and thus show reduced sensitivity.

negligible reduction in the statistic. The ultimate target of MLStat would be recovering the marginal events in the base search.

A. Injection study

To test the performance of MLStat rigorously, an injection study was performed in three chunks of LIGO data containing the events GW151012, GW151216, and GW170729, respectively. In total, 5720 simulated BBH and BNS merger signals, isotropic in sky locations and inclinations, having chirp distances between 5 Mpc to 300 Mpc, were injected in the real LIGO strain. The masses for BBH injections were sampled such that the distribution of total mass was uniform up to $100 M_{\odot}$ while component masses were between $2 M_{\odot}$ and $98 M_{\odot}$. Whereas, for BNS injections the component masses were sampled uniformly in the range $1 M_{\odot}$ to $3 M_{\odot}$. The waveform approximant used for BBH injections was SEOBNRv4 [74], while for BNS it was SpinTaylorT5 [79]. The injections were recovered using search pipelines based on MLStat and the ranking statistic used in PyCBC analysis of GWTC-1 (base statistic here). The IFAR values of the recovered injections were assigned using the background of the respective chunks and statistics. The sensitive spacetime volume $\langle VT \rangle$ estimates were made in different chirp mass bins [9]. Ratios of the $\langle VT \rangle$ estimates with MLStat and base statistic are plotted in Fig. 3 for different chirp mass bins. An overall improvement of 30% in sensitivity can be observed for chirp masses above $5 M_{\odot}$. Though, for chirp masses below $5 M_{\odot}$, the improvement in sensitivity with MLStat was not expected as the training as well as data preprocessing is limited to higher masses, the improvement is evident mainly because of the aggressive reduction in overall background. Thus, even if the low P_{CBC} values result in lower values of MLStat as compared to base statistic, the overall IFAR values for these injections show a small improvement.

IV. ANALYSIS OF O1 AND O2 DATA

We repurpose the offline analysis data of PyCBC search described in GWTC-1 [9,80]. Initially, we analyzed two chunks of data from O1 and O2 that contained the low significance events GW151012 and GW170729, respectively. The analysis consisted of ~ 5.9 days of coincident data during October 8–20, 2015 for O1 and ~ 5.3 days of coincident data starting from July 27 to Aug. 5, 2017 for O2. As described in the next section, after finding the event GW151216 with an improved estimated significance, the corresponding chunk consisting of ~ 5.1 days of coincident data during December 3–18, 2015 was also analyzed.

CWT maps of duration 1 second were created keeping the PyCBC triggers in the center. These images were then analyzed with the ML classifier to get the respective P_{CBC} values. We observed that the classifier is immune to

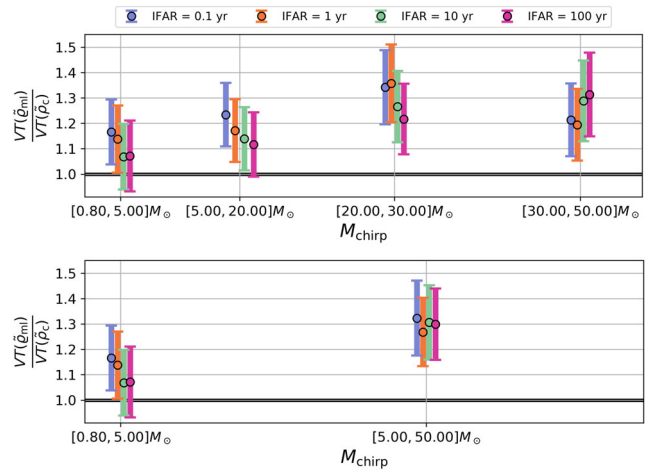


FIG. 3. Ratio of the sensitive volume-time (VT) estimates with MLStat and base statistic for different IFAR thresholds and four chirp mass bins obtained from the injection study is shown in the top panel. The bottom panel combines the higher mass bins to reduce the error bars, showing an average $\sim 10\%$ improvement in sensitive volume for low chirp masses ($< 5 M_{\odot}$) and overall $\sim 30\%$ improvement for higher chirp masses.

