GHENT
UNIVERSITY

Detection of Supernovae and their characteristics with Einstein Telescope

Marco Vanderpoorten

Promotors:
Dr. Gergely Dálya
Prof. Dr. Archisman Ghosh

University of Gent
Faculty of sciences
Department of Physics
Academic year 2022-2023

Preface

First of all, I want to thank Prof. Archisman Ghosh and Dr. Gergely Dálya for introducing me to the world of gravitational wave astronomy and supernova physics. They were always there to help and fix my Linux and especially BayesWave troubles.
I also want to thank my girlfriend, family and friends for proofreading parts of my thesis even though not everything was straightforward for people with no physics background.

Abstract

Earlier studies have shown the potential of machine learning to determine the properties of supernovae out of their gravitational wave signature. Here, improvements on the performance of the algorithms have been made by choosing a suitable set of hyperparameters. 14 different 3D core-collapse supernova simulations were used on which noise samples resembling the sensitivity of the Einstein Telescope were added. The waveforms were then reconstructed using BayesWave. These reconstructions were then used as training data for the algorithms. It has been shown that the morphologies of the detector sensitivities of LIGO and ET have little effect on the performance of the algorithms and the choice of hyperparameters, determined by the successive halving method, plays a significant role in the performance of the algorithms. When confronting the algorithms with an unknown supernova model, the predictions are not very accurate. It was concluded that more supernova models are necessary to prove the applicability of machine learning in practice.

From the improved sensitivity of the ET compared to LIGO, one could expect supernovae to be detectable from a larger distance. The detection limits of 7 supernova models are determined by simulating 2000 events over weeks worth of observation time. BayesWave then determines whether the injected waveform was a signal, a glitch or noise. It was proven that even the lowest energy supernovae are detectable up to the galactic center where the chances of a supernova event are the highest. Very energetic supernovae could in principle be detectable up to the Andromeda galaxy.

Contents

1 Introduction

Supernovae are one of the most energetic events in the Universe. One of the first observations of a supernova was already at the beginning of the 11th century when people around the world noticed a bright new dot in the sky with a visual magnitude of around -7.5 [54].
Many more observations of supernovae would follow in the next centuries, but one thing remained very speculative. The physics behind a supernova explosion. How can this explosion take place [53]? The major reason for these unanswered questions is the fact that the electromagnetic waves produced by supernovae, only escape the outer layer of the star hours to days after the explosion [23]. In other words, we only receive information about the event hours to days after it took place. This is a problem as the main physics of the supernova occurs in the first seconds of the explosion [25].
Fortunately, there are other ways to observe supernovae. One of them is via gravitational waves. As gravitational waves can travel through any material, they thus could give us insights into the processes that occur in the first seconds of a supernova.

Now I'll give you a brief summary of the content of this thesis. In the remaining part of this chapter, I will delve deeper into the theory of gravitational waves and supernovae to understand how these events can even produce gravitational waves. Then I'll explain how these waves can be detected and which problems could arise in detecting them. In the second chapter, I'll provide the necessary information to understand the work that was done. Also, an overview of the research will be given. In the remaining chapters respectively the results, possible future work and a general conclusion will be presented.

1.1 Gravitational wave theory

The theory behind gravitational waves is very complex and goes very deep. As the main subject of this thesis isn't theoretical but data analysis, I won't go as deep as possible. I rather give a simplified overview of the theory behind it to prove its utility in supernovae physics and astrophysics in particular. To do so, this and the next chapter are primarily based on both the lecture books written by Michele Maggiore[28, 29] and the course notes of the lecture "Relativiteitstheorie" by Prof. Van Acoleyen[6].

As we are looking at gravitational waves, we start with our leading theory of gravity: General Relativity, of which the basic equation is the Einstein equation:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{1.1.1}$$

With $R_{\mu\nu}$ and R respectively the Ricci tensor and Ricci scalar, $g_{\mu\nu}$ the metric and $T_{\mu\nu}$ the energy-momentum tensor. This equation is invariant under any "behaving"[1] coordinate transformation. Knowing this, we assume that there exists a physical situation where the weak-field limit holds. In other words, we assume the metric to be a small perturbation on top of the flat Minkowski metric:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \qquad |h_{\mu\nu}| \ll 1 \tag{1.1.2}$$

[1] $x^\mu \to x'^\mu(x)$ with $x'^\mu(x)$ invertible, differentiable and with a differentiable inverse.

This way, the Einstein equation can be massively simplified by only considering terms up to first order in the perturbation:

$$\Box\overline{h}_{\mu\nu} + \eta_{\mu\nu}\partial^\rho\partial^\sigma\overline{h}_{\rho\sigma} - \partial^\rho\partial_\nu\overline{h}_{\mu\rho} - \partial^\rho\partial_\mu\overline{h}_{\nu\rho} = -\frac{16\pi G}{c^4}T_{\mu\nu} \tag{1.1.3}$$

Where $\Box$ is defined as the flat space d'Alembertian, $\Box = \eta_{\mu\nu}\partial^\mu\partial^\nu$ and $\overline{h}_{\mu\nu}$ as:

$$\overline{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h \tag{1.1.4}$$

With $h = \eta^{\mu\nu}h_{\mu\nu}$. Now it can be shown that there is a remaining symmetry in the choice of our linearized reference frame. Namely:

$$x^\mu \to x'^\mu = x^\mu + \xi^\mu(x) \tag{1.1.5}$$

Where the derivatives of ξ^μ are of the order as $\overline{h}_{\mu\nu}$. With an appropriate choice of ξ we can now impose the Lorentz Gauge ($\partial^\nu\overline{h}_{\mu\nu} = 0$) which first of all reduces the degrees of freedom of $\overline{h}_{\mu\nu}$ from 10 to 6, but also greatly simplifies equation 1.1.3 to a wave equation:

$$\Box\overline{h}_{\mu\nu} = -\frac{16\pi G}{c^4}T_{\mu\nu} \tag{1.1.6}$$

Outside the source ($T_{\mu\nu} = 0$) the solutions of this equation are standing waves with a velocity equal to the light speed c. The name gravitational waves is thus justified. This picture can be simplified even more though. The Lorentz gauge doesn't fix the gauge completely and we can still do another gauge transformation. It can be shown that when $\Box\xi_\mu = 0$ the physical picture doesn't change. This gives us four more degrees of freedom that can be gauged away up to the point there are only 2 independent components in $\overline{h}_{\mu\nu}$ left. A common way to do so is by using the Transverse-Traceless(TT) gauge. This gauge sets $\overline{h}_{\mu\nu}$ to be transverse and traceless. A consequence of the tracelessness is the fact that $\overline{h}_{\mu\nu} = h_{\mu\nu}$. Using this gauge the form of the perturbative part of our metric becomes:

$$h_{\mu\nu} = \begin{pmatrix} 0 & & & \\ & h_+ & h_\times & \\ & h_\times & -h_+ & \\ & & & 0 \end{pmatrix} \tag{1.1.7}$$

Here are h_+ and $h_\times$ respectively the plus- and cross-polarization amplitudes. Gravitational waves thus have two polarizations. These amplitudes furthermore have a physical meaning: Consider a ring of test particles in a detector. Out of the geodesic equation, it can be shown that when a gravitational wave passes through, the plus-polarization will then cause the ring of particles to alternately be stretched into the x-direction and squeezed into the y-direction and vice-versa. The cross-polarization has the same effect but rotated over 45 degrees. A visual of this effect can be seen on figure 1.1. A general gravitational wave can then be expressed as a linear combination of both polarizations.

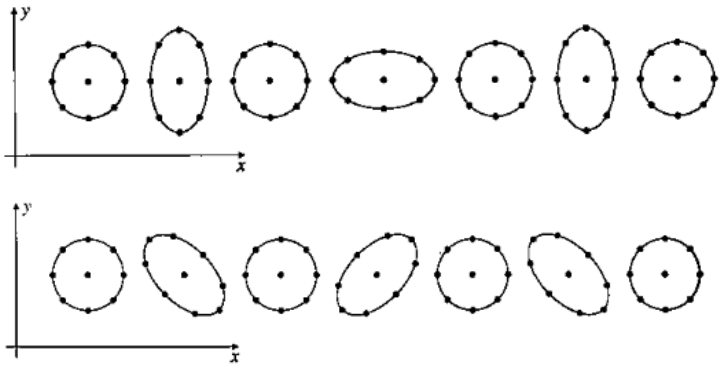

Figure 1.1: Physical effect of both polarizations of gravitational waves. In both cases, the ring is squeezed in one direction and stretched in the other. After half a wave period, the directions perpendicular get squeezed and stretched. Figure from Van Akoleyen [6].

Now that it is clear how gravitational waves propagate through space-time and how they interact with the detector, their generation will be explored and the proof that supernovae can indeed be detected using gravitational waves. For that, equation 1.1.6 has to be solved with a non-vanishing energy-momentum tensor. This is quite a lengthy and complicated derivation so just the results will be presented here. The solution for the trace-reversed perturbation becomes:

$$\overline{h}_{ij}(t,\vec{x}) = -\frac{2G}{c^6 R}\frac{d^2 I_{ij}}{dt^2}(t_r) \tag{1.1.8}$$

Here R is the distance from the source and I_{ij} the quadrupole moment tensor of the source, defined as:

$$I_{ij}(t) = \int d^3y y^i y^j T_{00}(t,\vec{y}) \tag{1.1.9}$$

From this equation, it is seen that the amplitude of gravitational waves drops as $1/R$, while the intensity of light drops as $1/R^2$. It also tells us that gravitational waves are only produced when the source has a non-vanishing second derivative of its quadrupole moment. Effectively this means that only accelerating systems in a non-spherical symmetric way radiate gravitational waves. The signals of these various sources can have varying morphologies and are only visible at certain frequency ranges. These different frequency ranges can then be used to classify the sources. Generally speaking though, these gravitational wave signals can be categorized by asking two questions: Whether the signal is continuous or not and if it can be modelled or not. This way every gravitational wave signal can be divided into 4 categories.

Continuous wave sources are sources that continuously produce gravitational waves with a close to constant frequency. The main example of such a source are neutron stars, if a neutron star has a deformity (so-called "mountain") on its surface, its rotation around its axis isn't spherically symmetric anymore and it will radiate gravitational waves. As these are not yet detected, we know that for the closest neutron stars, these mountains have to be really small (not larger than a couple of centimeters) [1].
The *stochastic background* is not really an individual source, but rather a superposition of gravitational wave signals of individual sources who are unable to be detected separately. The nature of these sources can go from binary black holes up to the gravitational wave equivalent of the Cosmic Microwave Background [16]. The first hints of observation of the stochastic background were made not too long ago by NanoGrav by looking at pulsar timing arrays [10]. The signal was not strong enough to claim a discovery though.

Compact binaries occur when two massive objects orbit each other. While doing so they emit gravitational waves which makes the system lose energy and the orbits of the massive objects smaller until they merge. It was this type of gravitational wave sources that was directly observed with LIGO in 2015 when 2 binary black holes merged [2].
And lastly, *bursts* are very short and energetic events. The primary example of such events are supernovae. As will be explained later, supernova explosions are very unlikely to be spherically symmetric and thus emit gravitational waves. In some cases also black hole binaries are treated as bursts as for intermediate-mass black holes, only their merge is visible in interferometers as a short energetic spike and not their inspiral phase [5].

1.2 Detection of gravitational waves

The detection of gravitational waves can be quite a tedious task as the amplitude of even the most energetic sources is of the order of the radius of a single proton. Most astrophysical sources generate gravitational waves with amplitudes only around one millionth of the radius of a proton. Nonetheless, have astronomers come up with ways to detect these incredibly small waves. At this point in time, two ways of detecting gravitational waves have shown their worth: Pulsar timing arrays and interferometers.

Pulsar timing arrays are a collection of millisecond pulsars, very fast-spinning neutron stars. These pulsars rotate around their axis with a period pretty much completely constant. When their emitted radio waves[2] is traveling to us, the space that that radio wave is going through can be influenced by gravitational waves making the travel time slightly different and thus making the apparent frequency of the pulsars slightly different. So by measuring their apparent frequency of rotation, one could infer the presence of gravitational waves. This method is only sensitive enough in the very low-frequency range where the stochastic background is the main source of gravitational waves.

Contrary to pulsar timing arrays, *interferometers* don't measure the time dilation caused by the gravitational waves but the length contraction. The primary example of interferometers is LIGO, the gravitational wave observatory which also detected the first direct proof of gravitational waves from binary black holes [2]. A simplified picture of LIGO can be seen on figure 1.2. A laser beam is shot into a beam splitter which is then split into two perpendicular arms of 4 kilometers long. In both arms two test masses with mirrors are present. These mirrors reflect the light a couple of hundred times. After that, both beams are collected again and go through a photodetector. Now the two beams are regulated in such a way that when they combine again, they are out of phase and the photodetector cannot detect any light. When a gravitational wave passes, the length of the arms varies by a tiny amount causing a slight phaseshift between both beams. This phaseshift can then be detected by the photodetector.

[2]Pulsars produce high-energetic jets which are well detectable when aligned to us. The signal that reaches us are then flashes of radio waves with a constant frequency.

