

New method to observe gravitational waves emitted by core collapse supernovae

P. Astone,¹ P. Cerdá-Durán,² I. Di Palma,^{1,3} M. Drago,⁴ F. Muciaccia,^{1,3} C. Palomba,¹ and F. Ricci^{1,3}

¹*INFN, Sezione di Roma, I-00185 Roma, Italy*

²*Departamento de Astronomía y Astrofísica, Universitat de València, 46100 Burjassot (Valencia), Spain*

³*Università di Roma La Sapienza, I-00185 Roma, Italy*

⁴*INFN LNGS and Gran Sasso Science Institute, 67100 L'Aquila, Italy*



(Received 30 October 2018; published 10 December 2018)

While gravitational waves have been detected from mergers of binary black holes and binary neutron stars, signals from core collapse supernovae, the most energetic explosions in the modern Universe, have not been detected yet. Here we present a new method to analyse the data of the LIGO, Virgo, and KAGRA network to enhance the detection efficiency of this category of signals. The method takes advantage of a peculiarity of the gravitational wave signal emitted in the core collapse supernova and it is based on a classification procedure of the time-frequency images of the network data performed by a convolutional neural network trained to perform the task to recognize the signal. We validate the method using phenomenological waveforms injected in Gaussian noise whose spectral properties are those of the LIGO and Virgo advanced detectors and we conclude that this method can identify the signal better than the present algorithm devoted to select gravitational wave transient signal.

DOI: [10.1103/PhysRevD.98.122002](https://doi.org/10.1103/PhysRevD.98.122002)

I. INTRODUCTION

The direct observation of gravitational waves (GWs) by the advanced kilometer-scale GW detectors, which have been operative in the period between 2015 and 2017, is a major milestone in physics and astrophysics [1–7]. So far, all the observed GW signals have been produced at the merger of compact binary systems. All but one correspond to black hole binaries with total mass in the range of tens of solar masses. The observation of the binary neutron star merger in 2017 [5] is a crucial milestone of the multi-messenger astronomy because of the combined detection of GW and electromagnetic observations [6,7].

In the future we expect that the LIGO, Virgo, and KAGRA interferometers will observe other astrophysical phenomena. The collapse of the core of massive stars ($\sim 10\text{--}100 M_{\odot}$), in particular those producing core-collapse supernovae (CCSNe), was considered a potential source of detectable GWs already at the epoch of resonant bar detectors. GW are emitted by aspherical mass-energy dynamics that include quadrupole or higher-order gravitational contributions. If this asymmetric dynamics is present in the preexplosion stalled-shock phase of CCSNe, we should have the chance to observe the violent death of massive stars also via the gravitational channel. GW bursts from CCSNe encode information on the core dynamics of a dying massive star and may enlighten the mechanism driving supernovae.

Early analytic and semianalytic estimates of the GW signature of stellar collapse and CCSNe gave optimistic signal strengths ($\sim 10^{-2} M_{\odot} c^2$), while modern multidimensional

simulations predict emission frequencies in the band of ground-based laser interferometers (10 Hz–10 kHz) with total emitted GW energies in the range $10^{-12}\text{--}10^{-8} M_{\odot} c^2$. These predictions suggest that even advanced interferometers will only be able to detect GWs from CCSNe at distances lower than 1–100 kpc with an optimistic rate event of the order of 1/25 yr. In other models of more extreme scenarios, involving nonaxisymmetric rotational instabilities, centrifugal fragmentation and accretion disk, the emitted GW signals may be sufficiently strong to be detectable to distances of (10–15) Mpc. At this distance the Virgo cluster is included in the sphere centered on Earth and, as consequence, the potential detection rate of the advanced detectors increases up to values higher than 1 yr.

A credible CCSNe scenario is based on a collapse of the star's iron core (see, e.g., [8,9], for recent reviews), which results in the formation of a protoneutron star (PNS) and an expanding hydrodynamic shock wave. The shock gets immediately stalled by presence of a continuous accreting flow. On a timescale of $\sim 0.2\text{--}1$ s, a yet-uncertain supernova mechanism revives the shock that reaches the stellar surface and produces the spectacular electromagnetic emission of a type-II or type-Ib/c supernova. If the shock fails to revive, a black hole is formed with no or very weak electromagnetic signature associated. In this work we refer generically as CCSNe to any core collapse event, regardless if the final outcome is a supernovae or the formation of a black hole.

The supernova type classification is based on the explosion light curve and spectrum, which depend largely

on the nature of the progenitor star. The time from core collapse to breakout of the shock through the stellar surface and first supernova light is minutes to days, depending on the radius of the progenitor and energy of the explosion.

Any core-collapse event generates a burst of neutrinos that releases most of the protoneutron star's gravitational binding energy ($\sim 10^{53}$ erg $\sim 0.15 M_{\odot} c^2$) on a timescale of the order of 10 s. This neutrino burst was detected from SN 1987A and confirmed the basic theory of CCSNe [10].

Multidimensional simulations of core-collapse supernovae are currently at the frontier of research in the field following the two main basic explosion paradigms: the *neutrino-driven* mechanism, thought to be active for slowly rotating progenitors and responsible for the most common SNe, and the *magneto-rotational mechanism*, active only for fast rotating-progenitors and responsible for rare but highly energetic events, like hypernovae and long GRBs.

Several groups worldwide are currently attacking this problem with two- and three-dimensional simulations using the world's most powerful supercomputers. Multiple challenges arise during the numerical modeling: (i) accurate solution of the neutrino transport equations during the evolution; (ii) incorporation of the complete interactions of electron, muon and tau neutrinos and their anti-particles with matter; (iii) use of high resolution to resolve numerically fine structure features in the convective and turbulent flow around the protoneutron star; this is of special importance for the development of magneto-rotational instabilities in fast-rotating progenitors; (iv) accurate (general relativistic) description of gravity; (v) use of sophisticated equations of state to describe the behavior of matter at high densities. The different groups studying the problem use different approaches to tackle each of these challenges and, to this point, no one has carried out a definitive three-dimensional simulation including all the physical ingredients and with sufficiently high resolution to give the worldwide community confidence in the results.

Despite the problem complexity, these calculations give acceptable remnant neutron-star masses and predicted already few distinct signatures of GW signals in both the time and frequency domains. The core-bounce signal is the part of the waveform which is best understood [11], and it can be directly related to the rotational properties of the core [12–14]. However, fast-rotating progenitors are not common and their bounce signal will be probably difficult to observe in typical galactic events, due to its high frequency and low amplitude.