changes in CWT maps corresponding to small translations in time, thus allowing us to round off the trigger GPS times to the first decimal place. Analyzing multiple triggers lying within a time window of 0.1 second is thus avoided. The P_{CBC} values were recorded for all the triggers from both the detectors and the coincident analysis was carried out with MLStat. The improvement in significance of GW151012 and GW170729 with MLStat is shown in Fig. 4. The IFAR of GW151012 (GW170729) increases from 5.84 (0.73) years in base statistic to 218.1 (55.8) years in MLStat, thus making them very confident detections.

A. Estimation for full foreground

Though doing a full analysis of any of the observational runs is beyond the scope of the current work, we wish to get an estimate of what we should expect from the extended analysis. We make an assumption that the factor by which the background reduces as a function of base statistic remains similar across all the chunks in an observational run. We then note the improvement in the significance of the full set of foreground triggers against the background of the analyzed chunk of that run and estimate their improved IFAR values with respect to the background of the original chunks they belonged to. Calculating these MLStat IFAR values of full O1 and O2 foregrounds, we report the event GW151216 in O1 with improved significance, which has been discussed in the literature before [37,40,41,81–83]. The IFAR for this event improved from 0.0173 years in GWTC-1 to 1.656 years. When the chunk containing the event was analyzed with MLStat, the actual IFAR value was obtained to be 1.453 years, remarkably close to the one predicted using the approximate projection mentioned

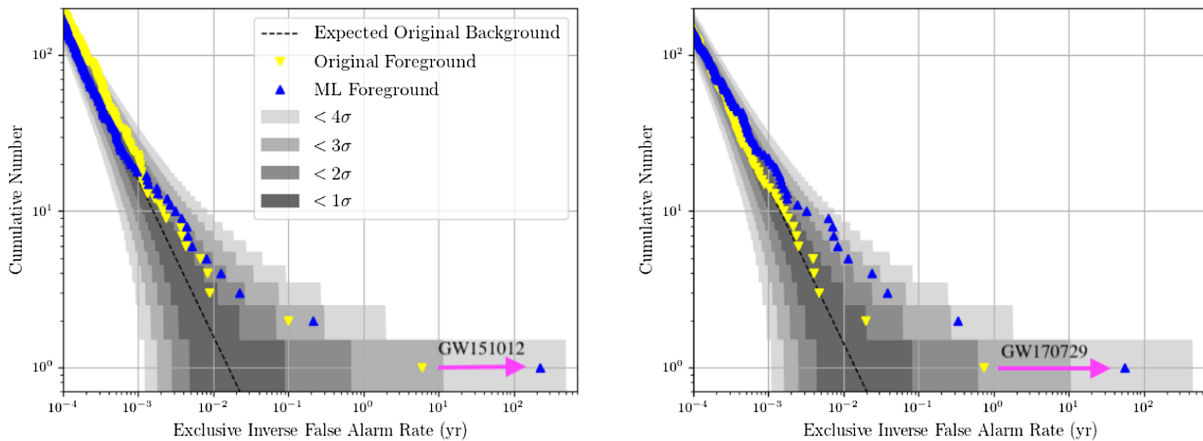


FIG. 4. Significance improvement for GW151012 (left) and GW170729 (right): Cumulative histograms of foreground events in base statistic (yellow) and MLStat (blue) along with the expected background plotted against the inverse false-alarm rate. Shaded regions show the sigma intervals for Poisson uncertainty. The loudness of first ~ 20 foreground events with MLStat in O2 is not a systematic bias and is rather an effect of the low number statistics which has been observed with other statistics before (see Fig. 2 and Fig. 3 in [9]). With MLStat, IFAR value of the most significant foreground event GW151012 (GW170729) improves from 5.84 (0.73) years to 218.1 (55.8) years. We also observed that the event sequence according to IFAR values was considerably shuffled in MLStat. The second most significant event viz., 151016, in O1 chunk has IFAR 0.21 years.