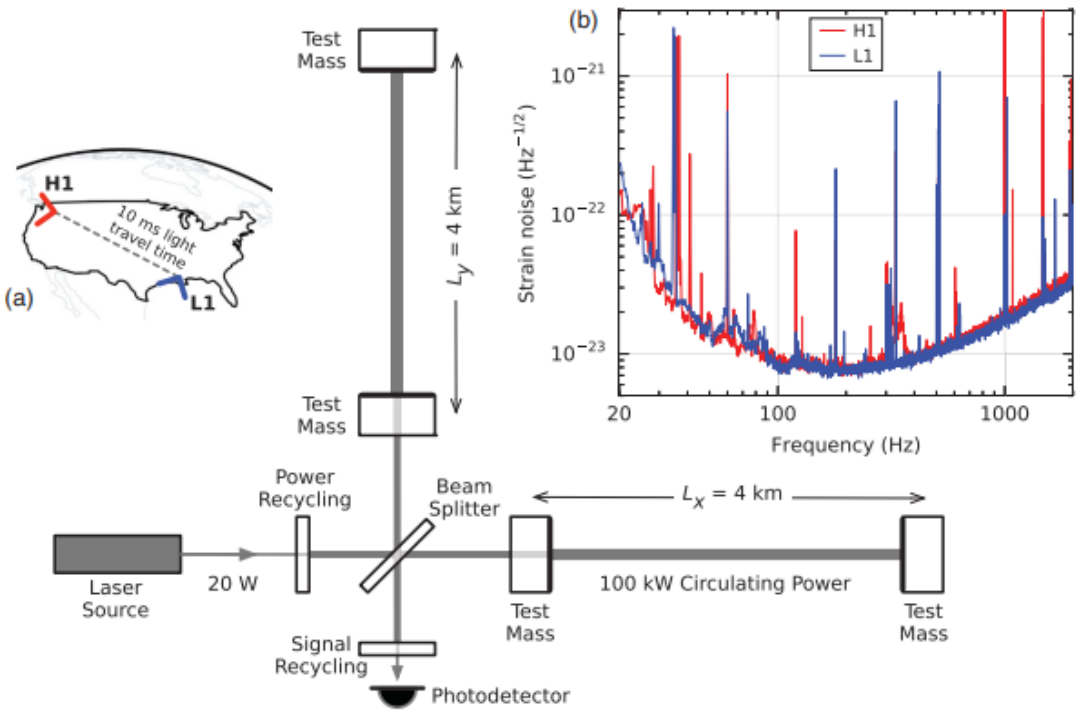

Figure 1.2: A schematic overview of the LIGO interferometers accompanied by their sensitivity curve. (figure from Abott et al. [2])

This technique of course has limits. At some point, other phenomena influence this phaseshift. These phenomena are called noises and can be visualized using a sensitivity curve like on figure 1.2. In gravitational wave interferometers, there are three main sources of noise. Seismic, thermal and quantum noise [30].
Seismic noise is caused by the ground motions and vibrations. These displace the test masses slightly so that the resulting phase shift is slightly altered. Sources of seismic noise can go from earthquakes to vehicle traffic. This noise can be reduced by providing suspension filters to reduce the vibrations and by actively monitoring the motions of the ground and correcting the position of the test masses accordingly. This kind of noise is dominant for low frequencies (up to around 10 Hz for LIGO).
Thermal noise arises from the movement of the atoms of the distinct parts of the detector like the mirrors and suspensions. This noise is particularly visible at the resonant frequencies of these parts, which can be seen as spikes in the sensitivity curve at specific frequencies. This source of noise is prevalent at medium frequencies (10-500 Hz for LIGO).
Quantum noise exists purely because of the quantum nature of light. Random quantum fluctuations can influence the light intensity in the photodetector and thus makes the photodetector measure the incoming light intensity slightly wrong. Quantum noise is the main noise source for high frequencies (from 500 Hz for LIGO).

Other than LIGO, there are also other existing and proposed detectors using the principle of interferometers. Virgo and KAGRA as already existing examples, located in respectively Italy and Japan. Future detectors are proposed like the space-based observatory LISA and the Einstein Telescope[3], from which the latter is the one used in this thesis. The sensitivity curve of the Einstein Telescope (ET) can be seen on figure 1.3 compared with the LIGO sensitivity curve.

[3]One of the proposed places to place the detector is on the Belgian-Dutch-German border.

One can see that the sensitivity of ET is generally one order of magnitude better than the sensitivity of LIGO and for lower frequencies even more. This is mainly because of the design choices of ET. As explained earlier at lower frequencies the seismic noise is dominant. Therefore will ET be built underground as there the ground motion is a lot less. Furthermore, will the arms of ET be 10 kilometers long instead of the 4 kilometers of LIGO. Because the characteristic strain is defined as the fractional change in length over the arms lengths, using longer arms will decrease the measurable characteristic strain. Also, as the noise is inversely proportional to the lengths of the arms [49], longer arms mean less quantum noise. Also, the geometry of ET is different than the ones of LIGO. Instead of two perpendicular arms, a triangular shape has been proposed in which every corner one detector is located. This has a purpose, to be able to measure both the polarizations of the incoming gravitational wave, at least two detectors are needed [28]. The third one can then be used as a valuable check. Every detector is then fed by two interferometers. One optimized for lower frequencies and one optimized for larger frequencies. This way the frequency range of the detector is much larger than LIGO's.

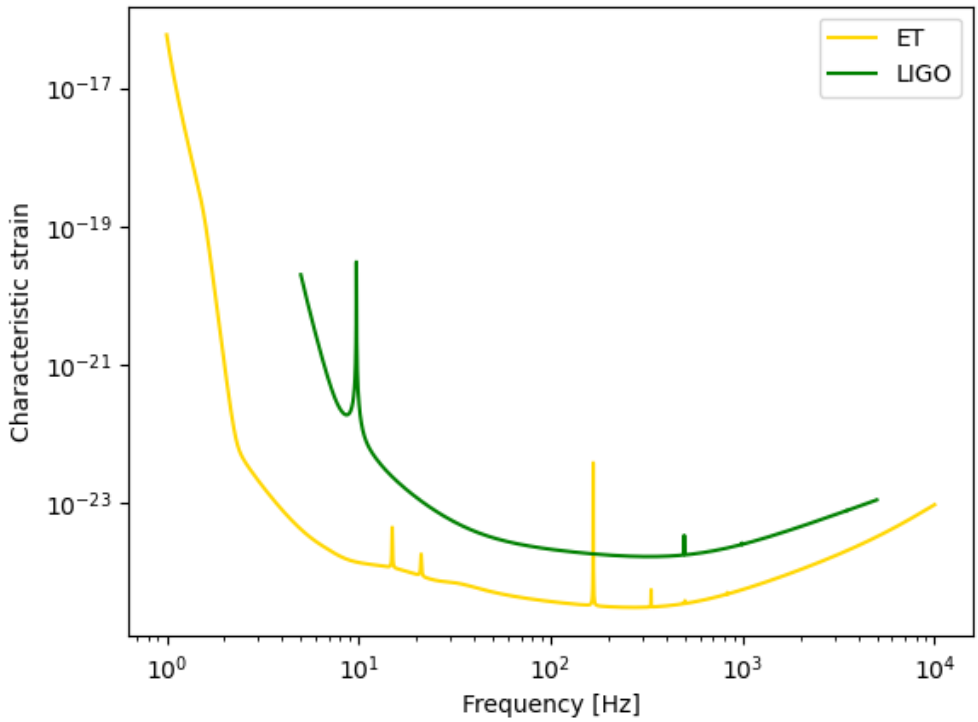

Figure 1.3: Sensitivity curves for both LIGO and the Einstein Telescope. Supernovae are generally observed within the 100-1000 Hz range. In that frequency range, the Einstein Telescope's sensitivity is around one magnitude better than LIGO.

1.3 Supernovae and gravitational waves

1.3.1 Core collapse supernova

To truly understand what supernovae are, one must dive into the stellar evolution. When a star has more than 0.08 solar masses, its gravity is strong enough to heat the core sufficiently so it can start nuclear fusion. The fusion of hydrogen to helium gives the star enough outward pressure to resist gravity. This balance is very stable and only ends when there isn't as much hydrogen left in the core of the star as necessary to efficiently undergo nuclear fusion. The star can still fuse hydrogen in the shells around the core, while the star becomes a red giant. At some point, there also isn't enough hydrogen to fuse in these shells. Now there is no outwards pressure anymore and the core contracts until the temperature is high enough to fuse helium into carbon. Eventually, this fuel is also exhausted. What happens next depends on the mass of the star. When the star has a mass of less than 8 solar masses, it contracts until the core is so dense that electrons are so close together that they produce an outwards pressure by Pauli's exclusion principle. Then the star becomes a white dwarf and dies. There is a limit to how much mass this white dwarf can have though, the Chandrasekhar limit of 1.44 solar masses[4]. When the star has a higher mass, gravity can overcome this degeneration pressure of the electrons and starts fusing subsequently carbon, neon, oxygen and silicon in shells until the core consists mainly of iron. As iron is the element with the highest binding energy, it will not fuse further and suddenly there is no outwards pressure anymore to counteract gravity and the core collapses yet again.

Now the star[5] will explode and cause a supernova. Before the discussion of this explosion, it should be kept in mind that this is not the only way to explode a star. When the previously mentioned white dwarfs are part of a binary or tertiary star system, they can steal mass by their gravity from their colleagues and exceed the Chandrasekhar limit this way. This explosion is fundamentally different than core-collapse supernovae though. They undergo a thermonuclear explosion, which radiates light in a fixed way. Because of this, the distance of such an explosion from us can easily be determined and these explosions are actually used as ways to determine the distance of other astrophysical phenomena close to it. They have one problem in gravitational wave astronomy however. They are much less energetic than core-collapse supernovae and are thus not really detectable in gravitational waves. This thesis will thus only focus on core-collapse supernovae.

As there is no outwards pressure anymore in the star, matter from the outer layers starts falling on the core. The density in the core keeps rising due to the constantly increasing gravity. Now there are two scenarios possible. The first is the case where the core bounces back at nuclear densities because of the nucleon's degeneracy pressure. If there is too much mass falling onto the core, even this degeneracy pressure can't withstand gravity and the star will collapse into a black hole. This bounce of the core already produces gravitational waves. It can be seen in its gravitational wave signature as a large spike in the beginning. Nevertheless, it has to be noted that this bounce often doesn't have a gravitational wave emission comparable to other processes during supernovae [19]. Only when the proto-

[4]Mind the difference between the mass of the star's core and the initial mass of the star. As the star blows away matter in its lifetime, it can start with a mass of up to 8 solar masses and still not exceed the Chandrasekhar limit.

[5]The core of the star before collapse is often called a proto-neutron star

neutron star has a significant rotation, this feature will be important in its waveform. Estimations suggest that only about 1% of massive stars have the appropriate conditions for this bounce to be significant in the gravitational wave signal [24].
The other effects of supernovae are also important for another reason. Simulations have shown that only a bounce doesn't possess enough energy for the star to go supernova. The initial shock wave would be stalled by the still infalling matter and would fade. Over the years, a lot of different scenarios were proposed to solve this problem. These will be explored in the upcoming subchapters.

1.3.2 Explosion mechanism

The explosion mechanism or the shock revival mechanism is the process that revives the shockwave after it has been stalled. One way to do so is by a mechanism that is called "neutrino heating", proposed by Bethe and Wilson [11]. In a last attempt to save the star from collapsing the core undergoes what is called "URCA"-processes [50]. These are processes that produce neutrinos. In supernovae mainly electrons and protons react to form neutrons and neutrinos. The majority of these neutrinos then escape carrying away even more energy and causing the star to collapse even quicker. Some neutrinos are trapped in the ultra-dense core though. When the shockwave is created, the neutrinos are freed and can travel outwards. Now the main idea behind neutrino heating is that although their cross-section is incredibly small, the amount of neutrinos is so high that they still interact considerably with the stalling shockfront, providing the shockfront with enough energy to revive the shockwave.
The mass flow resulting from this mechanism as well as the neutrino emission is suspected to be anisotropic and thus is expected to give a contribution to the gravitational wave signal. Müller et al. [33] found that these contributions could indeed be relevant in the gravitational wave signal.

Another way to revive the shockwave is when a strong magnetic field is present in a rapidly rotating proto-neutron star proposed by Bisnovatyi-Kogan [12]. When strong enough, the momentum of the core can be transferred to the outer layers and thus provide more energy to the shockwave to be revived again. Another possibility to revive the shockwave with magnetic fields is a magnetic instability where a jet-like shock is launched along the rotational axis, exploding the star. Takiwaki and Kotake [52] have shown that only stars with extreme values of their magnetic fields produce a relevant contribution to the signal.

Besides these two models, other more exotic explosion models were proposed. These go from a shockwave resulting from a phase transition of the stellar matter into quark matter [55] to fragmentation of the star by extremely rapid rotation [40]. These models are very speculative and have no numerical evidence backing them up.
Astronomers estimate that around 99% of the supernovae are neutrino driven and the remaining 1% MHD driven [51]. There remains one problem with the neutrino heating mechanism however. It doesn't always provide enough energy for the shockwave to be completely revived again. Therefore other mechanisms which could provide that extra push on top of the neutrino heating were proposed. These will be discussed in the following subchapters.