In addition, during the post-bounce evolution of the newly formed proto-neutron star (PNS), the convection determines the excitation of highly damped modes in the PNS by accreting material and instabilities (SASI), with a peculiar GW emission. In this case the GW waveforms last for about 200–500 ms until the supernova explodes (see e.g., [15]) or, in the case of black hole formation, the typical duration is 1 s or above [16]. The peculiarity is that the

signal frequencies raise monotonically with time due to the contraction of the PNS, whose mass steadily increases. Characteristic frequencies of the PNS can be as low as ~ 100 Hz, specially those related to g-modes [15–20] and SASI [16,19,20], which make them a target for ground-based interferometers with the highest sensitivity at those frequencies.

These information can be used in the search of the GW signal embedded in the detector noise, with the perspective to increase the confidence detection of signal emitted in the deeper universe. In the case of the search of GW binary systems the dominant analysis technique is the matched filter, based on models computed in the general relativity (GR) framework. Then, the posterior probability distribution for the signal parameters are estimated from the noisy detector data using probabilistic Bayesian methods. These techniques can be used in the case of a detailed prediction of the waveform, a case different to the present one. For this reason the approach used in the past for CCSNe relies on fully unbiased algorithms, which do not require any assumptions about the GW morphology. In general, these algorithms assign a loudness measure to each event, whose significance is evaluated by computing the rate at which the background noise produces events of equal or higher loudness (false alarm rate, FAR).

Currently, in the CCNSe case we can take advantage of the signal peculiarity, in particular that associated to monotonically raise of the frequency related to the g-mode excitation. The aim of this paper is to present a search strategy of events in coincidence in the advanced detector network, characterized by a raising monotonic behavior in the time-frequency plane, similar to the one observed in numerical simulations.

This strategy is based on machine learning techniques. These are tools applied even to big chunks of data in different contexts, analyzed with minimal human supervision, and able to resolve ambiguity and tolerate unpredictability. In this framework pattern recognition, seen as practical outcome of the machine learning technique which divides data into classes, is a data analysis approach widely used for recognizing regularities in images. This approach has been tested already on GW data in particular for the real time detection and the parameter estimation of binary black hole mergers [21–24]. Here we present a method helpful for the search of signals associated to the supernovae explosions. In the following sections, after the discussion of the scientific problem of the detection of the transient signal due to the supernovae explosion, we will describe the phenomenological waveforms generated to simulate the CCNSe GW signal, the architecture of the convolutional neural network and the whole method developed to recognize the signal. Finally, to validate the method, we present the results obtained by injecting waveforms in Gaussian noise with the spectral behavior of the LIGO and Virgo Advanced detectors.

II. PHENOMENOLOGICAL WAVEFORMS FOR CCSNE

The first step of a signal search based on machine learning technique is to provide a training data set, where the GW signals are present. It follows that we have to produce GW templates representative of CCSNe covering the parameter space of possible core collapse events. At present the outcome of multidimensional numerical simulations is a limited set of GW waveforms because of the massive amount of computational resources needed to produce each of them. In addition the progenitors models used in these simulations can be biased: most of them are developed with the aim to compare the model prediction with the observations of the supernovae SN1987A or they are focused to the case of fast-rotating progenitors, a small fraction of the total number of observable events. Furthermore, due to the numerical challenge of these simulations and the various approximations used, it is unclear how close the existing numerical templates are to the actual GW signal for a specific type of progenitor. Therefore, the existing numerical templates seems to have just a partial coverage of the CCSNe parameter space.

For this reason, to validate our search method, we use a more flexible approach. We have developed a parametrized phenomenological waveform designed to match the most common features observed in the numerical models of CCSNe and we devote the next section to present the simulation used to produce the template bank, which covers a wider parameter space.

A. Reference numerical simulations

We base our phenomenological templates on the numerical simulations by [15,16,18–20,25]. In all these works, the authors present the gravitational waves signals extracted from core-collapse simulations, plot spectrograms of the signal, and interpret those spectrograms in terms of excitation of modes (g-modes and SASI modes) and convection. A more detailed analysis and interpretation of the waveforms in terms of eigenmodes of the proto-neutron star (PNS) has been carried out by [26–30]. These simulations were performed in two and three dimensions (2D/3D), using either a modified Newtonian potential or general relativity (XCFC approximation) and a neutrino treatment with different degrees of sophistication (from a simple leakage to Ray-by-Ray+ transport). The progenitors used are nonrotating stars (except for [16]) with solar metallicity (except [16] and some models of [15]) and correspond to zero-age main-sequence masses in the range 8–40 M_{\odot} . This kind of progenitor is most likely to form type II supernovae and in some cases a failed supernova (unovae, in which a black hole is formed). We focus in this work exclusively in this kind of progenitors. A galactic supernova (or an unnova) is very likely to have such a nonrotating progenitor, so the features presented in these

work are the most relevant for a possible detection. We note that the fraction of CCSNe associated to the collapse of rapidly rotating core is probably below 1% (see discussion in [31]).

Waveforms from the collapse of nonrotating progenitors have the next features identified by several of authors:

- (1) *Bounce signal*: in practice almost nonexistent. Only fast rotating models give a strong signal at bounce [12–14].
- (2) *Prompt convection*: Some models show prompt convection right after bounce, which lasts for 50–100 ms at about 100Hz (see, e.g., [18,32]). The amplitude of this signal is currently under debate and it may depend on fine details of the numerical simulations and on the equation of state.
- (3) *Excitation of g-modes of the PNS*: basically all simulations in the literature show this feature. Its frequency starts around 100 Hz and grows in time as the mass of the PNS grows creating a characteristic raising arch in the spectrogram. It may start right after the bounce or with some delay (up to $\sim 200ms$). The signal last until the onset of the explosion or the formation of the black hole. This signal has been identified as the lowest order $l = 2$ g-mode (2g_1) of the inner core of the PNS [29].
- (4) *SASI modes*: SASI modes are observed in models in which the SASI is active [16,19,20]. It starts at ~ 100 Hz, usually with some delay after bounce, and its frequency grows in time, albeit at a lower pace than g-modes. Its frequency growth is close to linear rather than an arch.
- (5) *Memory*: The explosion and the anisotropic neutrino emission, leave a low frequency signal in the range ~ 1 –10 Hz (e.g., [18,32]), usually described as a memory effect.

B. Parametrized templates

We concentrate this work in the g-modes, the most common feature of all models, which also are responsible for the bulk of the GW signal in the postbounce evolution of the PNS. The aim of our phenomenological template is to mimic the raising arch observed in core-collapse simulations. To this end we will consider a toy model for the GW emission in CCSNe. The idea is that at each time in the postbounce evolution, the PNS is in quasihydrostatic equilibrium and any perturbation will excite the eigenmodes of the system, in particular g-modes. Nonspherical eigenmodes, in particular $l = 2$ modes, will emit GWs at some characteristic frequencies corresponding to these eigenmodes. This premise has been shown to be a quite accurate description of most of the waveforms [27–29]. These modes are continually being excited by the hot bubble surrounding the PNS, in particular by convective motions and SASI [18,32]. At the same time these excited modes are damped by the PNS conditions (e.g., by

the existence of convective layers that do not allow for buoyantly supported waves) and by the presence of nonlinearities and instabilities (e.g., turbulence). Therefore, the GW emission can be modeled as a damped harmonic oscillator with a random forcing, in which the frequency varies with time.

