PROCEEDINGS OF SCIENCE

Multimessenger challenges for the detection of core collapse supernovae

Irene Di Palma,[a,b,c,*] **P. Cerdá-Durán,**[d] **M. Drago,**[a,b] **M. López,**[e,f] **F. Ricci**[a,b] **and A. Veutro**[a,b]

[a] *Università di Roma La Sapienza, Department of Physics, 00133 Roma, Italy*
[b] *INFN, Sezione di Roma, 00133 Roma, Italy*
[c] *Department of Physics, Ariel University, Ariel, Israel*
[d] *Departamento de Astronomía y Astrofísica, Universitat de València, Dr. Moliner 50, 46100, Burjassot (Valencia), Spain*
[e] *Nikhef, Science Park 105, 1098 XG, Amsterdam, The Netherlands*
[f] *Institute for Gravitational and Subatomic Physics (GRASP), Utrecht University, Princetonplein 1, 3584 CC Utrecht, The Netherlands*

E-mail: irene.dipalma@uniroma1.it, pablo.cerda@uv.es, marco.drago@uniroma1.it, m.lopez@uu.nl, fulvio.ricci@uniroma1.it, alessandro.veutro@uniroma1.it

The recent discovery of gravitational waves and high-energy cosmic neutrinos, marked the beginning of a new era of the multimessenger astronomy. These new messengers, along with electromagnetic radiation and cosmic rays, give new insights into the most extreme energetic cosmic events. Among them supernovae explosion is one of the challenging targets of this new astronomical approach.Gravitational waves, much like neutrinos, are emitted from the innermost region of the core collapse supernova and thus convey information on the dynamics in the supernova core to the observer. They potentially carry information not only on the general degree of asymmetry in the dynamics of the core collapse supernova, but also more directly on the explosion mechanism, on the structural and compositional evolution of the protoneutron star, the rotation rate of the collapsed core, and the nuclear equation of state. The development of a new machine learning algorithm will be described to further improve the detectability of a gravitational wave signal from core collapse supernova and the results obtained will be discussed.

38th International Cosmic Ray Conference (ICRC2023)
26 July - 3 August, 2023
Nagoya, Japan

*Speaker

PoS(ICRC2023)1573

1. Introduction

The advanced generation of gravitational wave detectors has proved their capacity of detecting gravitational waves (GWs) from astrophysical processes. While gravitational waves have been detected from mergers of binary black holes and binary neutron stars, core collapse supernovae (CCSN), have not been detected yet and they still represent a puzzle to solve. We had confirmation of the basic CCSN theory through the detection of MeV neutrinos from the SN1987A [1]: the collapse of a massive star's core is driven by the release of gravitational energy and the vast majority of this energy is realised in neutrinos. However, the details of the mechanism of the explosion are still an open question and the astronomical community is trying to disentangle the supernova explosion mechanism with intense studies.

Massive stars ($M > 8M_{\odot}$) spend most of their lives burning hydrogen into helium, which settles in the core and, when temperatures increase sufficiently, burns into heavier nuclei until iron is reached. The resulting iron core is supported by electron degeneracy pressure. When the core reaches the Chandrasekhar mass, pressure cannot support its own weight anymore and collapses. The collapse of the inner core is stopped abruptly when nuclear saturation density is reached at the center and the inner core bounces back. Shortly after the core bounce neutrino emission carries away energy from the post-shock layer. In the present models of the neutrino driven supernovae explosions, the intense hydrodynamic mass motion has to play a significant role in the energy transfer by the neutrino flux to the the medium behind the stagnating core-bounce shock, reviving the outward shock motion and thus initiates the SN explosion. Due to the weak coupling of neutrinos in the region of this energy deposition, in the hydrodynamic models of the explosions a large variety of physical ingredients are needed [2, 3]. This so called neutrino driven mechanism [4], is the dominant theory to explain CCSN explosions in slowly rotating progenitors.

In a supernova explosion, GWs are generated in the inner core of the source, so that this messenger carries direct information of the inner mechanism. The feasibility of this scenario will be supported by the joint observation of neutrino and gravitational wave emission from CCSN, by assessing the correlation between neutrino emission and collapsed core motion. Although the phenomenon is among of the most energetic in the universe, the amplitude of the gravitational wave impinging on a detector on the Earth is extremely faint. For a CCSN in the center of the Milky way, a rare event, we could expect amplitudes of the metric tensor perturbations ranging between $10^{-21} - 10^{-23}$. To increase the detection probability we should increase the volume of the universe to be explored and this can be achieved both by decreasing the detector noise and using better performing statistical algorithms.