1.3.3 Convection

Convection in general a flow of a physical quantity originating from a gradient. In the case of supernovae, a lot of gradients are created by the outgoing shockwave, the most important ones being the entropy gradient and the lepton gradient. In supernovae this convection is assumed to occur in different phases [14] with the most important phase being "Prompt convection". Prompt convection occurs immediately after the bounce of the core passes and is caused by both the gradients mentioned earlier. The other phases occur after the shockwave has passed and can be caused by various kinds of gradients.
The contributions of convection to the gravitational wave signal are believed to compete with other sources. Murphy et al. [35] have also shown that these different phases of convection are also visible in the supernova's gravitational wave signature by the time they arrive at the detector. Prompt convection happens immediately after the bounce and after a quiescent period of 100ms, the other phases start.

1.3.4 Standing Accretion Shock Instability (SASI)

At this point, the shockwave is traveling outwards. Blondin et al. [13] have found though that this radial shockwave can oscillate in a linear and/or spiral motion. This non-radial instability is the standing accretion shock instability or simply SASI. Its effect on the oscillation of the shockwave is primarily characterized by the l=1,2 modes of the spherical harmonics. It causes a back-and-forth sloshing of the shockfront, which carries energy outwards very efficiently and can lead to an asymmetric explosion.
In gravitational wave signals, this asymmetrical flow of mass is particularly strong in lower frequencies where gravitational wave interferometers are most sensitive. Murphy et al. [35] have also shown that this signal approximately takes place 200 to 800 ms after the bounce.

1.3.5 g-, f-, p- modes

Because of the SASI mechanism, the proto-neutron star is continuously excited by the accreting matter. The restoring forces to these oscillations can produce powerful gravitational wave signals. These forces can have different natures. G-, f- and p-modes are caused respectively by gravity, surface and pressure forces. Simulations have shown that these signals are very prevalent in gravitational wave signals and the vast majority of existing supernova models exhibit these kinds of modes, mainly g-modes [51].

2 Methods

2.1 Supernova models

As supernovae are unmodelled sources, an exact calculation of the strain of the waveforms of supernovae is impossible. Therefore physicists tried to understand the mechanisms of supernova explosions through numerical methods. These simulations were run with different equations of state and different approaches to approximated general relativity, neutrinos and other physical processes. These factors also influence the produced waveform of the supernova.
In this research, different supernova models are thus used to get a larger variety of possible outcomes. To get a picture as close to reality as possible, this thesis limits itself to numerical simulations in 3 dimensions. The used supernova models and their characteristics are briefly explained in Appendix A. The relevant characteristics in this thesis are summarized on table 1.

2.2 BayesWave

As mentioned before, it is a challenge to detect gravitational waves as their effects are very small. Even with sophisticated detectors like LIGO and ET, the gravitational wave signal is not easily reconstructed over the noise. In general, one needs a good model for the signal and for the noise. When accurate models exist for the incoming waveform, matched filtering is an effective technique even when the noise is poorly understood. Unfortunately, other methods are necessary for unmodelled sources like supernovae. BayesWave is a software designed to reconstruct gravitational burst signals [17]. As the name suggests it uses Bayes' Theorem in its algorithm:

$$p(\mathbf{h}|\mathbf{s}, M) = \frac{p(\mathbf{h}|M)p(\mathbf{s}|\mathbf{h}, M)}{p(\mathbf{s}|M)} \tag{2.2.1}$$

With $\mathbf{h}$ the gravitational waveform, $\mathbf{s}$ the data under model M. $p(\mathbf{h}|\mathbf{s}, M)$ is the posterior distribution, which determines the probability of a gravitational waveform given the data. $p(\mathbf{h}|M)$ is the prior distribution, which encompasses all the initial assumptions of the waveform. $p(\mathbf{s}|\mathbf{h}, M)$ is the likelihood, the probability of seeing the data when assuming a specific waveform and model, and lastly is $p(\mathbf{s}|M)$ the evidence, which quantifies how well the data represents a given model with no knowledge of the gravitational waveform.

The BayesWave approach is now as follows: It defines the data as $\mathbf{s} = \mathbf{R} \star \mathbf{h} + \mathbf{n}_G + \mathbf{g}$. With $\mathbf{R}$ a time-delay operator that describes the gravitational wave detector network response, $\mathbf{n}_G$ Gaussian noise modeled by another algorithm: BayesLine and $\mathbf{g}$ is the glitch model. The signal and glitch model are then found by a Reversible Jump Markov Chain Monte Carlo (RJMCMC) using an unspecified amount of Morlet-Gabor (sine-gaussian) wavelets. Using this information the posterior distribution is calculated, also via a RJMCMC algorithm. Lastly is it also useful to determine whether the data is best described by noise, glitch or a signal. For that, the evidence must be calculated. This is done again with an RJMCMC algorithm to calculate the following integral: $p(\mathbf{s}|M) = \int p(\mathbf{h}|M)p(\mathbf{s}|\mathbf{h}, M)d\mathbf{h}$

2.3 Machine learning

Machine learning focuses on the development of algorithms that enable computers to learn difficult tasks without being specifically programmed to do so. Machine learning can be split into two categories: supervised and unsupervised learning. Supervised learning models are trained using labeled data, which is data for which the correct desired output is known. In this way, the model can adjust itself. In unsupervised learning however, the computer gets no prior knowledge of the dataset and tries to find hidden correlations in it. In this thesis, only supervised learning will be used. In the follow-up of this chapter, some machine learning algorithms and dimensionality reduction techniques will be explained based on the machine learning lecture book "The Elements of Statistical Learning" [22].

Machine learning has several pitfalls that can cause the model to perform poorly on real-life data. Two of these are overfitting and bias. Overfitting can occur when the training set is very large. One would expect that it would perform better when the computer gets fed a lot of learning information. This is only true up to a certain point. When the training data is too large or the algorithm too advanced, the algorithm will learn the data by heart and as a result will perform worse on new data. Bias is a phenomenon that occurs when the training data does not reflect the distribution in real life. The results of the algorithms will then be biased towards a class that is more represented than in real life.

2.3.1 Dimensionality reduction techniques

Principal component analysis (PCA)

The first dimensionality reduction technique is the principal component analysis or simply PCA. What it effectively does, is mapping the data on a lower dimensional space, in which the majority of the information is still contained. The dimensionality of the dataset is then reduced by picking an arbitrary amount of components that possess the most information about the initial data. The way it does that is by fitting a high-dimensional ellipsoid to the data. The axes of this ellipsoid are then the principal components of the dataset. Then when reducing the dimensionality, the principal axis with the largest variances are the ones that store the most data and the whole data is then projected onto those principal components.

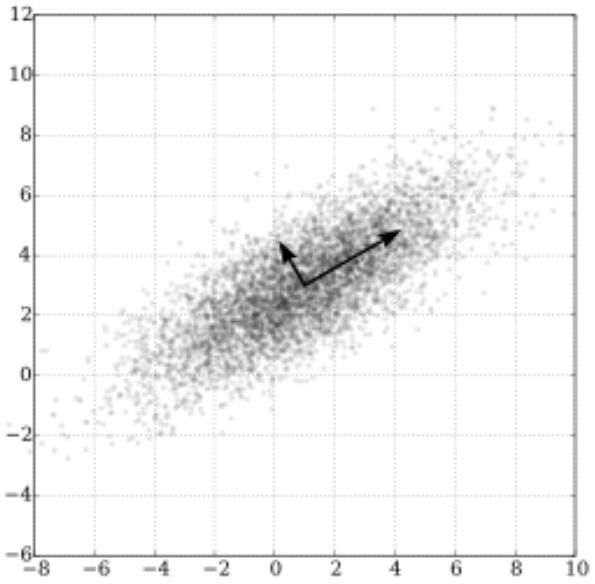

Figure 2.1: Visualization of the principal axis of a Gaussian distribution (from Wikipedia).

Uniform Manifold approximation and projection (UMAP)

The second dimensionality reduction technique is the uniform approximation and projection (UMAP). It differs from PCA as it is a non-linear reduction technique. It tries to fit the dataset on a lower-dimensional Riemannian manifold. The way it does that is mathematically very technical and lies beside the scope of this thesis. More reading can be done in McInnes et al. [31].

2.3.2 Classification algorithms

There are two different classes of supervised machine learning algorithms: classification and regression algorithms. Classification algorithms predict discrete values of a certain feature. Aside from the Neural networks, three other classification algorithms were used: decision trees, support vector machines and K-nearest neighbor.

Decision Trees (DT)

Decision trees are arguably the most intuitive machine learning algorithms in existence. They split the training data horizontally or vertically, one dimension at a time. This way the algorithm tries to reduce the training data into different 'boxes' in such a manner that in every box only one class remains. When confronted with the test data, it follows the made cuts to arrive at a specific cell that has a class assigned to it. The specific test data point is then classified in that class. Now the question at hand is how does the algorithm decide at which place in the high-dimensional space to make that cut. It does this in such a way as to reduce the so-called impurity as much as possible. In other words, make the cut in such a way that on both sides of the cut, one class is as dominant as possible. There are different definitions for this impurity. The one that is used during this thesis is the "Gini-impurity" which is defined as:

$$I_G(p) = \sum_k p_k(1 - p_k) \tag{2.3.1}$$

Here p is the percentage of samples in a cell of a certain class k. As mentioned before, the algorithm keeps making the cuts until there is only one class left in every cell, or the impurity is 0 in every cell. This can be prone to overfitting. A way to deal with this problem is "pruning" or in other words, reducing the length of the tree to improve the accuracy of the tree. This can go from setting a minimum impurity the cell has to have before it can be split to hard coding a limit to the number of cells the tree may possess.

Support vector machines (SVM)

Support vector machines work roughly speaking on the same principle as decision trees. They try to cut the parameter space into different parts and assign a class to every part. The way they do it is fundamentally different however. To truly understand what is happening, a visual representation is very helpful like on figure 2.2. The algorithm tries to find a hyperplane that separates the classes perfectly. This can leave some degrees of freedom in the problem. As can be seen on the figure, a slightly rotated line, still separates the classes just fine. Therefore the hyperplane is chosen such that the margin between the closest samples of each class with the hyperplane is maximal. This way the hyperplane, which is the best at generalizing the problem, will be chosen.
Support vector machines have one big advantage compared to decision trees, which is the possibility to make non-linear cuts in the parameter space. Support vector machines optimize the parameters of a hyperplane of a general form, which can be a linear plane but also a polynomial hyperplane. They do that by using the so-called "kernel trick". In this thesis, only the linear and polynomial kernels will be used.

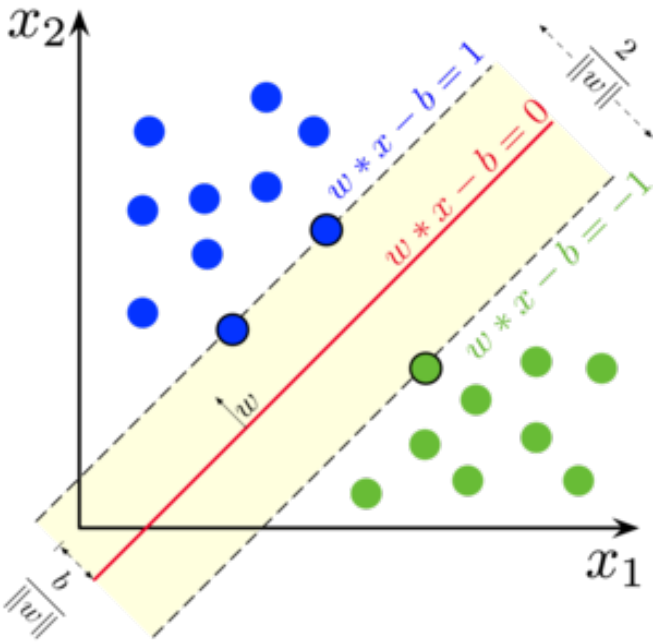

Figure 2.2: An example of how support vector machines decide where to make their cut. In this case, a linear kernel was used. figure from Scikit-learn.

K-nearest neighbor (KNN)

K-nearest neighbor is different from the other algorithms as it doesn't actively learn anything from the data. It stores the sample data in memory and uses that to classify new data. It does that by looking at the closest training points in the parameter space and classifying the test point as the one dominant in those closest training points. The amount of training points that are used to classify the test data is an important hyperparameter in the algorithm and should be chosen with care. In binary classifications like the case in this thesis, it should be chosen as an odd number to prevent a draw.

2.3.3 Regression algorithms

Other features can have continuous values like the progenitor mass and angular velocity. Although there exists alterations to the previous algorithms that can deal with continuous values, an alternate route was chosen. Outside of neural networks, only the effects of linear regression and the least absolute shrinkage and selection operator (LASSO) on the data were studied.

Linear regression (LR)

The first algorithm is linear regression, a simple algorithm that fits the data to a linear function of the form $y = \sum_i w_i d_i$. It tries to determine the ideal weights w_i in such a way that the residual sum of squares of the training set is minimal. This method can be expanded to non-linear regression like polynomial regression.

Least absolute shrinkage and selection operator (LASSO)

Lastly, LASSO works in an analog way as linear regression. It tries to fit the data to a linear function. There is one fundamental difference between the two though. Instead of minimizing the residual sum of squares, an extra term is added to the error calculations as a penalty to keep the weights from growing extensively. This penalty term is of the following form: $\alpha \sum_i w_i$. This also makes the algorithm non-linear in the training set.