above (see Fig. 5). This confirmed the detection of GW151216 with MLStat while also increasing the reliability of the approximation. The comparison of foreground and background triggers with MLStat and base statistic is shown in Fig. 6. Notice that the background and most of the foreground gets pushed to lower values with MLStat except the event GW151216 (marked with star) which shows minimal change in statistic value due to the high P_{CBC} values recorded in both the detectors. The CWT maps of the events that become significant with our analysis are shown in Fig. 7.

We list the combined foreground of O1 and O2 with $\text{FAR}_{\text{ml}} < 1$ per month in Table I. GW170818 was not detected by P_{CBC} and thus does not show up in our

analysis. The recorded P_{CBC} values for GW170817 are very low in both the detectors as it falls out of the parameter space used for training the classifier and the CBC-tracks for such long duration signals may not be visible in the CWT map. However, as mentioned towards the end of Sec. III, the loudness of this event results in MLStat being only slightly smaller than the base statistic. We intend to build a more comprehensive ML tool better covering the parameter space of CBCs to include neutron stars as well in the follow-up work.

We report the parameter estimation results for GW151216 using the fully Bayesian code BILBY [84,85]. For the analysis, we estimate the noise power spectral density with BAYESWAVE [86] using 16 sec of data around

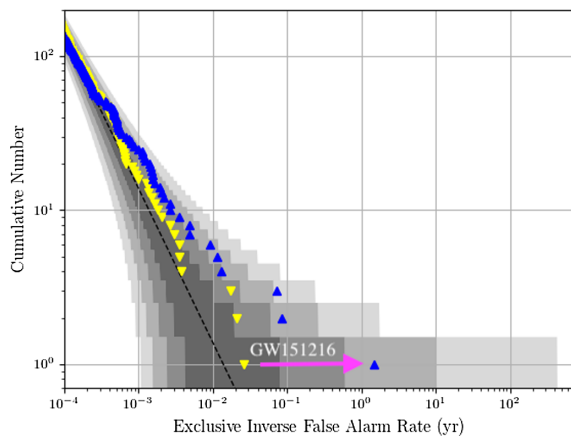


FIG. 5. Significance improvement for GW151216: With MLStat, IFAR value of GW151216 improves from 0.0173 years to 1.453 years making it the most significant foreground event of the chunk.

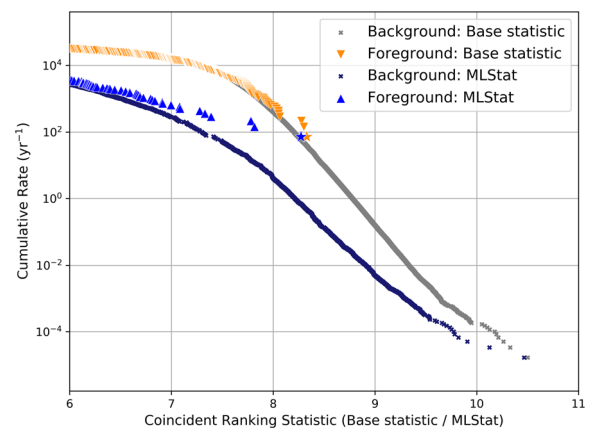


FIG. 6. Comparative search results with MLStat and base statistic for the chunk containing GW151216 (marked by star). Notice the reduced background and a better separated foreground with MLStat.