Following these arguments, we can model the strain measured at the detector as the solution of:

$$\partial_{tt}h + \frac{\omega(t)}{Q}\partial_t h + \omega(t)^2 h = a(t), \quad (1)$$

where $\omega(t) \equiv 2\pi f(t)$ is the angular frequency corresponding to the eigenmode excited in the PNS, Q is the Q-factor, which we consider to be constant for simplicity, and $a(t)$ is an acceleration driving the signal (the random forcing).

We model the frequency as a 2 degree polynomial:

$$f(t) = f_0 + f_1(t - t_{\text{ini}}) + f_2(t - t_{\text{ini}})^2; \quad t \in [t_{\text{ini}}, t_{\text{end}}], \quad (2)$$

where t refers to the postbounce time and f_0 , f_1 , and f_2 are three coefficients determining the behavior of the frequency evolution. t_{ini} and t_{end} correspond to the beginning and end of the signal, being $t_{\text{end}} - t_{\text{ini}}$ its duration. Note that the beginning of the signal t_{ini} do not necessarily have to coincide with the time of bounce ($t = 0$), so it is possible to incorporate in the model the typical delays (up to 200 ms) observed in numerical simulations.

Instead of using f_1 and f_2 directly it is more convenient to define $f_{1s} \equiv f(t = 1 \text{ s})$ and t_2 , the latter being the time at which the polynomial has a maximum. Given that the spectrograms of numerical simulations are not showing any maximum in the evolution of the features characterized as g-modes (at least in the preexplosion phase), the value of t_2 has to fulfill that $t_2 \geq t_{\text{end}}$. Using f_0 , f_{1s} and t_2 as parameters, the frequency can be expressed as

$$f(t) = f_0 + \frac{2(f_{1s} - f_0)(t_2 - t_{\text{ini}})}{(2t_2 - t_{\text{ini}} - 1)(1 - t_{\text{ini}})}(t - t_{\text{ini}}) \quad (3)$$

$$- \frac{f_{1s} - f_0}{(2t_2 - t_{\text{ini}} - 1)(1 - t_{\text{ini}})}(t - t_{\text{ini}})^2, \quad (4)$$

where time is expressed in seconds.

To mimic the sudden downflows observed in numerical simulations responsible for the excitation of the g-modes, we model $a(t)$ as a series of N instantaneous accelerations of the form $a_n \delta(t - t_n)$, $n = 1, \dots, N$, with values of t_n distributed randomly in the interval $[t_{\text{ini}}, t_{\text{end}}]$ and with a random amplitude a_n in the interval $[0, a_{\text{max}}]$. a_{max} is a normalization constant, which we chose to be $\propto \omega^2$. There is an arbitrary constant in the choice of a_{max} , which is not of relevance for this work, because the templates are scalable to any desired amplitude. This constant could be calibrated

in the future to generate distance-dependant templates, although this is beyond of the scope of this work. Also the dependence of this amplitude with ω should be explored in the future.

Finally, instead of using N as a parameter for the template, we use:

$$f_{\text{driver}} \equiv \frac{N}{t_{\text{end}} - t_{\text{ini}}}, \quad (5)$$

which is the driver frequency, i.e., the number of triggers per unit time introduced by the forcing. Physically, f_{driver} is related to the characteristic frequency of the random perturbations exciting modes in the PNS. Since these perturbations are expected to be driven by convection and SASI, its typical frequency is few hundred Hz.

In total we have the next set of 7 free parameters for the parametrized phenomenological template: t_{ini} , t_{end} , f_0 , f_{1s} , t_2 , Q , and f_{driver} . For a given set of parameters we solve numerically Eq. (1) by means of the first order symplectic-Euler method. The computed waveform has some stochastic nature due to the randomness of the amplitude and time of the instantaneous accelerations. Therefore, for a given set of parameters one can generate different realizations, depending on the seed used for the random number generator. This allow us to have variability between waveforms corresponding to different realizations of the same model, something that has been observed when running numerical simulations, e.g., for simulations using different random perturbations in the initial model (see, e.g., [33]).

All the waveforms generated for this work have a sampling rate of 20 kHz, and have been padded with zeroes to a total duration of 2 s (in all cases longer than the duration of the signal) with the signal centered in the interval. To avoid errors from the numerical integrator we use a time-step 10 times smaller than the inverse of the sampling rate.

An example of a waveform generated by our method can be seen in Fig. 1. The parameters used for this example are $t_{\text{ini}} = 0$, $t_{\text{end}} = 0.8 \text{ s}$, $f_0 = 100 \text{ Hz}$, $f_{1s} = 700 \text{ Hz}$, $t_2 = 1.25 \text{ s}$, $Q = 10$, and $f_{\text{driver}} = 200 \text{ Hz}$. Note that the signal (upper left panel) only has power in regions where there are impulsive accelerations (lower left panel). Also the features in the spectrogram (upper right) follow closely the prescribed value for the frequency (lower right). We note that the waveforms used in this study try to represent the typical features observed in simulations of neutrino-driven CCSNe. This features are observed in numerical simulation by all groups in the community and their origin is well understood (g-modes in the protoneutron star). These features are not expected to disappear in future, more detailed, numerical simulations, although the parameter space of possible values for the waveform may change in the future. A difference with respect to other works on the detection of CCSNe GWS is that they make direct use of the waveforms from numerical simulations, while our