2.3.4 Neural Networks (NN)

Neural networks exceed the previous algorithms by far in complexity. They are inspired by the biological neural networks that constitute animal brains. Essentially, they do that by taking an array of input values and doing some complicated computations on them. These computations are mathematically very difficult and so only a short overview of how they work is given.
A good way of understanding how they work is by using an example, image recognition. Every pixel in an image has a certain scalar value corresponding to its color. The collection of these values is then treated as the input for the neural network. Now these values will be sent through some "hidden layers" of the network. These hidden layers consist of an arbitrary amount of "neurons" that take on a scalar value based on the input array. The value of the neuron is calculated by multiplying the value of the input pixel with a specific weight assigned to the connection and performing some "activation" function on that value[6]. In other words, every arrow on figure 2.3 has a weight assigned to it, which is then used to calculate the value of the neuron on the end-point of the arrow. This step is repeated until the output neurons have values. In regression, these values are then the values of the researched characteristic. In the case of classification, the class of the neuron with the highest value will then be assigned to the input data.

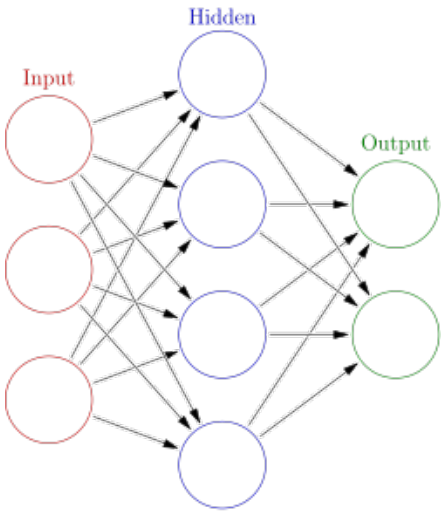

Figure 2.3: A visual representation of a simple neural network. It takes an array of input values and performs a set of calculations to get to an output value (from Springboard.com).

[6]In this thesis, the sigmoid function is used as the activation function: $f(x) = \frac{1}{1+e^{-x}}$

At the start, the weights of the neural network are not optimized and the algorithm would label the input data randomly. The network still needs to be trained. It does that by minimizing the cost function. A function that takes all the initial weights and gives back how bad the algorithm is doing. By trying to make the cost function as small as possible, the weights of the network will be more optimized and the algorithm will perform better on new data. Without going into too much detail, this minimizing is done by an algorithm called "backpropagation". Roughly speaking, altering the weights which have the most effect on the cost function so that the cost function is minimized as much as possible. This process can be very time-consuming for computers so often a simplified method is preferred called "stochastic gradient descent". It calculates how the cost function is minimized the most just as backpropagation but by only looking at a randomly selected small subset of the data instead of the whole data.

2.4 Projects overview

Now that some background has been established of the underlying physics and methods used in this thesis, it is time to elaborate the research that has been done. The latter can be split into two distinct parts. In the first part, it was researched how well the properties of a supernovae event could be reconstructed by only looking at its gravitational wave signature with the aid of machine learning. This project will be given the name "Machine learning project". In the second part, it was investigated from how far supernovae could be detected with ET. This project will be referenced as the "Distance project" later on. In the upcoming subchapters, the goals and methods of both projects will be explained thoroughly, starting with the machine learning project.

2.4.1 Machine learning project

As mentioned before, gravitational waves could give an insight into what happens inside a star when it goes supernova. But as supernovae are unmodelled sources of gravitational waves, determining the characteristics of a supernova event by its gravitational wave signature isn't straightforward. As could be seen in Appendix B, supernovae waveforms are very complicated and determining the main characteristics with the naked eye is impossible. One possible solution to this problem is by using machine learning. In this project, the goal is to see how well the characteristics of supernovae can be determined by using machine learning solely on its gravitational wave signature. This research is a follow-up on work done by Dálya et al. [18]. Here, the performance of the algorithms on data measured by ET will be compared to data measured by the LIGO-VIRGO-KAGRA network. Einstein Telescope is generally one order of magnitude more accurate than LIGO, but there are some fluctuations on that. At some frequencies, it is a bit less and at others a bit more. This could influence the results of the machine learning algorithms. Then, it will be shown that the performances of the machine learning algorithms can be significantly improved compared to Dálya et al. [18]. Lastly, also a more thorough study of the applicability of machine learning in practice will be performed.

The set-up is comparable to Dálya et al. [18]: First, a set of waveforms for training the machine learning algorithms is needed. Unfortunately, no real supernova events have been detected using gravitational wave detectors, so supernova models were used as training data. After that, a new set of supernova models were given to the algorithm. Its predictions were then compared with the known values of the supernova characteristics from those models.

In table 1 a summary of all the used waveforms are given alongside their relevant characteristics for which the machine learning algorithms were tested. Namely, its mass, rotational velocity, explosion mechanism and the presence of SASI and prompt convection. Other characteristics, which were too rare or observed in the vast majority of models were left out as it would be pointless to try to make conclusions out of such a dataset. Because of the limited amount of supernova models, one could expect to see some bias as it is impossible to make a training set that resembles the distribution in real life. For some characteristics, we don't even know the true distribution like the presence of prompt convection or SASI.

Waveform family	Waveform Identifier	Mass[$M_\odot$]	Ang. v[rad/s]	Exp M	SASI	prompt
Andresen et al. 2017 [9]	s11	11.2	0	neutrino	False	False
Andresen et al. 2019[8]	m15fr	15	0.5	neutrino	True	False
Bugli et al. 2021[15]	H	35	10	MHD	False	False
Kuroda et al. 2016[26]	TM1	15	0	neutrino	True	False
Kuroda et al. 2017[27]	s11	11.2	0	neutrino	True	True
Morozova et al. 2018[32]	M10_SFHo	10	0	neutrino	False	True
Müller et al. 2012[34]	N20-2	20	0	neutrino	True	True
O'Connor et al. 2018 [36]	Mesa20_pert	20	0	neutrino	True	False
Ott et al. 2013[38]	s27-fheat1.00	27	0	neutrino	True	True
Pan et al. 2020[39]	s40_fr	40	1	neutrino	True	True
	s40_sr	40	0.5	neutrino	True	True
Powell and Müller 2019[41]	s18	18	0	neutrino	True	True
Powell and Müller 2020[42]	m39	39	3.4	neutrino	True	True
Powell et al. 2021 [43]	Z100_SFHx	100	0	neutrino	True	False
Radice et al. 2019[44]	s9	9	0	neutrino	False	True
	s60	60	0	neutrino	False	True
Scheidegger et al. 2010[47]	R4E1FC_L	15	9.4	MHD	False	True

Table 1: Used supernova models for the machine learning project, accompanied with their relevant characteristics. Respectively their mass, angular velocity, explosion mechanism and the presence of SASI and prompt convection. The latter is not to be confused with general convection. Only a limited number of rotating and MHD waveform families exist.

The majority of waveform families had multiple models. The choice of specific waveforms was such to fill the parameter space as homogeneously as possible. One exception is the explosion mechanism as there only exist two different MHD models at the time of writing this thesis. Every waveform family was represented by 1000 randomly generated waveforms. In the case where there were multiple models used of a single family, the amount of generated models was evenly distributed to get to 1000 waveforms in total.
For the generation of the waveforms, the data files of the authors of these waveforms were used, which can be found on an internal GitLab page here: LIGO GitLab. This data was then randomly orientated and the signal was made exactly 10 seconds long to make a buffer so the whole signal for sure is visible and so there is no bias later on when exposing it to the machine learning algorithms. To get an as accurate picture as possible also the rotation of the Earth and the location of the detector have to be dealt with. Luckily the latter is already built in BayesWave. Only the time of the signal has to be passed through. All the waveforms were generated 10 minutes apart from each other so the total simulation time was $6 \times 10^5 s$ or around a week so Earth's rotation is taken into account. After all the waveforms were generated, a stationary Gaussian noise, that resembles the ET sensitivity curve, was added on top of the waveforms.

The resulting waveforms will then be transformed in such a way that the signal-to-noise ratios of the waveforms lie between 10 and 100. The reason this was done, is twofold. First of all do the different models generally have different energy which means that at different distances, some waveforms would not be recognizable while others would be almost fully reconstructed. To give every waveform the same weight in the machine learning algorithms, working with a signal-to-noise ratio over distances is a good idea. Furthermore is it reasonable to assume that for improved sensitivities, the results would not change for new or improved detectors which would make the SNR a way to give results for every existing or future detectors. The validity of this last point is something that will also be explored in this thesis, as mentioned earlier. Szczepańczyk et al. [51] have shown that supernovae are at best distinguishable at an SNR of around 10, seen in figure 2.4. Generating waveforms below that limit would thus be pointless as BayesWave wouldn't even be able to find the waveform over the noise. On the other hand, Raza et al. [46] showed that at SNR 100 the waveforms were almost completely reconstructed (see figure 2.5). Going beyond this limit wouldn't give us any new insights as the waveforms would just keep being almost fully reconstructed. Finally, this weighting to SNR is done by multiplying the signal by the ratio of the old and new SNR.

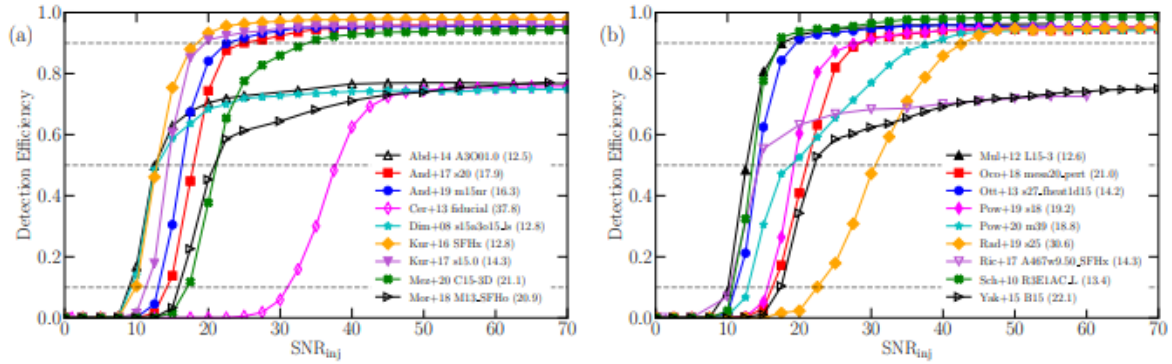

Figure 2.4: Detection efficiency as a function of the signal-to-noise ratio for LIGO O5 from Szczepańczyk et al. [51]. The supernova models only are detectable when they have a signal-to-noise ratio of at least 10.

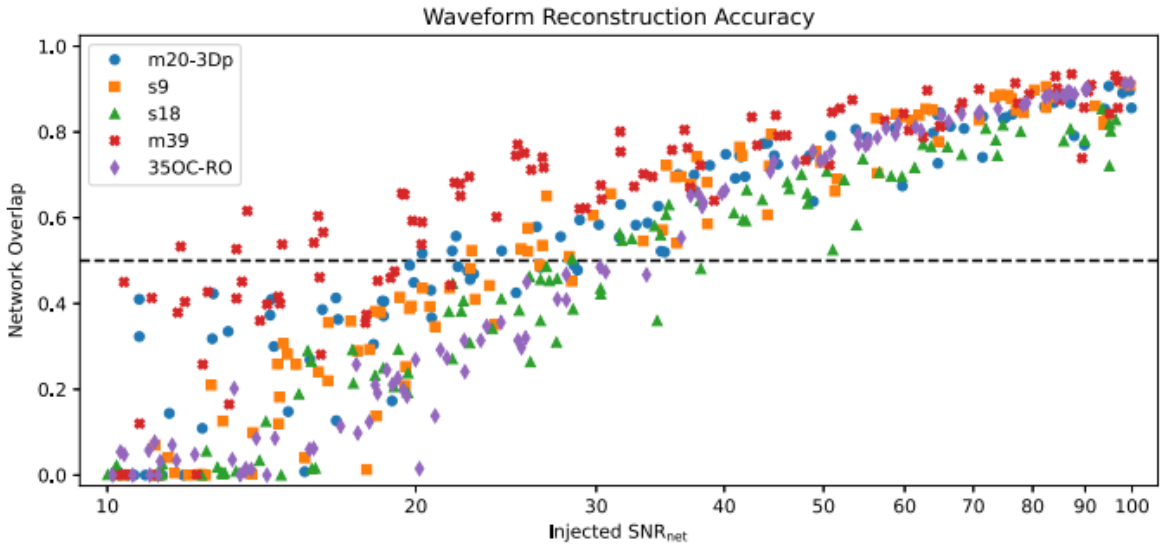

Figure 2.5: Waveform overlap as a function of signal to noise ratio for LIGO-VIRGO from Raza et al.[46]. At signal-to-noise ratios of around 100, the waveforms are quasi fully reconstructed.