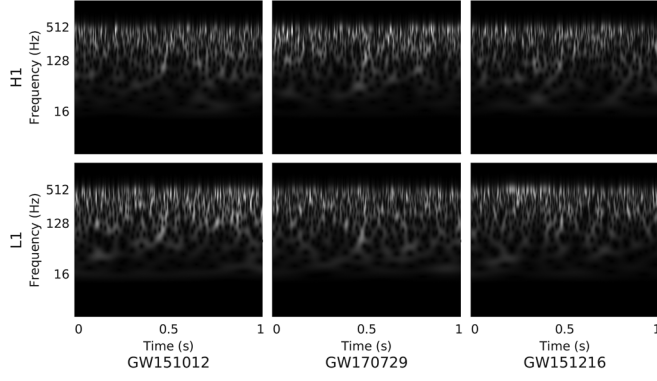


FIG. 7. CWT maps of events that gain significance with MLStat. Whitened strain data of duration one second, bandpassed between 16 Hz and 512 Hz, is used to generate the grayscale images that are fed to the classifier.

GW151216. The posteriors with SEOBNRV4HM_ROM model [87], which includes higher order modes, are shown in the Fig. 8 for mass ratio, chirp mass, luminosity distance, and effective spin. Other detailed analyses for the event which take into account various waveform models and priors can be found in [81–83]. We find that GW151216 has component masses of $46.8_{-17.7}^{+12.5} M_{\odot}$ and $23.6_{-9.6}^{+5.6} M_{\odot}$ with the inverse mass ratio of ~ 2 with the luminosity distance of 2269_{-1033}^{+1193} Mpc. Also, 90% credible intervals for effective

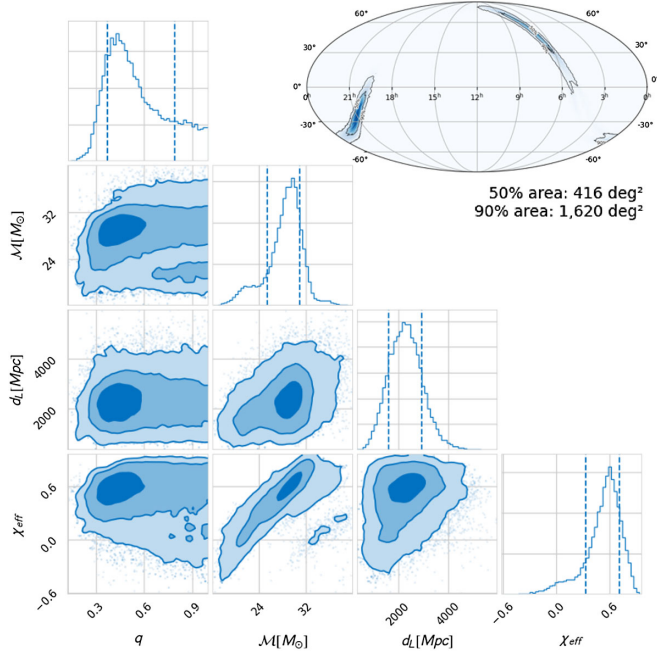


FIG. 8. Bayesian parameter estimation posteriors for the event GW151216 using SEOBNRV4HM_ROM waveform model. We can clearly see the evidence for unequal masses with component masses of $46.8_{-17.7}^{+12.5} M_{\odot}$ and $23.6_{-9.6}^{+5.6} M_{\odot}$ and nonzero spin with χ_{eff} of $0.56_{-0.5}^{+0.23}$ and luminosity distance of 2269_{-1033}^{+1193} Mpc. The corresponding sky localisation map is also overlaid.

inspiral spin parameter show nonzero spins with χ_{eff} of $0.56_{-0.5}^{+0.23}$. The expected posteriors from the parameter estimation also serves as a check that the event is indeed a BBH merger. We also plan to analyse the other marginal events in detail with data quality checks and parameter estimation.

V. CONCLUSIONS

As the sensitivities of ground-based GW interferometers improve, the rate of detection of astrophysical events is going to increase, posing a challenge to the analyses to cope up with the large number of events. The major hurdle in this task arises from the terrestrial and instrumental glitches, and their occurrence may increase with sensitivity. This issue may become even more severe with time as the density of events observed in the distant Universe with statistically lower SNRs is expected to increase. Improving the fraction of true events in the set of potential triggers can significantly reduce this burden. Perhaps in the future, ML will also help us in expanding the dimension and volume of the parameter space for astrophysical searches.