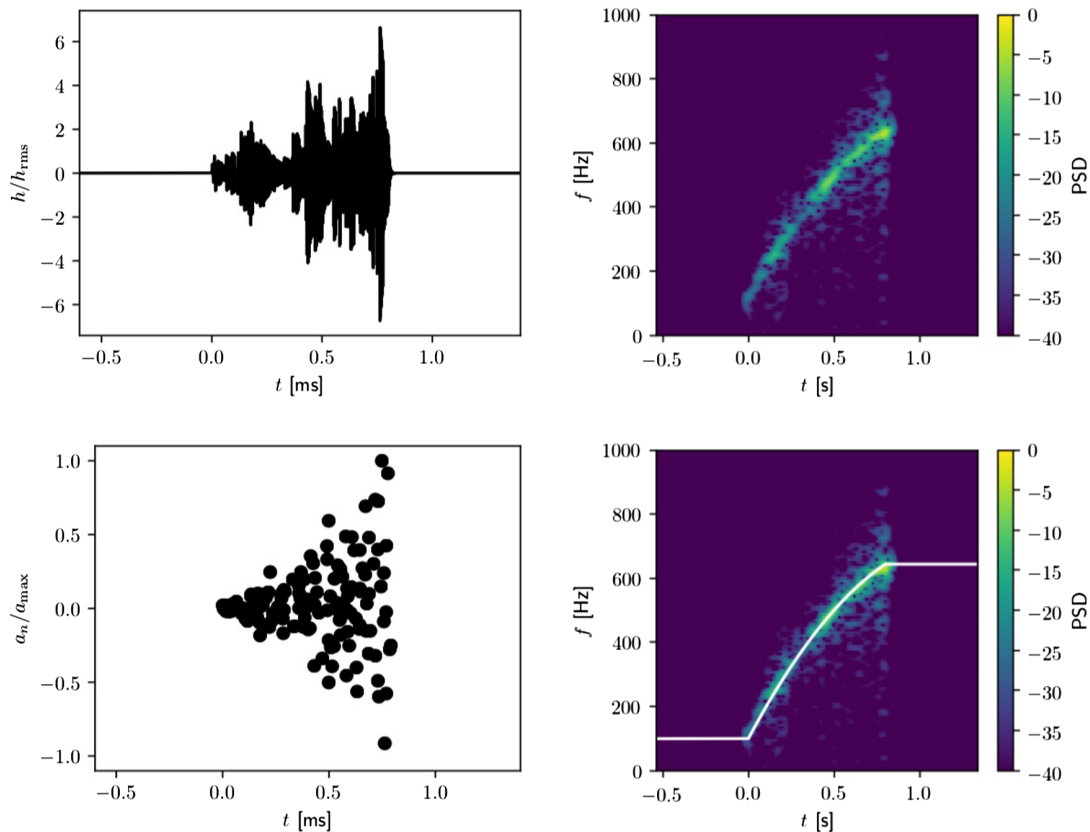


FIG. 1. Example of a realisation of a phenomenological template. We plot the strain (upper left), the corresponding spectrogram (upper right), the coefficient of the impulsive acceleration (lower left), and the frequency of the harmonic oscillator overlotted to the spectrogram (lower right).

approach allows us to choose a relatively wide parameter space that is able to encompass the results from future numerical simulations. There is also room for improvement for the parametrized templates and we aim at making a more comprehensive comparison between these templates and numerical waveforms, however this is out of the scope of this work.

C. Parameter space and template bank

The range of possible values for the 7 free parameters defining the waveforms can be obtained by comparing with the spectrograms of the numerical simulations discussed in Sec. II A. The range of values that we propose here are based on a simple inspection of the work by [15,16,18–20,25], with certain room such that we can accommodate any of these models inside our parameter space. The parameters f_0 , f_{2s} , and t_2 are based in the inspection of the frequency evolution of the predominant feature in the spectrogram. f_{driver} is set to the typical values of SASI and convective motions frequency, as we argue above. The duration of the signal is based on the minimal and maximal duration of all waveforms from nonrotating progenitors (fast rotating progenitors can have a longer duration [34]). Finally, the parameter Q controls the width of the feature in

the spectrogram. Numerical simulations show a wide variety of widths for these features. While some simulations show relatively narrow features (e.g., [15]), which would correspond to $Q \sim 10$, in other cases the signal in the spectrogram is very broad (e.g., [18]), corresponding to ($Q \sim 1$). Note that Q is limited to values larger than $1/2$, otherwise the oscillations become overdamped. Table I shows the parameter space explored in this work.

TABLE I. Parameter space of the phenomenological templates. The second and third columns indicate the range (maximum and minimum, respectively) for each parameter. The fourth shows the value used to generate the template bank in this work. Note that not all combinations are possible since $t_{1s} > t_{\text{ini}}$ has to be fulfilled.

| Parameter | Minimum | Maximum | Test value |
|--------------------------|--------------------|----------|------------|
| t_{ini} [s] | 0 | 0.25 | 0 |
| t_{end} [s] | 0.2 | 1.0 | 0.5 |
| f_0 [Hz] | 100 | 600 | 100 |
| f_{1s} [Hz] | 400 | 2000 | 1100 |
| t_2 [s] | $> t_{\text{end}}$ | ∞ | 0.66 |
| Q | 1 | 10 | 10 |
| f_{driver} [Hz] | 100 | 600 | 600 |

Given that this work is a proof-of-concept of the methods proposed, we use a single test value within the parameter space (see Table I), and we created a template bank containing 100 different realisations of this parameter set. This value is representative of a typical CCSNe waveform and in similar, e.g., to model M15 in [15]. Therefore, the templates used in this work do not cover the all possible CCSNe scenarios and serve just as an example. A deeper analysis covering the whole range of possible CCSNe scenarios will be developed elsewhere.

III. THE METHOD

In this paper we use simulated data having the spectral behavior of the advanced detectors LIGO and Virgo [35], with a standard Gaussian noise assumption. We inject on these data the randomly generated waveforms described in the previous section. The SNR is defined as the square sum of the ratio of the reconstructed waveform in the frequency domain (\tilde{h}_+ , \tilde{h}_\times) and the amplitude spectral density $S_k(f)$ of each detector k :

$$\text{SNR} = \sqrt{\sum_k \int \frac{\tilde{h}_+^2 + \tilde{h}_\times^2}{S_k(f)} df} \quad (6)$$

Our method is essentially a two procedure steps:

- (i) the data preprocessing derived in part from the software pipeline coherent wave burst (cWB) [36], which prepares time-frequency images of the interferometer data;
- (ii) a convolutional neural network (CNN), which provides the classification of images in the noise or noise + signal classes;

A. Data preprocessing

The first step of the analysis is a preprocess based on the initial part of the pipeline coherent wave burst (cWB) of burst search.

cWB is the GW transient signal algorithm in use by the LIGO and Virgo collaborations that made the first alert of the GW150914 signal [37]. It is an algorithm to measure energy excesses over the detector noise in the time-frequency domain and combining these excesses coherently among the various detectors. This is performed introducing a maximum likelihood approach to define the ratio among the probability of having a signal in the data over the probability of only noise. The algorithm is unbiased in the sense that does not depend on expected waveform, making it open to a wide class of transient signals. The algorithm has been recently improved by implementing a new method of estimation of event parameters, which considers assumptions on the polarization state (circular, linear, elliptical, etc...) [38,39].

cWB looks for power excesses in the time-frequency domain using Wilsong-Daubechies-Meyer wavelet transform [40], which allows a better characterization of spectral

features with respect to the Fourier transform. The discrete wavelet transform are performed at different resolutions (wavelet levels), each one is an independent and complete representation of the original data. The likelihood approach allows us to combine different time-frequency levels, having a unique wavelet representation adapted to the characteristic of the signals.