Now the waveforms are ready to be sent through BayesWave. In his MSc thesis, Raza[45] investigated the optimal hyperparameters for using BayesWave in supernova research. He recommended using 4 million RJMCMC iterations, a frequency range from 16 Hz to approximately 2 kHz and a segment length of at least 4 s. All of these parameters were used here except for the number of iteration steps. As BayesWave is a very time-expensive algorithm, it was chosen to use only 400.000 iteration steps instead of 4 million so the total simulation time was only weeks instead of months. This reduction in accuracy could of course come with a price. As for especially high SNR signals, the algorithm needs more iteration steps to converge to a steady distribution[45]. Moving on, The data which will be let loose on the algorithms have to be figured out. The straightforward answer would be the strain of the waveforms, but this could potentially lead us to trouble. As all waveforms were made exactly 10 seconds long, there is a clear start and end point of the signal visible. The machine learning algorithms could use this information to learn and thus wouldn't be trained to find the characteristics in the signal itself. Obviously, this is not ideal, so another approach was taken. Of every waveform, the power spectral density in a given frequency range was used to train the machine learning algorithms. 90% of the total waveforms were used to train and the other 10% to test the algorithm's performance. This split was randomized every time. When researching the applicability of this method in practice, all the waveforms of one model were completely left out of the training data. The waveforms of that one model then served as test data. Lastly, the input data was also scaled to have a mean of 0 and unit variance as some machine learning algorithms do not cope well with unscaled data and was reduced in dimensionality by the techniques described in section 2.3.1.

2.4.2 Distance project

The second part of this thesis is about determining the distance at which a supernova could be detected with ET. As the latter has a proposed sensitivity to be a magnitude better than current detectors, it would be interesting to see how much more ET would detect in the years it would be operational. This work was done in an analog way as earlier studies about the detection limit of LIGO O3 and O5 by respectively Abbott et al. [4] and Szczepańczyk et al. [51].

Again a collection of supernova models were used in this project. This time the selection was made based on the gravitational wave energy, as one could expect high-energy supernovae to be more detectable at a greater distance. Two low, medium and high-energy supernovae models were chosen together with the Morozova M10_SFHo model as in Szczepańczyk [51] et al. coherent WaveBurst was used instead of BayesWave and the last named waveform was badly reconstructed there. Here it was researched whether this bad reconstruction was model or algorithm-depended. A summary of the used waveforms can be seen in table 2.

Waveform family	Waveform Identifier	$E_{GW}[M_{\odot}c^2]$
Morozova et al. 2018 [32]	M10_SFHo	1.66×10^{-8}
Obergaulinger et al. 2020 [37]	O	unknown
O'Connor et al. 2018 [36]	mesa20_pert	9.5×10^{-10}
Powell and Müller et al. 2019 [41]	s18	1.6×10^{-8}
Powell and Müller et al. 2020 [42]	m39	7.5×10^{-8}
Radice et al. 2019 [44]	s9	1.6×10^{-10}
Scheidegger et al. 2010 [47]	R3E1AC_L	2.2×10^{-7}

Table 2: Used supernova models for the distance project with their gravitational wave energy. The Obergaulinger waveform's exact gravitational wave energy is unknown, but the model has been postulated as a high-energy waveform.

Of each model, 2000 waveforms were created at a random distance from the detector. The random values were not drawn from a uniform distribution though. As the range in distance was very high, it was chosen to represent the data on a log scale and thus the random values were drawn from a log-uniform distribution. apart from this, the generation of the waveforms happened in the same way as explained before in the machine learning project. The scaling with SNR also wasn't done. Now the waveforms are ready to be sent through BayesWave. Now another mode was chosen. When switched on, BayesWave calculates the chance that the injected signal was indeed a signal, a glitch or noise. Whatever BayesWave picks as the most probable outcome, is the outcome that was assigned to that waveform. Lastly, all the hyperparameters of BayesWave that were used in the machine learning project will also be used here. One remark though, as the number of iteration steps could influence the behavior at large SNR, the effect of using fewer iteration steps was first researched.

3 Results

BayesWave was used as a search algorithm and reconstructed the waveform. Two different plots made by BayesWave can be seen on figure 3.1 and 3.2 for respectively low and high SNR. There, a strain vs time-plot and the power spectrum of the reconstructed waveform can be found. The strain is expressed in withened strain, which is in units of noise standard deviations. So a value of 1 means 1 sigma away from the median noise. The power is expressed as a power spectral density which has units 1/Hz. In both plots, the grey line is the noise, the green line the injected signal and the purple line is the signal BayesWave was able to reconstruct. In the remainder of this thesis, only the power spectra are used. The difference in reconstruction between a low and high SNR is clear.

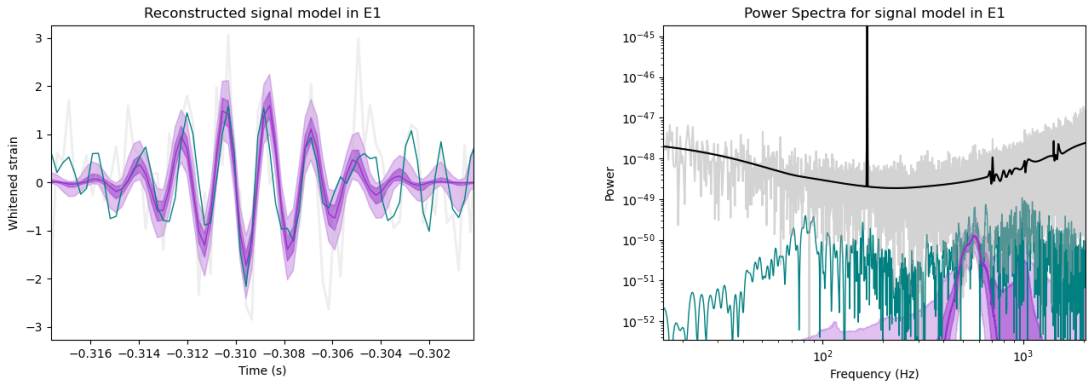

Figure 3.1: The reconstruction by BayesWave of the strain vs time- and the power spectrum for the O'Connor '18 mesa20_pert model for low SNR (17.1).

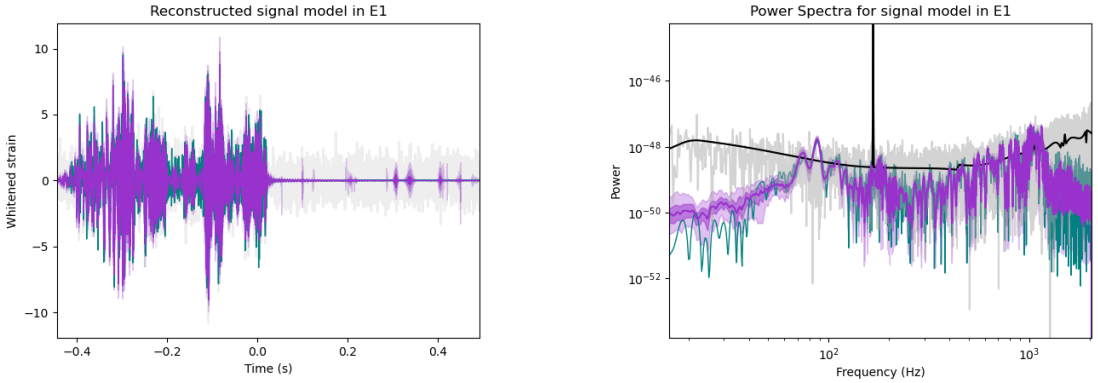

Figure 3.2: The reconstruction by BayesWave of the strain vs time- and the power spectrum for the O'Connor '18 mesa20_pert model for high SNR (93.4).

3.1 Machine learning project

Difference Einstein Telescope and LIGO-VIRGO-KAGRA

Now the first question asked, is whether the performance of the algorithm depends on the detector. For that, the algorithms were trained on 1000 waveforms for 10 supernova models. These are the ones mentioned in table 1 excluding the Bugli '21, Morozova '18, Müller '12 and Powell '20 models. To be able to compare the performance between the algorithms across the different supernova characteristics, the performance of the classification algorithms is measured in balanced accuracy (BA). This is defined as the average over all classes of the amount of correctly predicted samples per class divided by the total number of samples in that class. In our case, there are only two classes, True or False. For the regression algorithms, another unit has to be used. The normalized root mean square error (NRMSE) was used and is defined here as:
$NRMSE = \frac{1}{\sigma_O}\sqrt{\frac{1}{n}\sum_{i=0}^{n}(y_{i,O} - y_{i,P})^2}$
With $y_{i,O}$ the real values of the model and $y_{i,P}$ the predicted values by the algorithms. σ_O is the standard deviation of the test set.

Every algorithm was then trained 25 times to test the difference between ET and the LVK-network. The results can be seen on figures 3.3 and 3.4. There the difference is defined as $BA_{ET} - BA_{LVK}$ for the classification algorithms and $NRMSE_{ET} - NRMSE_{LVK}$ for the regression algorithms. On the plots, the means are plotted together with the 1σ-ranges with a bin-size of 5 SNR. One can see that the difference varies a lot between different algorithms, dimensionality reduction techniques and characteristics. It is obvious that the errors coming from the fluctuations of the mapping between the sensitivities of both detectors have to be negligible, if any, compared to the systematic errors caused by the algorithms and dimensionality reduction techniques. The results obtained in this thesis thus can be generalized to any existing or future gravitational wave detector.

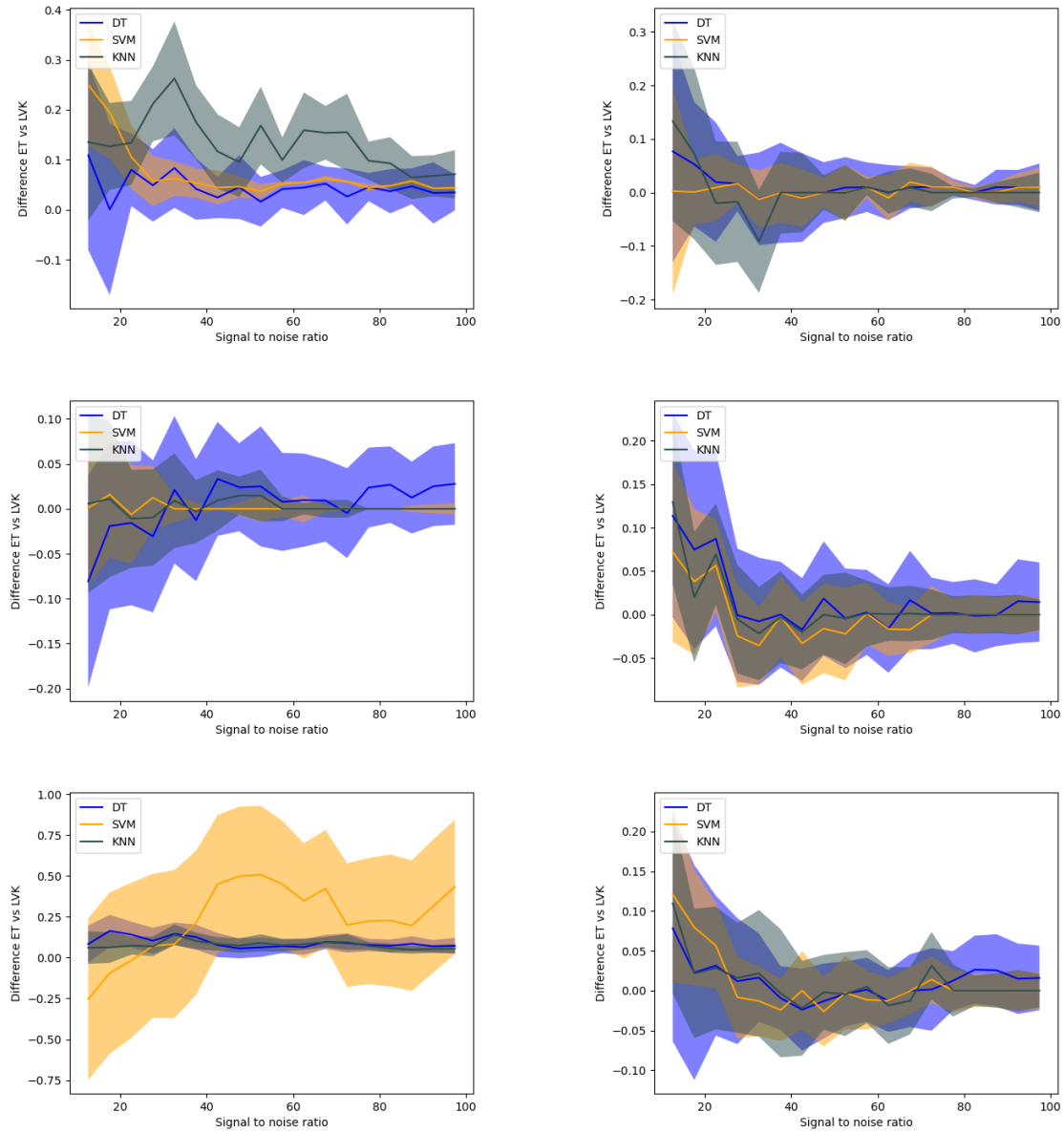

Figure 3.3: Difference in accuracy between the Einstein Telescope and the LVK-network for for respectively the PCA and UMAP reduction techniques. On the first row the results for the explosion mechanism are plotted, on the second and third row are respectively the results for the presence of prompt convection and SASI plotted. The shaded areas show the 1σ-range over 25 realizations.