In this work, we demonstrated the capability of machine learning to improve the significance of CBC signals and to discard false triggers by integrating it with the existing analysis framework of PyCBC. We used transfer learning with InceptionV3, a pretrained image based classifier, for effective identification of binary black hole mergers against glitches in LIGO data. We repurpose the PyCBC offline search data to reanalyze the matched filter triggers for two chunks of data from O1 and O2 that contained the events GW151012 and GW170729. We use the retrained InceptionV3 network to classify the continuous wavelet transform maps, a representation of time series data in time-frequency domain, of these triggers and get the P_{CBC} values which are used to construct a new ranking statistic MLStat—a simple extension of the statistic used by PyCBC in the first GW transient catalog (GWTC-1). This helped in breaking the degeneracy between the real CBC signals and the noise transients that result in an increased background. For validating the performance of MLStat against the base statistic, we performed a rigorous injection study with simulated BNS and BBH signals which showed on the average $\sim 10\%$ increase in sensitive volume for chirp masses between $0.8\text{--}5 M_{\odot}$ and $\sim 30\%$ for chirp masses higher than $5 M_{\odot}$. With the present training range being restricted to and data preprocessing favourable for the stellar mass BBHs ($m_1, m_2 \in 2\text{--}98 M_{\odot}$; $M \leq 100 M_{\odot}$), the ML network is not expected to identify signals in the low chirp masses. However, due to aggressive reduction in overall background with MLStat, the detection sensitivity for low chirp mass signals is still enhanced. The sensitivity across the parameter space should further improve once we incorporate binaries involving neutron stars and intermediate mass BBHs in our training set. The preclassification

information may also be fed to the ML model to make its performance more robust.

With MLStat, we achieve a considerable reduction in background and an improved separation of foreground. We report 1 to 2 orders of magnitude reduction in false alarm rates for the low significance events GW151012 and GW170729 with MLStat as compared to the values obtained by PyCBC in GWTC-1. This is also the first time a machine learning based search algorithm was able to detect all the CBCs in GWTC-1 with same or better significance. We also confirm the detection of the event GW151216, which was not included in GWTC-1 lists of confirmed and marginal events due to lack of significance. While the existence and nature of this event is debated [37,41,81–83], for demonstration of completeness of the process of obtaining a new detection using MLStat, we perform parameter estimation for this event with an improved waveform model with higher order modes SEOBNRV4HM_ROM [87]. It is worth noting that the posteriors of parameters and the sky localization contours do not suggest any obvious presence of a glitch and thus, strengthen the possibility of GW151216 being of astrophysical origin. Considering the ability of our method to distinguish the false triggers from astrophysical events, the list of subthreshold events reported with MLStat may be more reliable for astrophysics (e.g., population studies) but further follow-up through a data quality check would still be required.

That, by tuning a generic ML algorithm and introducing a simple extension to the ranking statistic, we could achieve these significant improvements, shows the enormous

potential in ML, provided we can adapt it with fine understanding of the problem in hand. There is ample scope to improve our present analysis and, in general, several avenues may be explored to introduce machine learning based algorithms in GW analyses.

ACKNOWLEDGMENTS

Authors express thanks to Ninan Sajeeth Philip, Shivaraj Kandhasamy, and the LIGO-Virgo-KAGRA Collaboration for their valuable comments and suggestions. Some of the results in this work have been derived using the PESUMMARY package [88]. S. J. and N. M. acknowledge support of Council for Scientific and Industrial Research (CSIR), India. NM expresses thanks to the Max Planck Society and the Leibnitz University, Hannover. B. G. acknowledges support of the Max Planck Society and the University Grants Commission (UGC), India. We acknowledge the use of GWA and the LDG clusters at IUCAA (Sarathi) and Caltech for the computational work. The follow-up analysis for GW151216 is performed with the HPC clusters HYPATIA at the Max Planck Institute for Gravitational Physics, Potsdam-Golm. S. M. and S. A. acknowledge support from the Department of Science and Technology (DST), India, provided under the Swarna Jayanti Fellowships scheme. This document has been assigned The Inter-University Centre for Astronomy and Astrophysics (IUCAA) Preprint No. IUCAA-03/2020 and LIGO Document No. LIGO-P2000399.