In our method, the preprocess is based on the wavelet transform applied to whitened data, since the time-frequency contains both the signal and the detector noise. cWB extracts from the network data a list of triggers above a defined threshold [41].

While in [42] the time-frequency likelihood is used as input of a neural network for the identification of binary systems, here we apply a different approach more suited for CNN: a fixed time-frequency level, i.e., the one which splits the available frequency band in 64 pixels, while the time size is fixed to 256 pixels.

The frequency upper limit is set to 1024 Hz as we did in [41], and the 256 pixels correspond to a time window of 2 s. If the signal is too short in time to reach the number of 256, the pipeline includes adjacent pixels to image borders on left and right. The extension on the left is randomly chosen in a uniform distribution between zero and the number of missing pixels, while the right one is the complement to the total number.

In practice, starting from the time domain data of the two LIGOs and one Virgo interferometer, we produce three time-frequency sets of images, one for each detector. Because the gravitational-wave signal must be present in at least two detectors we developed a technique to visually enhance the coincidences among all the interferometers of the network. The method consists in using primary colours for the spectrograms of each detector: red (R) for LIGO-Hanford, green (G) for LIGO—Livingston and blue (B) for Virgo (see Fig. 2). Then, the three single-colored spectrogram are stacked together to give as output an RGB image (see Fig. 3). The RGB spectrogram is a compact representation of the data to make evident the cross correlation between different detectors. This is an efficient way to prepare our data for the image recognition task that we will perform with our convolutional neural network.

B. The convolutional neural network

Machine learning has become in recent years a cornerstone for many fields of science and it has been adopted more and more as a valuable asset [43]. It has attracted much interest due to significant theoretical progress and due to the increased availability of large amounts of computing power (GPUs) and easy to use software implementations of standard machine learning techniques.

CNN are designed to deal with spatially localized data, such as those found in images [44]. The network architectures of the CNN can be complex and include operations

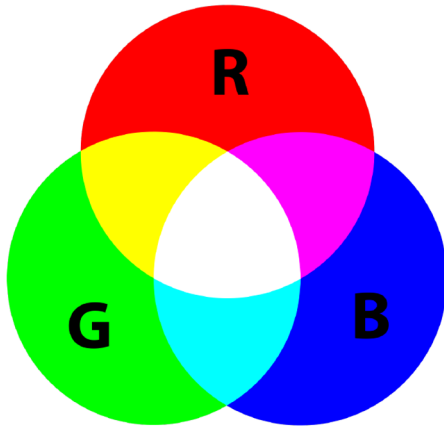


FIG. 2. The mechanism of additive color synthesis. LIGO Hanford is assigned to red, LIGO Livingston to green and Virgo to blue. A triple coincidence will appear in white, while a double coincidence in yellow, magenta or cyan, depending on which couple of detectors is involved.

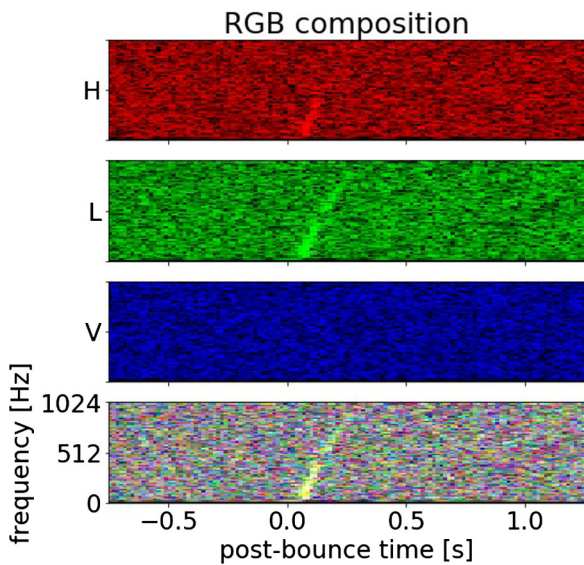


FIG. 3. From the top; the spectrogram of LIGO Hanford is red, then that of LIGO Livingston is green and Virgo is blue. At the bottom: the RGB image obtained by stacking the previous three spectrograms. In this case, the signal is present just in Hanford and Livingston so that the combined signal at the bottom is in yellow.

that go beyond those performed by the individual neurons of the networks. These extensions have allowed CNNs to become the state-of-the-art solution to several categories of problems, most notably photographic image classification [45–47].

Our aim is to provide a clear evidence that the machine learning technique, in particular neural network, can be more efficient to the respect of other approaches to extract GW CCSNe signals, embedded in the detector noise and emitted in the far universe.

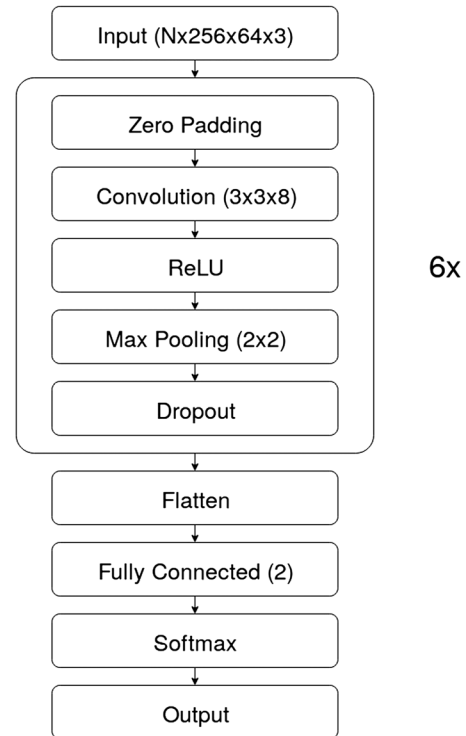


FIG. 4. Sketch of the architecture of our model.

The driving idea is to identify a set of N features in the data chunks, which are the outcomes of the CCSNe 3D simulations. This set of information is used to train a CNN that, thanks to its architecture, can proceed mostly in an automatic way in the learning process.

In this architecture the neuron acts as an image filter and its weight can be thought as a specific pattern. For example, patterns might include different orientation of edges or small patches of color. If the local region of the input matches that pattern, then the single neuron is activated. The input is scanned to look for the set of signal features and the process output is another image indicating where that pattern can be located.

1. The CNN architecture

The model definition, the training and the validation phases have been developed in the TFLearn [48] and Keras frameworks [49], both based on the TensorFlow backend [50].

The network is designed as simple as possible while still having enough variables for the optimisation [51]. In Fig. 4, we sketch the block diagram of the CNN. The images are inputs of the following sequence blocks: *ZeroPadding*, *Convolutional*, *Rectified Linear Unit (ReLU)*, *MaxPooling*, *Dropout*.