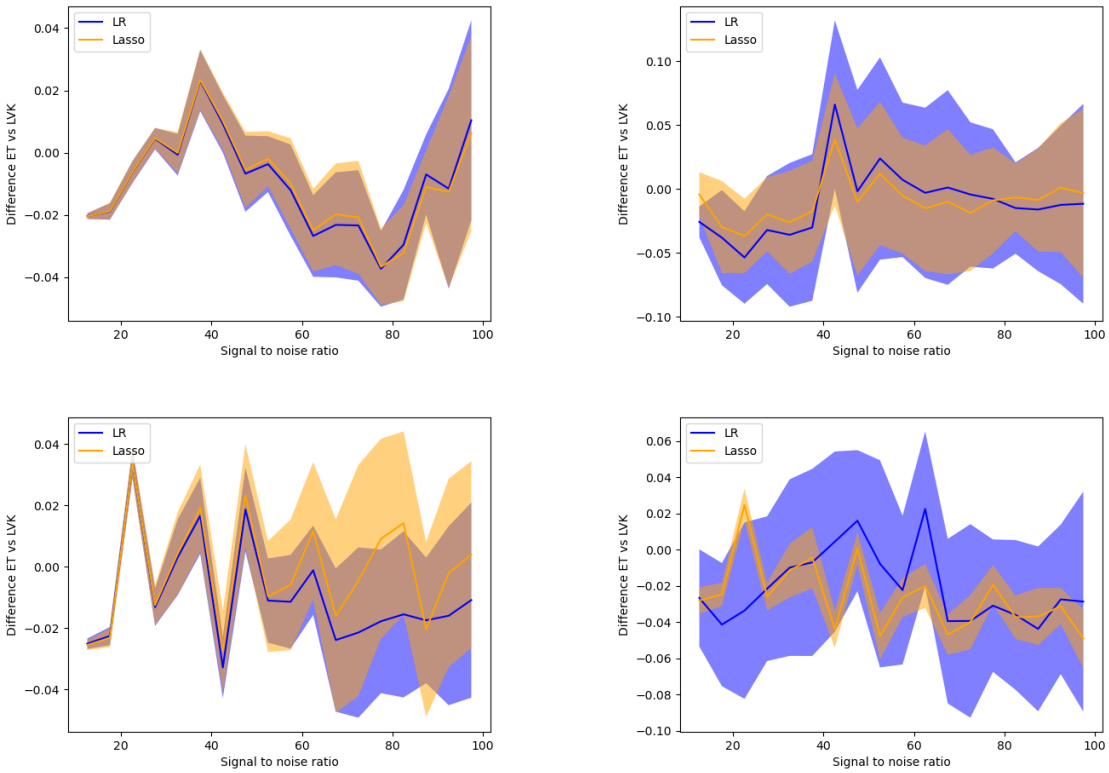

Figure 3.4: Difference in normalised root mean squared error between the Einstein Telescope and the LVK-network for respectively the PCA and UMAP reduction techniques. On the first row are the results for the progenitor mass, on the second the results for the angular velocity. The shaded areas show the 1σ-range over 25 realizations.

Improvements on Dálya et al. [18]

In their previous study, Dálya et al. reduced the dimensions up to around 800 and used the vanilla hyperparameters of the machine learning algorithms built in. The chances are very high that these sets of parameters are not the ideal ones for the problem at hand. The first change that was made was the number of dimensions the data was reduced to. The main reason for that is time-wise. The computation time it takes to reduce the dimensions exponentially grows when considering more than around 300 dimensions like what can be seen on figure 3.13. Although that figure was used for another part of this thesis, the time plot is relevant in this discussion too, unlike the NRMSE. It appeared that lower dimensions also could give equally as good results as higher dimensions. So it was chosen to reduce the dimensionality of the data to 50 dimensions to keep as much information but with reasonable time-loss.

The hyperparameters were treated on a completely different way. One option to determine the best set of hyperparameters is to go over every parameter one by one and search for the best-performing one. This leads to inaccuracies though, as some combinations of hyperparameters that independently don't result in the best performance, can together produce better results. Therefore another approach was opted: The successive halving method. In this method, the ideal set of hyperparameters is estimated by considering every collection of proposed parameters. It then calculates the best-performing sets with a small amount of test samples. The worst-performing half is then left out and the cycle begins again but now with more test samples. This is repeated until one set of hyperparameters remains. These are then used to train the algorithm and predict the values of the test set. This method is clearly visualized on figure 3.5.

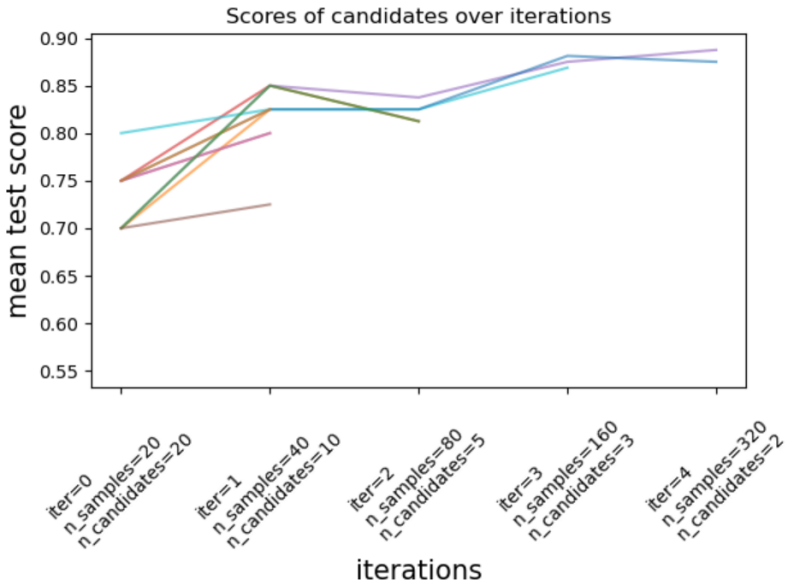

Figure 3.5: Visualisation of the successive halving method. The performance of the algorithm for every combination of proposed hyperparameters are calculated for a small amount of samples. The best half is then put to the test again to more samples. This carries on until just one set of hyperparameters remain (from Scikit-learn.org).

This way the machine learning algorithms are trained using all of the models referenced in table 1. The results of both the classification and regression algorithms are then found on figures 3.7 and 3.9. The shaded areas are again the 1σ-ranges but now with 10 realizations. Comparing these results to Dálya et al. (figures 3.6 and 3.8), every classification characteristic performs considerably better for both dimensionality reduction techniques. Every algorithm predicts the classes almost flawlessly for SNR approximately greater than 25, whereas originally this was only the case for some specific combinations of algorithms and supernova characteristics. Especially the troubles decision trees and K-nearest neighbors sometimes had, were caused by a suboptimal choice of hyperparameters. Also, the standard deviations are now smaller than before, meaning that the algorithms are more consistent in their predictions. And how do neural networks compare to the other algorithms? Using PCA, the advanced machine learning algorithm doesn't perform remarkably better than the other, more simple algorithms. Using UMAP, it even performs worse. It seems to be overfitting the data and does not add anything valuable on top of the simpler classification algorithms.

The same conclusion can be said about the regression algorithms. The performance differences are even more explicit than with the classification algorithms. When using PCA, the standard deviations are lower. The opposite is true when using UMAP. One can see that the steep increase for high SNR using PCA still is present and even more pronounced. This steep increase is probably caused by the choice of hyperparameters in BayesWave. As talked about before, the amount of iteration steps was chosen to be 400.000 instead of the recommended 4.000.000, which could lead to non-convergence for high SNR signals. This non-convergence is now visible as an increase in NRMSE. The bad performance of LASSO also seems to be caused by a bad choice of hyperparameters. Now for every combination of supernova characteristics and dimensionality reduction technique, LASSO performs comparably to linear regression. Lastly, neural networks prove to be a valuable addition to the regression algorithms. When the data is reduced using PCA, it performs remarkably better than the other two algorithms and for medium-range SNR also better than any algorithm using UMAP.

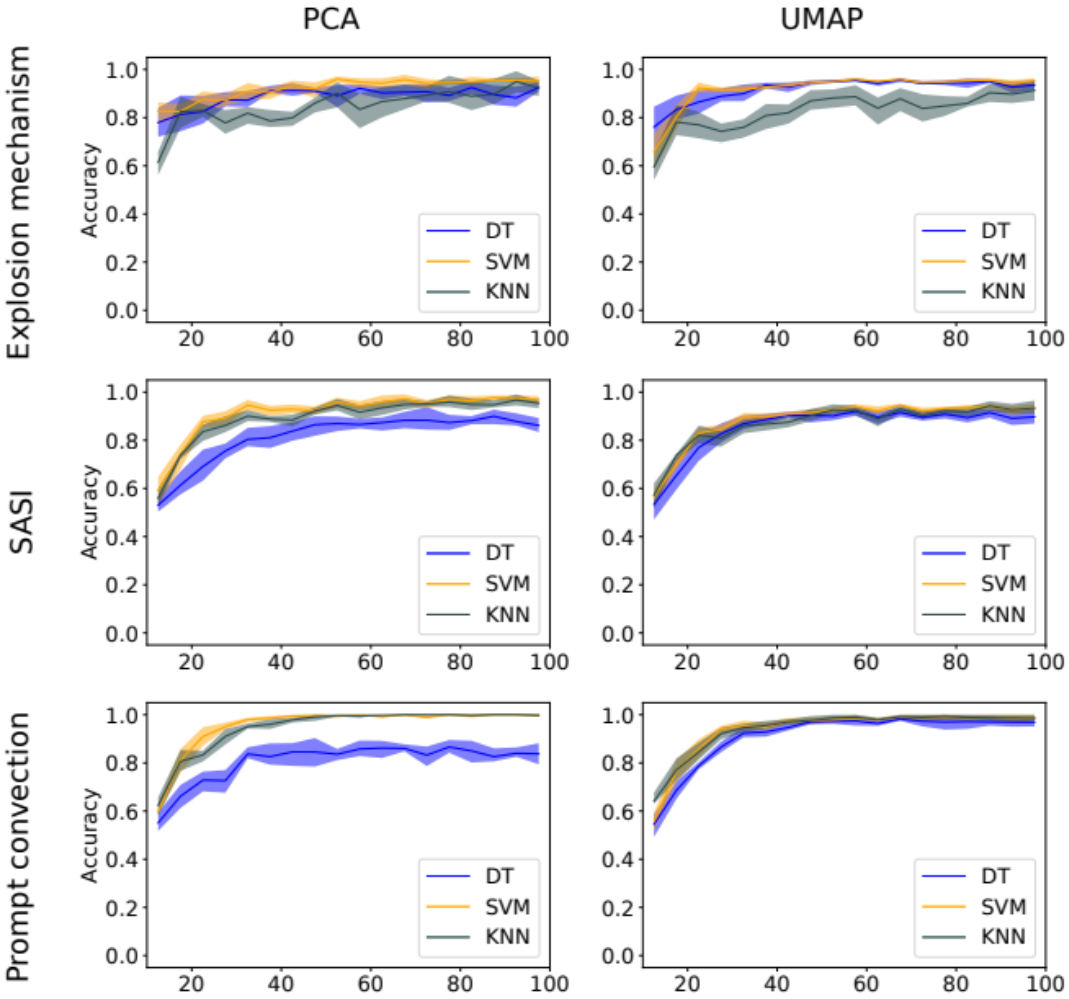

Figure 3.6: Original balanced accuracy of the predictions by the classifier algorithms as a function of the SNR. The shaded areas are the 1σ-ranges over 10 realizations. On the left are the results for PCA and on the right for UMAP. The rows represent respectively the results for the explosion mechanism, presence of prompt convection and SASI (from Dálya et al. [18]).