Multi-dimensional simulations of core-collapse super- novae 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 modelling: i) accurate solution of the neutrino transport equations during the evolution; ii) incorporation of the complete in- teractions of electron, muon and tau neutrinos and their anti-particles with matter; iii) use of high resolution to re- solve numerically fine structure features in the convective and turbulent flow around the proto-neutron 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 behaviour 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 world-wide community confidence in the results. Despite of 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 [5] and it can be directly related to the rotational properties of the core [6, 7]. These information can be used in the search of the GW signal embedded in the detector noise, with the perspec- tive to increase the confidence detection of signal emitted in the deeper universe.

2. Methodology

In the CCSNe 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 is to present a search strategy of events in coincidence in the advanced detector network [8, 9], characterised by a raising monotonic behaviour in the time-frequency plane, similar to the one observed in numerical simulations. Our strategy is based on machine learning techniques. These are tools applied even to big chunks of data in different contexts, analysed with minimal human supervision and able to resolve ambiguity and tolerate unpredictability.

To implement our search method, we consider a parametric phenomenological waveform designed to match the most common features observed in the nu- merical models of CCSN. We focus our attention on the g-modes excitation, the most common feature of all models developed so far to describe the CCSN phenomena, responsible for the bulk of the GW signal in the post-bounce evolution of the proto-neutron star. The convolution neutral network (CNN) in use is a specialized kind of Deep Learning algorithm to process data that has a known grid-like topology and can learn to differentiate a variety of input types due to its ability for pattern recognition [10]. In a CNN, the input is convolved with a filter, which varies according to the characteristics of the data since it can be *learned* by the network. The computations are performed at each step, as the filter is slided onto the input to compute the corresponding value in the output feature map. Despite of the automatic learning of the filter, some parameters need to be tuned by hand. For the development of our Mini Inception-Resnet network, including the model definition, the training and the validation phases, we have used the Keras frameworks, based on the TensorFlow backend. In this work we implement a *binary cross-entropy* instead and a sigmoid activation function for the output, i.e. we address the problem as a classification problem with a positive class (event class) and a negative class (noise class). Therefore, the output of the network is a probability vector θ, which contains the probabilities of the template belonging to one class or another. The classification task is then performed according to a pre-defined threshold θ^*, i.e. the template will be classified as *event class* only if this probability overcomes θ^*.

The Advanced LIGO-Virgo detector network collected data for three runs, during which more than 90 GW events have been identified [11, 12]. To assess the robustness of our method, we selected data from the second observing run of the Advanced GW detectors, without relaying on any neutrino information. In particular, we chose a stretch of real data even containing glitches, taken during August 2017, when Virgo joined the run [13]. The period includes about 15 days of coincidence time among the three detectors and we used this data set to generate about 2 years of time-shifts data to train and test the neural network as noise class. We define the starting time of each image every 2 s, the images containing the central time of injected signals are considered as *event class*, instead the ones without signals are the *noise class*. The injected signal is expected to be about 600-700 ms in duration, drawn everywhere in the image, with a small probability to be between two consecutive images. Such images are used for the training in any case, therefore the network can recognise also partial signature of the event. The *training set* for CCSN signals has been constructed injecting waveforms at fixed distances: 0.2, 0.4, 1, 2 and 3 kpc, for each distance, of the order of 70000 waveforms, with random sky localization. 75% of the set is used in the actual training while the remaining 25% is used for validation. In the *blind set* we injected a new ensemble of about 260000 simulated signals, generated by the phenomenological templates described before. The distance is chosen in a uniform distribution between 0.2 and 15 kpc, position in the sky are randomly chosen. This set is used to quantify the detection efficiency and to test the network. It is not involved in the training or validation procedure. For the *test set* we perform injections using CCSN waveforms from numerical simulations found in the literature considering 3D simulations of non-rotating progenitors representative of the neutrino driven mechanism. We injected about 65000 waveforms uniformly in distance and sky directions, from 100 pc to 15 kpc.

3. Results

We train the network using curriculum learning, where we start training with the easiest data sets, and then gradually the task difficulty is increased. During the validation process the the decision threshold $\theta^* = 65\%$ has been chosen. In this section we present the results obtained when we used the network trained and optimized in the previous section on the data of the *blind set* and the *test set*. The network has not been trained by any of the images of these two sets so they can be used for the final test of the performance of the network. The signals injected in the *blind set* correspond to waveforms generated by the same procedure used to generate the training set, while the injections in the *test set* correspond to realistic CCSN waveforms. In Fig. 1 on the left, we report the histogram of the injections in the real noise. Such plot shows the robustness of the decision threshold $\theta^* = 65\%$ even in the case of real detector noise.