The zero padding ensures that the size of the following convolutional output is still a power of two. Every convolutional layer in the network has the same kernel size (3×3) and number of filters (8).

Every convolutional unit has a *Rectified Linear Unit*—*ReLU*, i.e., an activation function defined as:

$$\text{ReLU}(x) = \max(0, x)$$

After the *ReLU* nonlinearity, a *Max pooling* [52] is performed. *Max pooling* is a downsampling process, which halves the image dimensionality, it reduces the computational cost by decreasing the number of parameters to learn and provides local translation invariance to the internal representation. After the *pooling*, the minimum possible contraction to 2×2 implies that the feature map area is shrunk by a factor of four and this operation returns the maximum output within a square neighbourhood.

The final step of every block is a soft dropout aimed to regularize the model and avoid overfitting.

Then, the whole process is repeated six times, then the flattened layer reshapes all the previous neurons in a one-dimensional vector.

This operation erases the information about topology, generally marking the boundary between the convolutional and fully connected part of the model. After the flattening layer, we set a fully connected layer with only two output neurons, one for each class, *noise* and *signal+noise*. Those neurons have a *softmax activation* function, in order to obtain class probabilities as the final output of our classifier. The softmax activation function is defined as:

$$\text{softmax}(\mathbf{z})_i = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad (7)$$

where the indices i and j run from 0 to n and the \mathbf{z} array is the output of the n neurons in the preceding layer.

2. The learning phase and validation

The output of the softmax layer is the probability p , while $q \equiv \{0, 1\}$ is the true image class. The learning phase is performed by a gradient descending algorithm of the loss function [43] toward lower values. To optimise our network we used the class of adaptive learning rate algorithms known as *Adaptive-moments* (Adam), because of its robustness and fast convergence.

In a first phase the CNN has to be trained to classify in the right class the images containing the signal. Artificial neural networks have a huge number of internal parameters to adapt during the learning phase. In general, the more parameters involved the more expressive the network is. More expressive power means a better ability to perform a given task, but also means longer and more difficult training, as well as higher computing resources needed. With the architecture described before we have a total of 3210 partially correlated trainable parameters; this number is enough to represent the knowledge required to successfully perform our classification task.

For the training phase we split the data in two different chunks: the train and the test set. The model is trained only

using the information from the first set, while the second one is never involved in this process. We feed the model with the train set while the test one is used to probe its generalization capabilities, i.e., it is effectively learning general features in our finite data set. The check of the learning process is done periodically to gain confidence on the classification efficiency performed on a completely new and independent set of data. Then, the information about this evaluation will be immediately discarded, without using them for the training: in this way, every successive evaluation will be like the first evaluation of a never-seen-before data set. The training phase stops when the gradient of the loss function is approaching zero within our arbitrary chosen interval of 10^{-4} .

The *curriculum learning* starts the training from higher values of the cost function (*easier classification*) and progressively decreases (*harder classification*). The training set is built using just images where we have signal-to-noise-ratio (SNR) higher than 4 in two detectors at least and we use different data set with decreasing SNR of the network: 40, 35, 30, 25, 20, 25, 20, 15, 12, 10, 8 and, for each of this value, we compute the CNN efficiency, η_{CNN} , defined as

$$\eta_{\text{CNN}} = \frac{\text{correctly classified signals}}{\text{all the signals at the input of CNN}} \quad (8)$$

and the CNN false alarm rate (FAR_{CNN}), defined as:

$$\text{FAR}_{\text{CNN}} = \frac{\text{misclassified noise}}{\text{all classified events}} \quad (9)$$

The class encode q of *noise* and *signal+noise* assumes just the value 0 or 1, while the output of the classifier is instead the real number p , the probability of the image to include a signal. Thus, to separate the two classes, we have to define the threshold θ_{CNN} . The choice is the result of a trade off between the false positive and the false negative classification. The main constraint is to minimize the signal

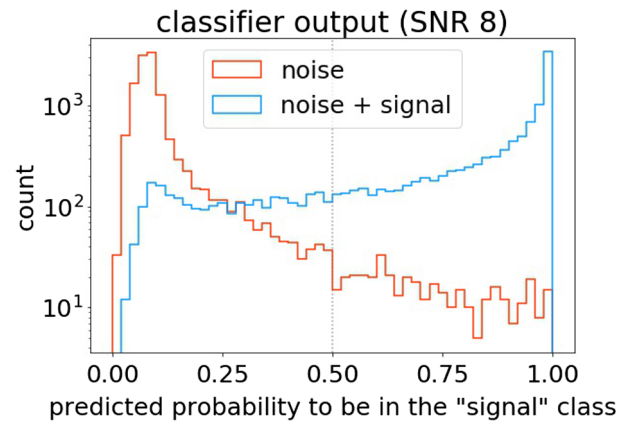


FIG. 5. Distribution of the classifier output for noise (red line) and signal + noise (light blue line) in the case of SNR = 8.

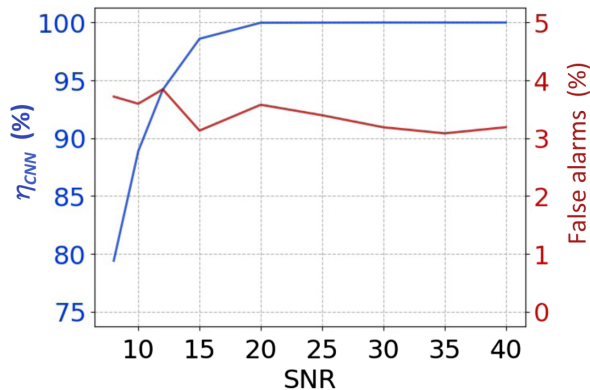


FIG. 6. η_{CNN} (Eq. (8)) for different SNRs (blue line) computed during the validation process at threshold $\theta_{\text{CNN}} = 0.5$. The red line is the false alarm probability [FAR_{CNN} Eq. (8)] associated to the various SNRs computed at the same threshold. The used data set is based on 20000 signals and noise events for each SNR, half used for training and half for validation.

dismissal, even if this implies to include some noise in the form of false positives.

In Fig. 5, we show an example of class classification histograms of the noise and signal + noise in the case $\text{SNR} = 12$. By choosing a threshold of 0.5 in the predicted probability to be in the signal class (see Fig. 5), we obtain the results shown in Fig. 6, where for SNRs between 8 and 15 the efficiency is higher than 80%, for SNR higher than 20 is 1 and the false alarm is confined in the range 3% and 4%.