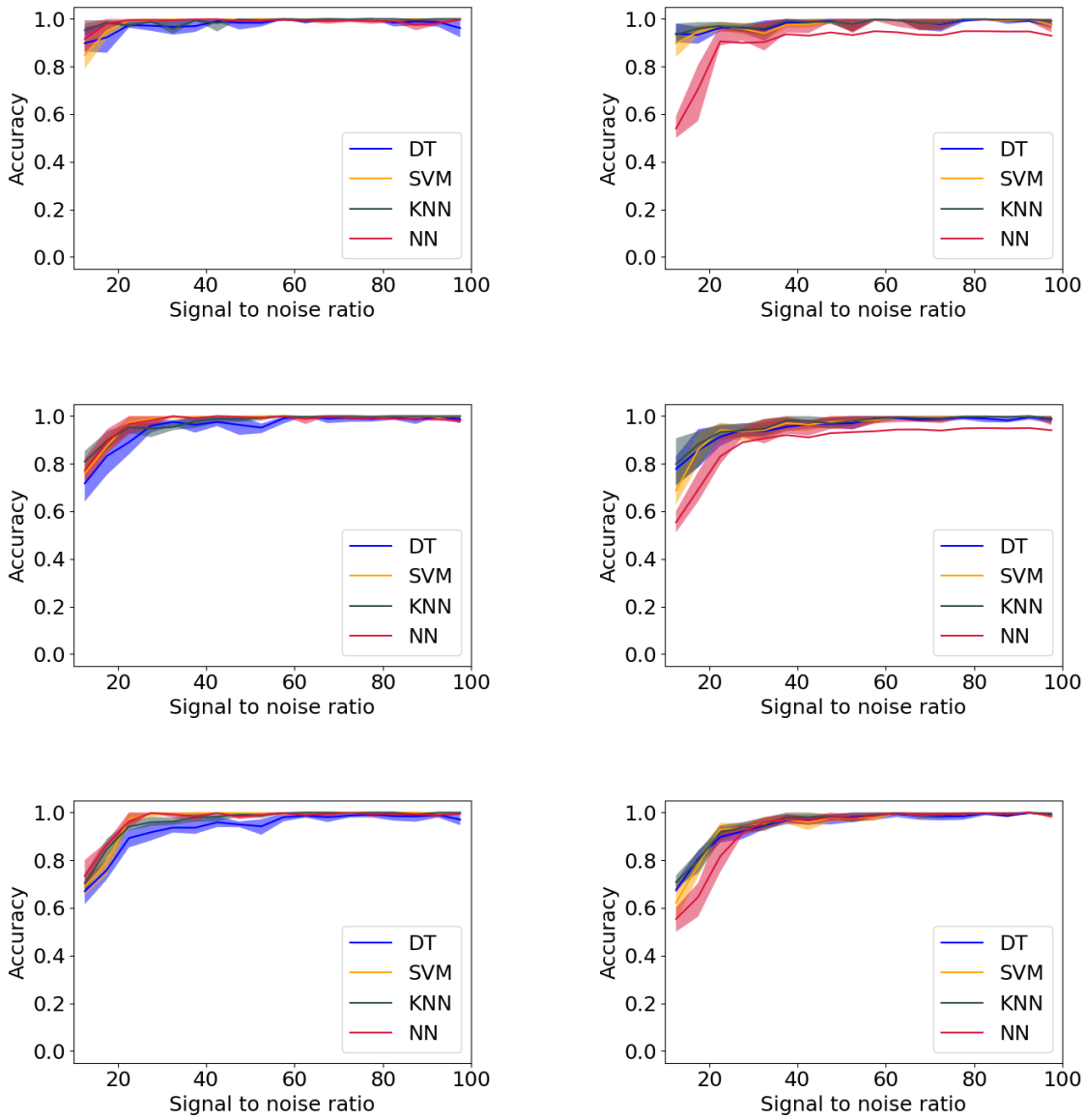

Figure 3.7: Improved balanced accuracy of the predictions by the classifier algorithms as a function of the SNR. The shaded areas are the 1σ-ranges over 10 realizations. On the left are the results for PCA and on the right for UMAP. The rows represent respectively the results for the explosion mechanism, presence of prompt convection and SASI.

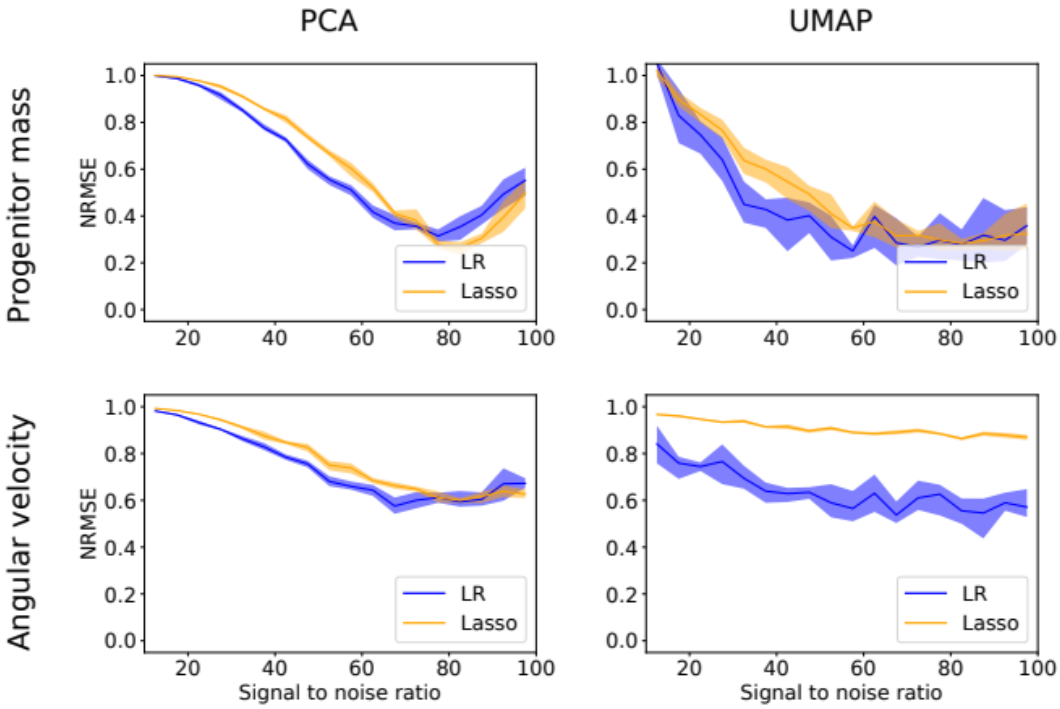

Figure 3.8: Original normalized root mean square error of the predictions by the regression algorithms as a function of the SNR. The shaded areas are the 1σ-ranges over 10 realizations. On the left are the results for PCA and on the right for UMAP. The rows represent respectively the results for the mass and angular velocity (from Dálya et al. [18]).

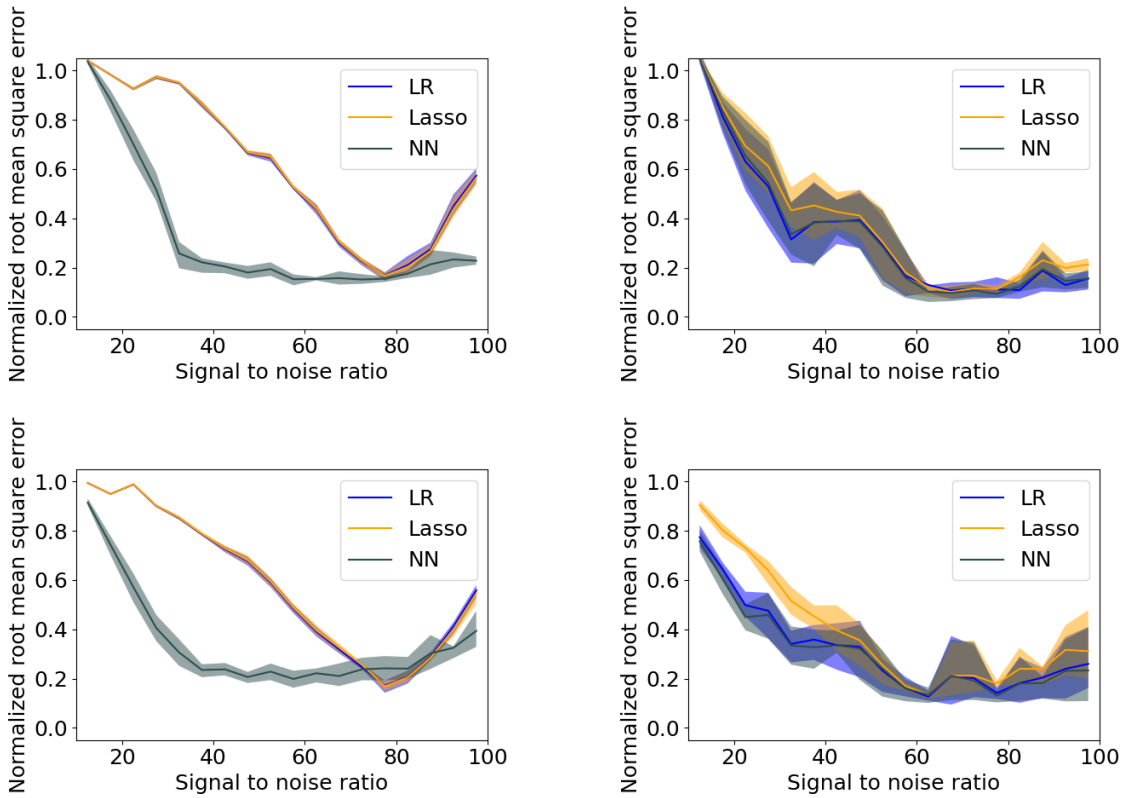

Figure 3.9: Improved normalized root mean square error of the predictions by the regression algorithms as a function of the SNR. The shaded areas are the 1σ-ranges over 10 realizations. On the left are the results for PCA and on the right for UMAP. The rows represent respectively the results for the mass and angular velocity.

Applicability in practice

The previous results show promising results in the applicability of machine learning in determining the properties of supernova explosions. There is one major possible problem that has to be kept in mind though. Now the algorithms are trained and tested on all the models at hand, but when a supernova event passes the detector, it could have another gravitational wave signature than the supernova models used for training. The machine learning algorithms thus should be able to interpolate the correct class from a different training set of supernova models. Especially in the classification algorithms, the previous results could be very misleading as when we look at the parameter space on figure 3.10, different models can form distinct islands even in lower dimensional space. Interpolating is then much more complicated.

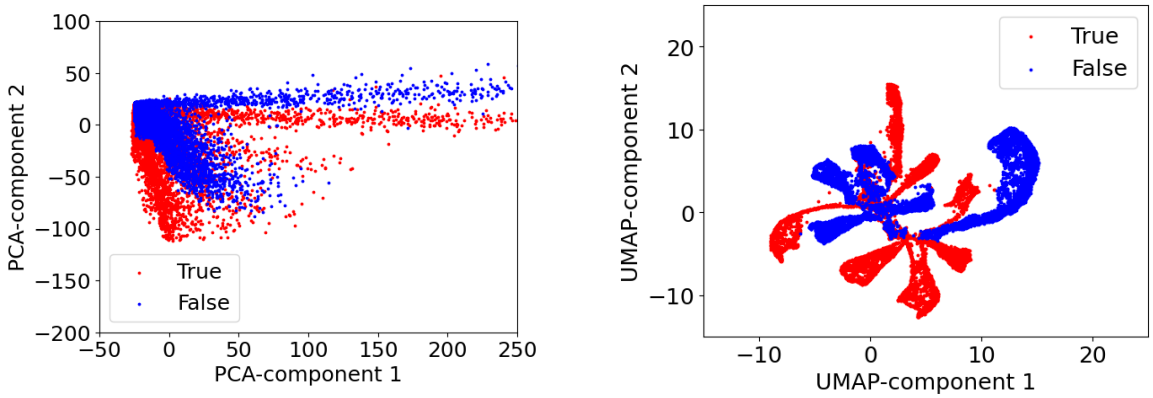

Figure 3.10: Representation of the entire dataset after reduction in dimensionality. The first two components after reduction by respectively the PCA and UMAP are plotted. The dataset is divided in two parts corresponding to the presence of prompt convection.

To test the applicability of the use of machine learning in practice, one model was completely left out of training and was completely used as test data. This was done for every used supernova waveform family[7]. A first test result can be found on figure 3.11 for the presence of prompt convection using PCA. Note that the shaded areas are now the 1σ-bounds of the 14 realizations of leaving out every supernova model once. The results are not good at all. The mean balanced accuracy is even below 0.50 and the variation is immense. The latter can be easily explained. As the models are on distinct islands, the accuracy of leaving one model out is an all-or-nothing prediction. It is just the question of whether the interpolation up to that specific island is correct or not. To prove this theory, also the balanced accuracy of one left-out model alone was determined. Two of them are seen on figure 3.12. For low SNRs, the algorithms are taking guesses, but for SNRs of 20 and higher, it is mostly a case of all-or-nothing, leading to very large standard deviations. For this reason, the 1σ-bounds will not be shown in the remainder of this thesis.

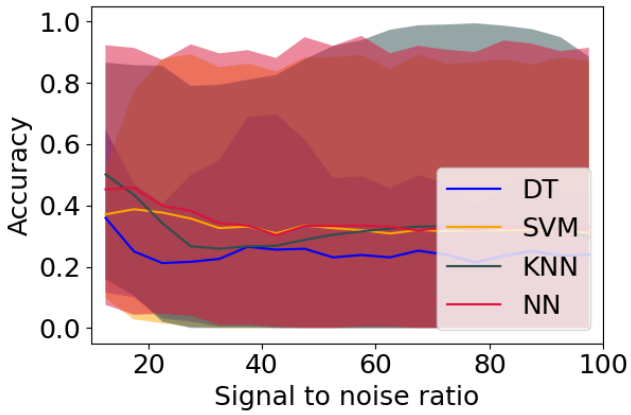

Figure 3.11: Balanced accuracy of the predictions of the classification algorithms on the presence of prompt convection using PCA when leaving out one model.

[7]Thus the waveform families of which multiple supernova models were used, were left out completely.