In Fig. 1 on the right we plot the Receiver Operating Characteristic (ROC) curve and we calculate the area under the curve (AUC). We note the high performance of the *test set* (AUC=0.79) compared with that obtained for the *blind set* (AUC=0.90). Even if we only trained our network with phenomenological waveforms from the template bank, such waveforms mimic the behaviour of the test set, which is the main reason behind such good results.

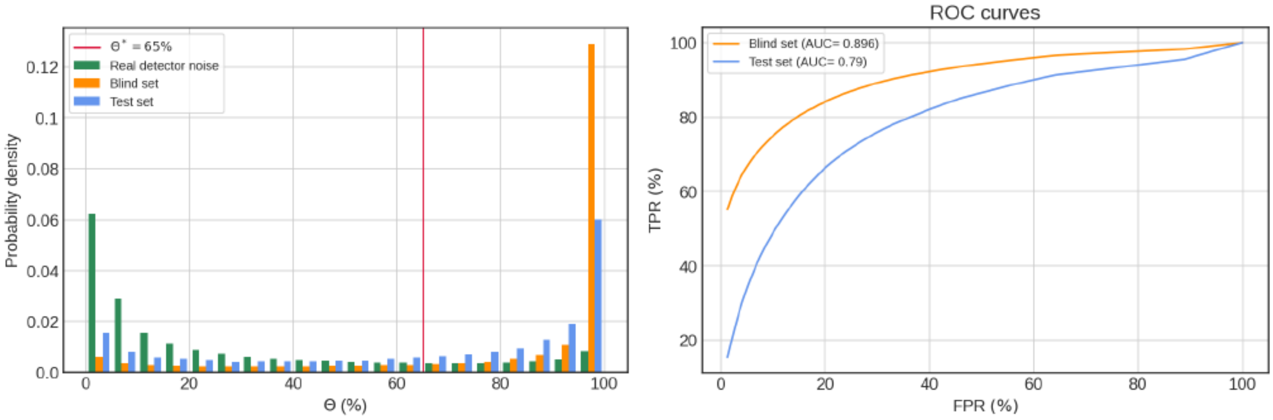

Figure 1: Left: Histogram of real detector noise and injections in real time as function of the probabilities predicted by Mini Inception Resnet. The vertical line represents the chosen decision threshold $\theta^* = 65\%$. Given the counts of the ith bin c_i and its width b_i, we define the probability density as $c_i/(\sum_i^N c_i \times b_i)$, where N is the total number of bins of the histogram. Right: True Positive Rate (TPR) as function of the false positive rate (FPR) in percentage. Performance of our neural network for the *blind set* and the *test set*. AUC is presented in the legend of the plot.

Another interesting graph that shows the resemblance between the *blind set* and *test set* is Fig. 2 on the left. Here we plot the efficiency η_{CNN} as a function of the distance.

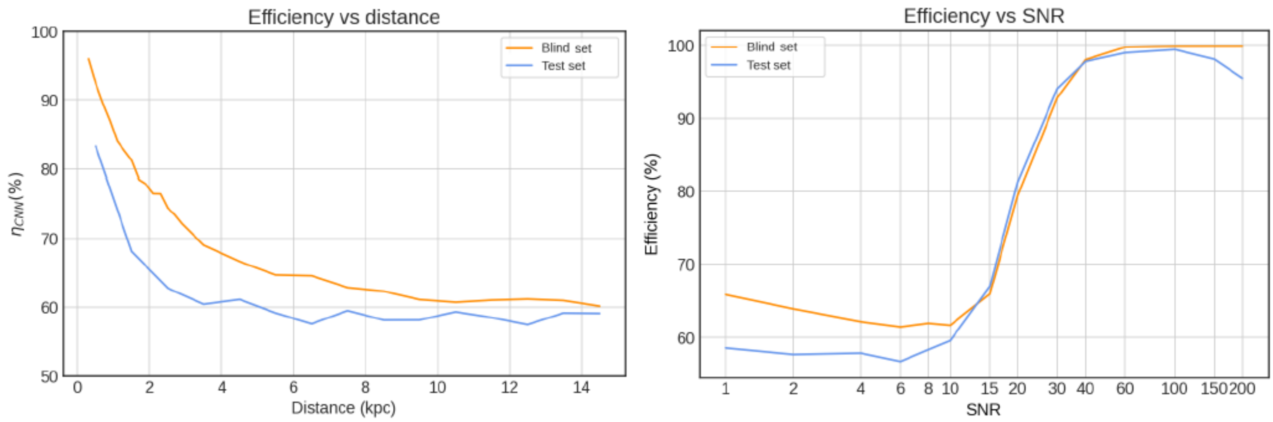

Figure 2: Left: The efficiency η_{CNN} as a function of the distance computed during the testing process in kpc. Right: The efficiency η_{CNN} as a function of the network SNR computed during the testing process