-
- [1] J. Aasi *et al.* (LIGO Scientific and Virgo Collaborations), *Classical Quant. Grav.* **32**, 074001 (2015).
 - [2] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 131103 (2016).
 - [3] F. Acernese *et al.*, *Classical Quant. Grav.* **32**, 024001 (2015).
 - [4] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 061102 (2016).
 - [5] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 241103 (2016).
 - [6] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **118**, 221101 (2017).
 - [7] B. P. Abbott *et al.*, *Astrophys. J.* **851**, L35 (2017).
 - [8] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **119**, 141101 (2017).
 - [9] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **9**, 031040 (2019).
 - [10] R. Abbott *et al.*, *Phys. Rev. X* **11**, 021053 (2021).
 - [11] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **102**, 043015 (2020).
 - [12] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **125**, 101102 (2020).
 - [13] R. Abbott *et al.*, *Astrophys. J.* **896**, L44 (2020).
 - [14] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **119**, 161101 (2017).
 - [15] B. P. Abbott *et al.*, *Astrophys. J.* **892**, L3 (2020).
 - [16] Y. Aso, Y. Michimura, K. Somiya, M. Ando, O. Miyakawa, T. Sekiguchi, D. Tatsumi, and H. Yamamoto (The KAGRA Collaboration), *Phys. Rev. D* **88**, 043007 (2013).
 - [17] T. Akutsu, M. Ando, K. Arai, Y. Arai, S. Araki, A. Araya *et al.*, [arXiv:2005.05574](https://arxiv.org/abs/2005.05574).
 - [18] B. P. Abbott *et al.*, *Living Rev. Relativity* **21**, 3 (2018).
 - [19] S. A. Usman, A. H. Nitz, I. W. Harry, C. M. Biwer, D. A. Brown *et al.*, *Classical Quant. Grav.* **33**, 215004 (2016).
 - [20] S. Sachdev, S. Caudill, H. Fong, R. K. L. Lo, C. Messick *et al.*, [arXiv:1901.08580](https://arxiv.org/abs/1901.08580).
 - [21] S. Klimentko, G. Vedovato, M. Drago, F. Salemi, V. Tiwari *et al.*, *Phys. Rev. D* **93**, 042004 (2016).
 - [22] T. Adams (LIGO Scientific and Virgo Collaborations), [arXiv:1507.01787](https://arxiv.org/abs/1507.01787).
 - [23] Q. Chu *et al.*, [arXiv:2011.06787](https://arxiv.org/abs/2011.06787).
 - [24] L. Zubakov and V. Wainstein, (1962).
 - [25] C. Cutler, T. A. Apostolatos, L. Bildsten, L. S. Finn, E. E. Flanagan *et al.*, *Phys. Rev. Lett.* **70**, 2984 (1993).