IV. RESULTS

In order to qualify the method, we compare its efficiency to the complete cWB procedure. The efficiency for each procedure is defined as the ratio of the number of events passing the procedure thresholds and the number of injected events. We simulate the background noise which is equivalent to six years of observation of advanced detectors, applying the usual time-shift procedure of the gravitational data analysis [36]. Then, we inject a subset of waveforms whose parameters are listed in Table I. Since cWB's performances change in function of the SNR, in order to have compared number of events for each SNR value, we inject more signal at lower SNR. For each SNR we build a set of 10000 time-frequency images and we combine them randomly with the same amount of noise images. The six years of observation time is accordingly reduced to the image selection.

The comparison between our method and the complete cWB approach is done through the cross correlation statistics, cc_{cWB} , [41]. In particular, the postproduction thresholds are set as reported in [41], relaxing just cc_{cWB} to the value 0.6, since we were not dealing with real noise. We compute the efficiency curves of cWB and our method for every SNR as function of the false alarm rate. In Fig. 7 we

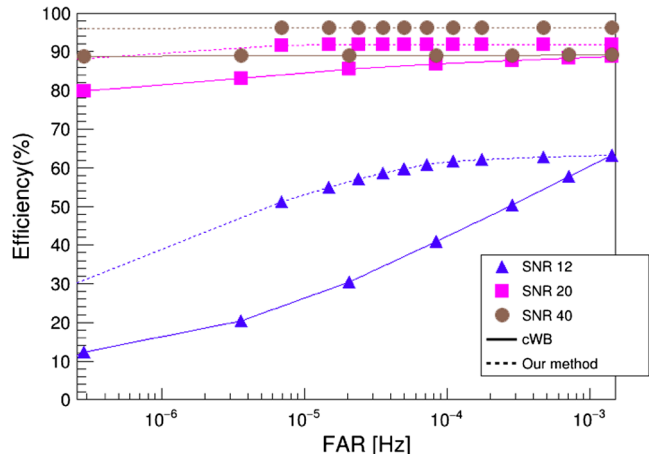


FIG. 7. Efficiency vs false alarm for cWB (continuous line) and our method (dashed line) in the three cases $\text{SNR} = 12$ (blue triangles), $\text{SNR} = 20$ (pink squares), and $\text{SNR} = 40$ (brown circles).

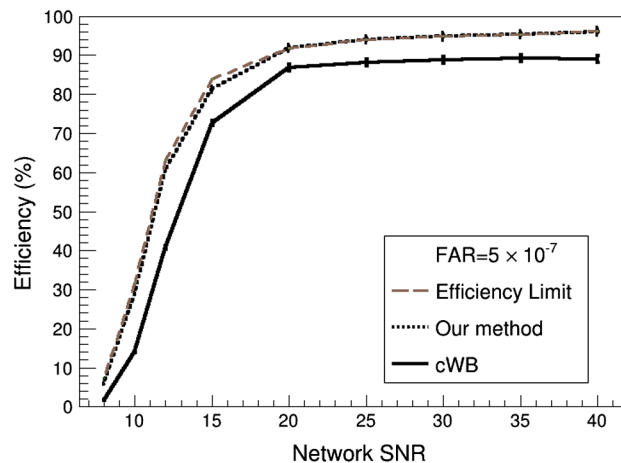


FIG. 8. Efficiency vs SNR in the case of complete cWB (continuous) and our method (dashed) for all SNRs. We report also the curve that shows the ratio between the input events of the CNN and the total injected events in function of SNR (brown). This curve sets the maximum efficiency that our method can achieve.

show just the case of $\text{SNR} = 12$ and $\text{SNR} = 40$ and we note that the efficiency of our method is better than that of complete cWB. The same results are obtained for all SNRs.

In Fig. 8, we plot the efficiency versus the signal to noise ratio of the network for the complete cWB and our method at the false alarm rate of about 7×10^{-5} Hz. Again, we note that our method has improved efficiency with respect to cWB.

In the same figure, we show also the ratio between the input events of the CNN and the total injected events in function of SNR, as listed in Table II. This curve sets the maximum efficiency that our method can achieve. The missing events depend on the cWB postproduction

TABLE II. Table of results at $\text{FAR} = 7 \times 10^{-5}$ Hz. In the columns we report: SNR—the network SNR, Injected—the number of injections, Input of CNN—the number of injections found in the pre-processing stage by cWB and used as input of CNN, Efficiency—the efficiency of the different methods.

| SNR | Injected | Input of CNN | Eff. cWB | Eff. our method |
|-----|----------|--------------|----------|-----------------|
| 8 | 233 994 | 16 878 | 1.7 | 5.9 |
| 10 | 44 946 | 14 237 | 14.1 | 29.0 |
| 12 | 19 500 | 12 333 | 40.8 | 60.7 |
| 15 | 13 500 | 11 346 | 72.6 | 81.3 |
| 20 | 12 000 | 11 029 | 86.9 | 91.8 |
| 25 | 12 000 | 11 290 | 88.2 | 94.0 |
| 30 | 12 000 | 11 385 | 88.8 | 94.8 |
| 35 | 12 000 | 11 450 | 89.3 | 95.4 |
| 40 | 12 000 | 11 534 | 89.0 | 96.1 |

threshold, so that whole improvement is achievable by a better tuning of the cWB inputs.

V. CONCLUSIONS

We have presented a nonlinear method based on convolutional neural network algorithm to extract CCSN

signals embedded in Gaussian noise with spectral behavior of Advanced LIGO and Virgo detectors. We compared the efficiency of the method for different signal to noise ratio to that of the algorithm used by the LIGO-Virgo collaborations to detect gravitational wave transient signals. The results show that our method has an higher efficiency and we conclude that using this new approach we can detect core collapse supernovae taking advantages of the peculiar features of the signal.

In the future, we plan to qualify the method using real detector data which are affected even by non-Gaussian noise.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the supports of the Istituto Nazionale di Fisica Nucleare of Italy and of the *Centro Amaldi* of the department of Physics of the University of Rome *Sapienza* (Grant No. 000008-17 Ateneo2017). P.C. acknowledges the support by the Spanish MINECO (Grants No. AYA2015-66899-C2-1-P and No. RYC-2015-19074) and the Generalitat Valenciana (Grant No. PROMETEOII-2014-069).