As we can see, at short distances there is a difference in efficiency between *blind set* and *test set* of $\approx 10\%$, but when we increase the distance, they seem to reach a lower limit at $\eta_{CNN} \approx 60\%$. In Fig. 2 on the right we also plotted η_{CNN} against the network SNR. For low SNR, the difference in efficiency η_{CNN} in the two case, *blind set* and *test set*, is around 10%, while for SNR > 15 we obtain the same efficiency. These final results assess the robustness of this method to detect CCSN signals embedded in the real detector noise.

4. Discussion

The search of CCSN signals is carried by a software algorithm whose architecture trains 98997 parameters. We trained the Mini Inception-Resnet using of about 26000 images corresponding to spectrograms of phenomenological waveforms injected in real noise of the three detector network LIGO-Virgo during the second observation run, and similar number of images without signals. We used the curriculum learning with decreasing value of the SNR for the training. At present the entire data processing is rather fast: the training and validation phase, performed in the real detector noise, is done in 2 hours and 21 minute using a GPU *Nvidia* Quadro P5000, while predicting the test set takes 3 ms for each 2 s long image. Given that we take advantage of the Keras/TensorFlow framework, widely used within the machine learning community, it should be easy to increase the complexity of our current CNN or to incorporate the latest developments in machine learning algorithms, with a reasonable increase of the computational cost of the signal search. We developed a new machine learning algorithm to further improve the detectability of a GW signal from CCSN. Regarding the applicability of our method for the GW detection, we have considered a detection threshold, $\theta^* = 65\%$, that results in a FAR of about 5% at SNR$\sim$ 15 (or a FPR of $\sim$ 10% at TPR= 50%). These values could be appropriate for an observation with high confidence of an event in coincidence with a neutrino signal. In those cases the neutrino signal is expected to be bounded within 20 s during the initial SNEWS alert [14] and very likely well within 1 s in the detailed analysis of high sensitivity neutrino detectors such as Super-K [15]. If the method were to be used in all-sky non-triggered searches, the range of values of FAR needed to make a detection with high confidence could be achieved by using values of θ very close to 100%. The efficiency of the algorithm in this regime is something that could be explored in future work. These results are very promising for future detections of GWs from CCSN, because the network allows us to observe more than half of the events within 15 kpc.

5. Acknowledgement

This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (//https://www.gw-openscience.org/ /), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain. PCD acknowledges the support from the grants PGC2018-095984-B-I00, PROMETEU/2019/071 and the *Ramon y Cajal* funding (RYC-2015-19074) supporting his research. M.L. is supported by the research program of the Netherlands Organisation for Scientific Research (NWO). In addition, IDP and FR acknowledge the support from the Amaldi Research Center funded by the MIUR program "Dipartimento di Eccellenza" (CUP:B81I18001170001), the Sapienza School for Advanced Studies (SSAS) and the support of the Sapienza grant RG12117A87956C66 and RM120172AEF49A82.

References

[1] Katsuhiko Sato and Hideyuki Suzuki, Phys. Rev. Lett., 58:2722–2725, 6 1987.

[2] H. Th. Janka. arXiv:1702.08825.

[3] H. Th. Janka, Neutrino Emission from Supernovae, pages 1575–1604. Springer International Publishing, Cham, 2017.

[4] H. A. Bethe. Supernova mechanisms. Rev. Mod. Phys., 62:801–866, 10 1990.

[5] K. Hirata et al., Phys. Rev. Lett., 58 1490, 1987.

[6] H. Dimmelmeier, J. A. Font and E. Müller, A&A, 393, 523, 2002.

[7] H. Dimmelmeier et al., Phys. Rev. D, 78, 064056, 2008.

[8] P. Astone et al., Physical Review D, 98, 12 2018.

[9] M. López et al., Physical Review D, 103, 6, 2021.

[10] Y. Bengio I. Goodfellow and A. Courville. Deep Learning. MIT Press, 2016. http://www.deeplearningbook.org

[11] B.P. Abbott et al., Phys. Rev. X 11, 021053, 2021.

[12] B.P. Abbott et al., Phys. Rev. X, 9(3):031040, 2019.

[13] R. Abbott et al., SoftwareX 13 (2021) 100658.

[14] P. Antonioli et al., New Journal of Physics, 6:114, July 2004.

[15] S. Fukuda et al., Nuclear Instruments and Methods in Physics Research A, 501(2):418–462, April 2003.