- [26] A. Pai, S. Dhurandhar, and S. Bose, *Phys. Rev. D* **64**, 042004 (2001).
- [27] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, *Phys. Rev. D* **85**, 122006 (2012).
- [28] L. S. Finn, *Phys. Rev. D* **46**, 5236 (1992).
- [29] L. S. Finn and D. F. Chernoff, *Phys. Rev. D* **47**, 2198 (1993).
- [30] B. P. Abbott *et al.*, *Classical Quant. Grav.* **33**, 134001 (2016).
- [31] L. K. Nuttall *et al.*, *Classical Quant. Grav.* **32**, 245005 (2015).
- [32] C. Messick *et al.*, *Phys. Rev. D* **95**, 042001 (2017).
- [33] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **6**, 041015 (2016).
- [34] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **96**, 022001 (2017).
- [35] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **93**, 122003 (2016).
- [36] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **6**, 041015 (2016).
- [37] A. H. Nitz, T. Dent, G. S. Davies, S. Kumar, C. D. Capano, I. Harry, S. Mozzon, L. Nuttall, A. Lundgren, and M. Tápai, *Astrophys. J.* **891**, 123 (2020).
- [38] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2016), pp. 2818–2826.
- [39] A. H. Nitz, T. Dent, T. D. Canton, S. Fairhurst, and D. A. Brown, *Astrophys. J.* **849**, 118 (2017).
- [40] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, *Phys. Rev. D* **100**, 023011 (2019).
- [41] B. Zackay, T. Venumadhav, L. Dai, J. Roulet, and M. Zaldarriaga, *Phys. Rev. D* **100**, 023007 (2019).
- [42] H. Gabbard, M. Williams, F. Hayes, and C. Messenger, *Phys. Rev. Lett.* **120**, 141103 (2018).
- [43] M. B. Schäfer, F. Ohme, and A. H. Nitz, *Phys. Rev. D* **102**, 063015 (2020).
- [44] E. Cuoco, J. Powell, M. Cavaglià, K. Ackley, M. Bejger *et al.*, *Mach. Learn. Sci. Technol.* **2**, 011002 (2021).
- [45] S. R. Green, C. Simpson, and J. Gair, *Phys. Rev. D* **102**, 104057 (2020).
- [46] D. George and E. A. Huerta, *Phys. Rev. D* **97**, 044039 (2018).
- [47] D. George and E. A. Huerta, *Phys. Lett. B* **778**, 64 (2018).
- [48] M. L. Chan, I. S. Heng, and C. Messenger, *Phys. Rev. D* **102**, 043022 (2020).
- [49] S. Coughlin *et al.*, *Phys. Rev. D* **99**, 082002 (2019).
- [50] T. Marianer, D. Poznanski, and J. X. Prochaska, *Mon. Not. R. Astron. Soc.* **500**, 5408 (2020).
- [51] K. Kim, T. G. F. Li, R. K. L. Lo, S. Sachdev, and R. S. H. Yuen, *Phys. Rev. D* **101**, 083006 (2020).
- [52] A. L. Miller, P. Astone, S. D’Antonio, S. Frasca, G. Intini, I. La Rosa, P. Leaci, S. Mastrogiovanni, F. Muciaccia, A. Mitidis, C. Palomba, O. J. Piccinni, A. Singhal, B. F. Whiting, and L. Rei, *Phys. Rev. D* **100**, 062005 (2019).
- [53] T. D. Gebhard, N. Kilbertus, I. Harry, and B. Schölkopf, *Phys. Rev. D* **100**, 063015 (2019).
- [54] M. Manteiga, I. Carricajo, A. Rodríguez, C. Dafonte, and B. Arcay, *Astron. J.* **137**, 3245 (2009).
- [55] J. S. Almeida and C. A. Prieto, *Astrophys. J.* **763**, 50 (2013).
- [56] T. Kuntzer, M. Tewes, and F. Courbin, *Astron. Astrophys.* **591**, A54 (2016).
- [57] S. Abraham, N. Mukund, A. Vibhute, V. Sharma, S. Iyyani, D. Bhattacharya, A. R. Rao, S. Vadawale, and V. Bhalerao, *Mon. Not. R. Astron. Soc.* **504**, 3084 (2021).
- [58] N. S. Philip, Y. Wadadekar, A. Kembhavi, and K. B. Joseph, *Astron. Astrophys.* **385**, 1119 (2002).