-
- [1] B. P. Abbott *et al.*, *Phys. Rev. Lett.* **116**, 061102 (2016).
 - [2] B. P. Abbott *et al.*, *Phys. Rev. Lett.* **116**, 241103 (2016).
 - [3] B. P. Abbott *et al.*, *Astrophys. J. Lett.* **851**, L35 (2017).
 - [4] B. P. Abbott *et al.*, *Phys. Rev. Lett.* **119**, 141101 (2017).
 - [5] B. P. Abbott *et al.*, *Phys. Rev. Lett.* **119**, 161101 (2017).
 - [6] B. P. Abbott *et al.*, *Astrophys. J. Lett.* **848**, L13 (2017).
 - [7] B. P. Abbott *et al.*, *Astrophys. J. Lett.* **848**, L12 (2017).
 - [8] H.-T. Janka, F. Hanke, L. Hüdepohl, A. Marek, B. Müller, and M. Obergaulinger, *Prog. Theor. Exp. Phys.* **2012**, 01A309 (2012).
 - [9] A. Burrows, *Rev. Mod. Phys.* **85**, 245 (2013).
 - [10] K. Hirata *et al.*, *Phys. Rev. Lett.* **58**, 1490 (1987).
 - [11] H. Dimmelmeier, J. A. Font, and E. Müller, *Astron. Astrophys.* **393**, 523 (2002).
 - [12] H. Dimmelmeier, C. D. Ott, A. Marek, and H.-T. Janka, *Phys. Rev. D* **78**, 064056 (2008).
 - [13] E. Abdikamalov, S. Gossan, A. M. DeMaio, and C. D. Ott, *Phys. Rev. D* **90**, 044001 (2014).
 - [14] S. Richers, C. D. Ott, E. Abdikamalov, E. O’Connor, and C. Sullivan, *Phys. Rev. D* **95**, 063019 (2017).
 - [15] B. Mueller, H.-T. Janka, and A. Marek, *Astrophys. J.* **766**, 43 (2013).
 - [16] P. Cerdá-Durán, N. DeBrye, M. A. Aloy, J. A. Font, and M. Obergaulinger, *Astrophys. J. Lett.* **779**, L18 (2013).
 - [17] K. N. Yakunin, P. Maronetti, A. Mezzacappa, O. E. B. Messer, E. Lentz, B. Bruenn, W. Raphael Hix, and J. A. Harris, in *Proceedings of the 50th Rencontres de Morions, 2015 Gravitation 100 years after GR*, edited by E. Augé, J. Dumarchez, J. Tran Thanh Vân, and ARISFS (2015), pp. 367–369.
 - [18] J. W. Murphy, C. D. Ott, and A. Burrows, *Astrophys. J.* **707**, 1173 (2009).
 - [19] T. Kuroda, K. Kotake, and T. Takiwaki, *Astrophys. J.* **829**, L14 (2016).
 - [20] H. Andresen, B. Müller, E. Müller, and H.-Th. Janka, *Mon. Not. R. Astron. Soc.* **468**, 2032 (2017).
 - [21] D. George and E. A. Huerta, *Phys. Lett. B* **778**, 64 (2018).
 - [22] J. Powell, A. Torres-Forné, R. Lynch, D. Trifirò, E. Cuoco, M. Cavaglia, I. S. Heng, and J. A. Font, *Classical Quantum Gravity* **34**, 034002 (2017).
 - [23] M. Razzano and E. Cuoco, *Classical Quantum Gravity* **35**, 095016 (2018).
 - [24] H. Gabbard, M. Williams, F. Hayes, and C. Messenger, *Phys. Rev. Lett.* **120**, 141103 (2018).
 - [25] K. N. Yakunin *et al.*, *Phys. Rev. D* **92**, 084040 (2015).
 - [26] H. Sotani and T. Takiwaki, *Phys. Rev. D* **94**, 044043 (2016).
 - [27] A. Torres-Forne, P. Cerdá-Durán, A. Passamonti, and J. A. Font, *Mon. Not. R. Astron. Soc.* **474**, 5272 (2018).
 - [28] V. Morozova, D. Radice, A. Burrows, and D. Vartanyan, *Astrophys. J.* **861**, 10 (2018).
 - [29] A. Torres-Forne *et al.*, arXiv:1806.11366.
 - [30] H. Sotani, T. Kuroda, T. Takiwaki, and K. Kotake, *Phys. Rev. D* **96**, 063005 (2017).
 - [31] S. E. Woosley and J. S. Bloom, *Annu. Rev. Astron. Astrophys.* **44**, 507 (2006).

- [32] A. Marek, H.-T. Janka, and E. Müller, *Astron. Astrophys.* **496**, 475 (2009).
- [33] A. Summa, F. Hanke, H.-T. Janka, T. Melson, A. Marek, and B. Müller, *Astrophys. J.* **825**, 6 (2016).
- [34] P. Cerda-Duran, N. DeBrye, M. A. Aloy, J. A. Font, and M. Obergaulinger, *Astrophys. J.* **779**, L18 (2013).
- [35] B. P. Abbott *et al.*, *Living Rev. Relativity* **21**, 3 (2018).
- [36] S. Klimentko, S. Mohanty, M. Rakhmanov, and G. Mitselmakher, *Phys. Rev. D* **72**, 122002 (2005).
- [37] B. P. Abbott *et al.*, *Phys. Rev. Lett.* **116**, 061102 (2016).
- [38] S. Klimentko, G. Vedovato, M. Drago, F. Salemi, V. Tiwari, G. A. Prodi, C. Lazzaro, K. Ackley, S. Tiwari, C. F. Da Silva, and G. Mitselmakher, *Phys. Rev. D* **93**, 042004 (2016).
- [39] I. Di Palma and M. Drago, *Phys. Rev. D* **97**, 023011 (2018).
- [40] V. Necula, S. Klimentko, and G. Mitselmakher, *J. Phys. Conf. Ser.* **363**, 012032 (2012).
- [41] B. P. Abbott *et al.*, *Phys. Rev. D* **95**, 042003 (2017).
- [42] S. Vinciguerra, M. Drago, G. A. Prodi, S. Klimentko, C. Lazzaro, V. Necula, F. Salemi, V. Tiwari, M. C. Tringali, and G. Vedovato, *Classical Quantum Gravity* **34**, 094003 (2017).
- [43] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT, Cambridge, MA, 2016).
- [44] K. Fukushima, *Biol. Cybern.* **36**, 193 (1980).
- [45] E. J. Kim and R. J. Brunner, *Mon. Not. R. Astron. Soc.* **464**, 4463 (2017).
- [46] Q. Feng *et al.*, *IAU Symp.* **325**, 173 (2017).
- [47] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, and P. Vahle, *J. Instrum.* **11**, P09001 (2017).
- [48] TFLearn: Deep learning library featuring a higher-level API for TensorFlow, <http://tflearn.org/>.
- [49] Keras: The python deep learning library, <https://keras.io/>.
- [50] TensorFlow: An open-source software library for machine intelligence, <https://www.tensorflow.org/>.
- [51] S. Haykin, *Neural Networks* (MacMillan College Publishing Company, New York, 2009).
- [52] Y. Zhou and R. Chellappa, in *IEEE International Conference on Neural network* (1988), Vol. 2, pp. 71–78.