- [59] S. C. Odewahn, S. H. Cohen, R. A. Windhorst, and N. S. Philip, *Astrophys. J.* **568**, 539 (2002).
- [60] X. Li, Y. Lin, and K. Qiu, *Res. Astron. Astrophys.* **19**, 8 (2019).
- [61] G. T. Richards, A. D. Myers, A. G. Gray, R. N. Riegel, R. C. Nichol, R. J. Brunner, A. S. Szalay, D. P. Schneider, and S. F. Anderson, *Astrophys. J. Suppl. Ser.* **180**, 67 (2009).
- [62] S. Abraham, N. S. Philip, A. Kembhavi, Y. G. Wadadekar, and R. Sinha, *Mon. Not. R. Astron. Soc.* **419**, 80 (2012).
- [63] C. M. Peters, G. T. Richards, A. D. Myers, M. A. Strauss, K. B. Schmidt, Ž. Ivezić, N. P. Ross, C. L. MacLeod, and R. Riegel, *Astrophys. J.* **811**, 95 (2015).
- [64] N. Mukund, S. Abraham, S. Kandhasamy, S. Mitra, and N. S. Philip, *Phys. Rev. D* **95**, 104059 (2017).
- [65] M. Zevin, S. Coughlin, S. Bahaadini, E. Besler, N. Rohani *et al.*, *Classical Quant. Grav.* **34**, 064003 (2017).
- [66] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, in *2009 IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2009), pp. 248–255.
- [67] A. Krizhevsky, I. Sutskever, and G. E. Hinton, in *Advances in Neural Information Processing Systems* (2012), pp. 1097–1105.
- [68] K. He, X. Zhang, S. Ren, and J. Sun, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778.
- [69] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed *et al.*, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015), pp. 1–9.
- [70] J. D. Creighton and W. G. Anderson, *Gravitational-Wave Physics and Astronomy: An Introduction to Theory, Experiment and Data Analysis* (John Wiley & Sons, New York, 2012).
- [71] L. Debnath and F. A. Shah, *Wavelet Transforms and Their Applications* (Springer, New York, 2002).
- [72] S. Bahaadini, V. Noroozi, N. Rohani, S. Coughlin, M. Zevin, J. Smith, V. Kalogera, and A. Katsaggelos, *Inf. Sci.* **444**, 172 (2018).
- [73] B. P. Abbott *et al.* (LIGO Scientific Collaboration), *Phys. Rev. D* **79**, 122001 (2009).
- [74] A. Bohé *et al.*, *Phys. Rev. D* **95**, 044028 (2017).
- [75] N. Qian, *Neural Netw.* **12**, 145 (1999).
- [76] B. Allen, *Phys. Rev. D* **71**, 062001 (2005).
- [77] A. H. Nitz, *Classical Quant. Grav.* **35**, 035016 (2018).
- [78] MLStat search data release, <https://github.com/shreejit-92/MLStat-1.git>.
- [79] P. Ajith, *Phys. Rev. D* **84**, 084037 (2011).
- [80] T. D. Canton and I. W. Harry, arXiv:1705.01845.
- [81] G. Ashton and E. Thrane, *Mon. Not. R. Astron. Soc.* **498**, 1905 (2020).
- [82] G. Pratten and A. Vecchio, arXiv:2008.00509.

- [83] Y. Huang, C.-J. Haster, S. Vitale, A. Zimmerman, J. Roulet, T. Venumadhav, B. Zackay, L. Dai, and M. Zaldarriaga, *Phys. Rev. D* **102**, 103024 (2020).
- [84] G. Ashton *et al.*, *Astrophys. J. Suppl. Ser.* **241**, 27 (2019).
- [85] I. M. Romero-Shaw, C. Talbot, S. Biscoveanu, V. D’Emilio, G. Ashton *et al.*, *Mon. Not. R. Astron. Soc.* (2020).
- [86] B. Bécsy, P. Raffai, J. N. Cornish, R. Essick, J. Kanner *et al.*, *Astrophys. J.* **839**, 15 (2017).
- [87] R. Cotesta, S. Marsat, and M. Pürrer, *Phys. Rev. D* **101**, 124040 (2020).
- [88] C. Hoy and V. Raymond, [arXiv:2006.06639](https://arxiv.org/abs/2006.06639).