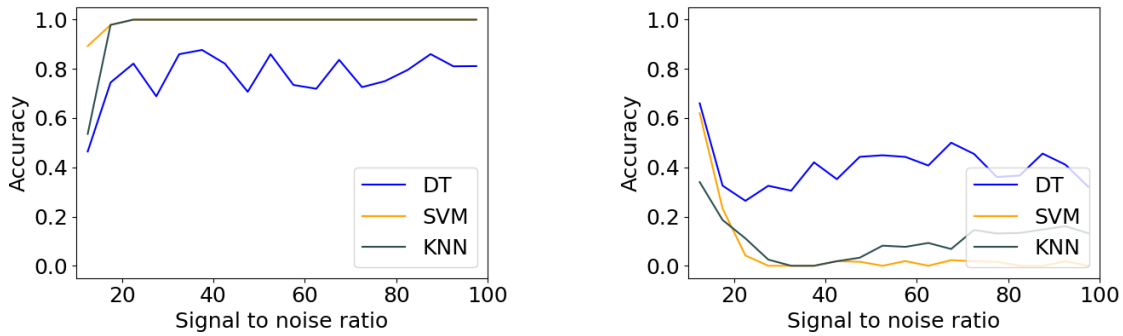

Figure 3.12: Balanced accuracy of the predictions of the classification algorithms on the presence of prompt convection using PCA when leaving the Andresen '19 m15fr and Andresen '17 s11 models out.

Now the question at hand is whether the hyperparameters can be tuned in such a way, that the predictions for an unknown model make sense. To do so, a similar approach to the previous part is performed with one key difference. The successive halving method is very accurate but also very time-consuming. Therefore it was opted to take another approach. For every relevant hyperparameter, a time-and-performance plot was made to determine its value. An example of the effect of the number of remaining dimensions after UMAP on the prediction of the mass by neural networks can be seen on figure 3.13. Note that now the definition of the NRMSE is slightly different from before. As we now test the algorithms with just one model left out, the standard deviation of the test data is 0, leading to infinity in the NRMSE. Therefore, the RMSE is normalized by dividing by the range of the supernova properties' values. Meaning dividing by 100 and 10 for respectively the progenitor mass and angular velocity.

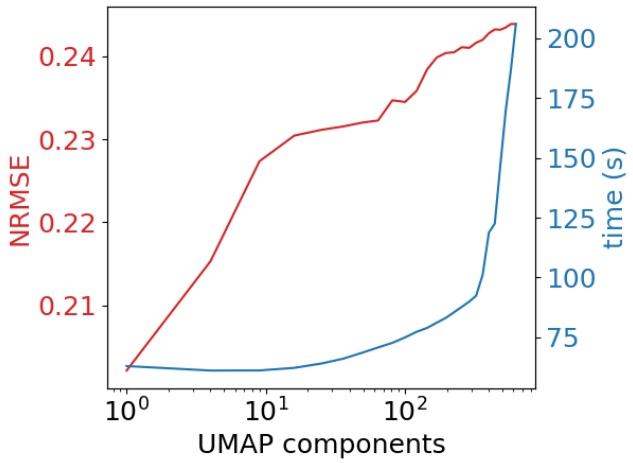

Figure 3.13: Time and normalized root mean square error of the predictions for the progenitor mass as a function of the number of remaining dimensions using UMAP and neural networks. For more than 300 components, the computational time increases exponentially. The lower the number of dimensions, the better the performance of the machine learning algorithms.

From figure 3.13, more can be learned. The performance of the machine learning algorithms increases when fewer dimensions are used. This effect was seen across all algorithms when predicting the angular velocity and using PCA instead of UMAP. This could be explained by overfitting, but with the best performance ones remaining only one dimension, one should be cautious. As only a limited set of different waveform families are available, it is possible that the most favorable outcome of the algorithm is just predicting a value that is close to the vast majority of the supernova models' values. Meaning that the algorithm didn't really learn anything from the physics behind the data. To see if the latter is the case, a dive under the hood was taken and the specific predictions made by the algorithms were plotted on histograms seen on figure 3.14, starting with just the predictions of the angular velocity. Once when the data was reduced to just one dimension and once when the data was reduced to 100 dimensions.
From table 1, one can see that the majority of supernova models have a angular velocity between 0 and 1 rad/s. The distribution should be peaked around those values and span the whole range of injected rotational velocities when the algorithm has learned anything from the data. This is not the case when the data is reduced to just one dimension. For every algorithm, the predictions are indeed sharply peaked at low values, but the algorithm doesn't bother trying to predict higher rotational velocities. As the majority of initial values are very low, this ultimately results in a lower NRMSE. When going to 100 dimensions, both linear regression and neural networks show a larger variety of predictions, providing more meaningful results. LASSO still has its sharp peak at around one rad/s and has hardly predicted any model to be non-rotating, while these include the majority of used supernova models. Linear regression and neural networks do predict a reasonable amount of non-rotating supernovae. The latter two algorithms seem to be more trustworthy in practice. A similar conclusion can be made when predicting the progenitor mass (figure 3.15) although less pronounced. Reducing the data with PCA also resulted in similar effects.

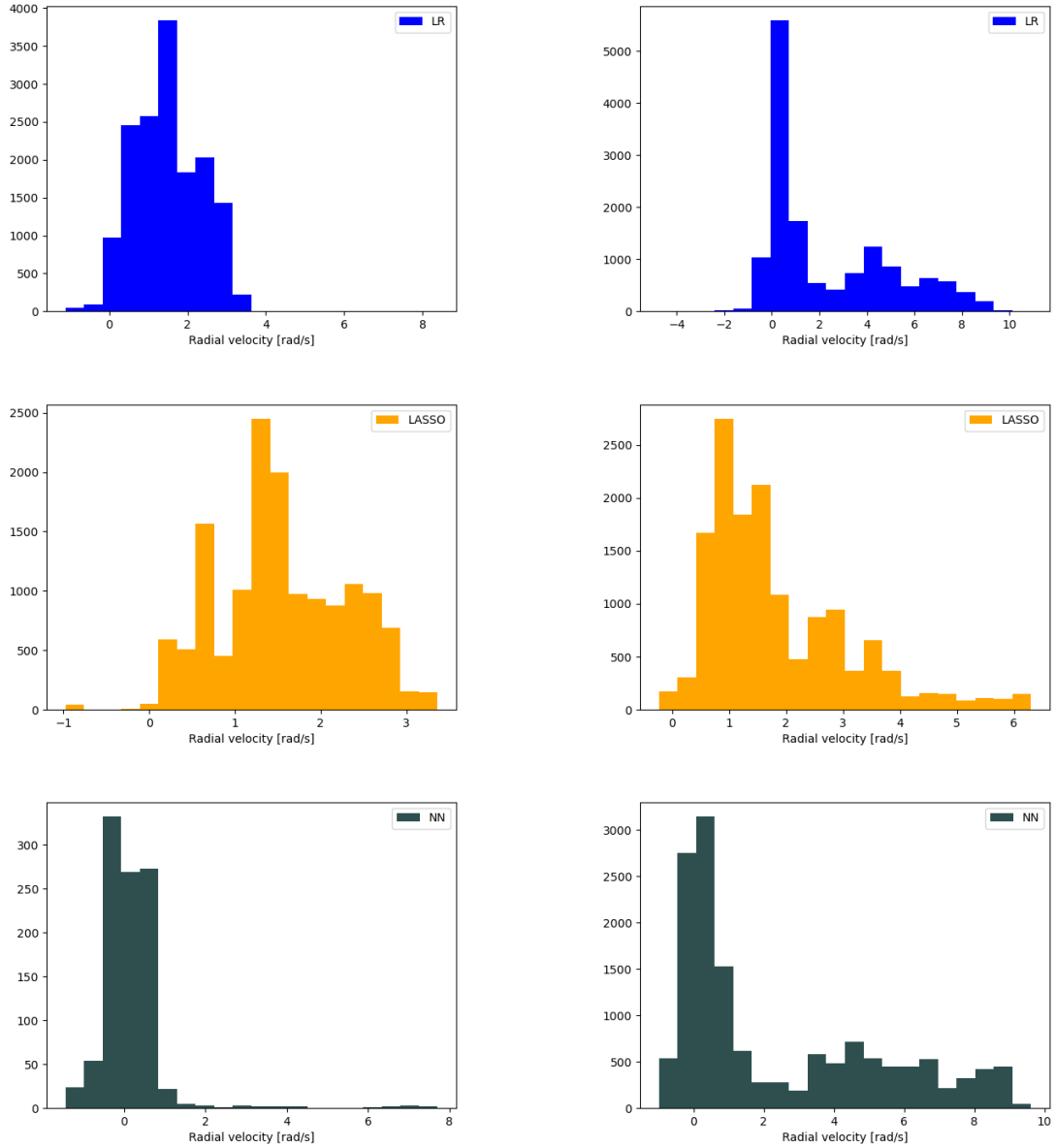

Figure 3.14: Distribution of the predictions made by the machine learning algorithms on the angular velocity while reducing the data with UMAP. On the left is the data reduced to one dimension and on the right to 100 dimensions. Only when using enough dimensions, the results of the machine learning algorithms are trustworthy.

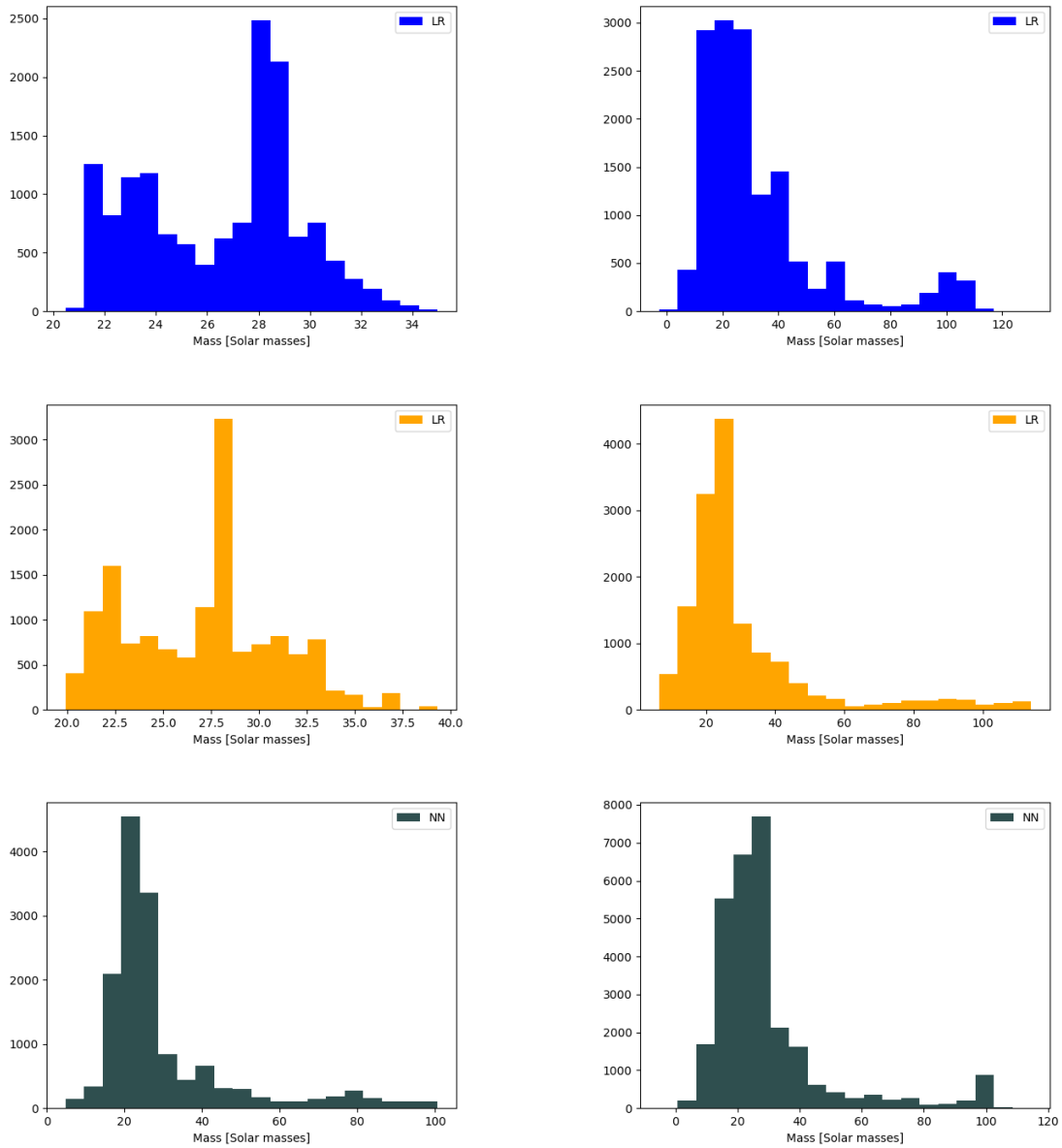

Figure 3.15: Distribution of the predictions made by the machine learning algorithms on the progenitor mass while reducing the data with UMAP. On the left is the data reduced to one dimension and on the right to 100 dimensions. Only when using enough dimensions, the results of the machine learning algorithms are trustworthy.

For this reason, the number of dimensions the data was transformed into was chosen as 100. The other relevant hyperparameters can be found in Appendix C. The performances of the machine learning algorithms on the angular velocity and progenitor mass are seen on figure 3.16. The most trustworthy algorithms perform worse than LASSO in general. This is a disappointing result as the predictions made by them represented the dataset. They were just wrong. When looking at only neural networks and linear regression, the combination PCA+LR performs the best for low SNRs, while the combination UMAP+ LR/NN performs the best for high SNRs. Note that the new definition of the NRMSE is more forgiving than the previous definition. These values thus can't be compared with earlier results but are definitely worse. The main issue is the limited amount of supernova models. A research with more models is needed to give a definite answer to whether machine learning can be used to infer the properties of a supernova accurately.

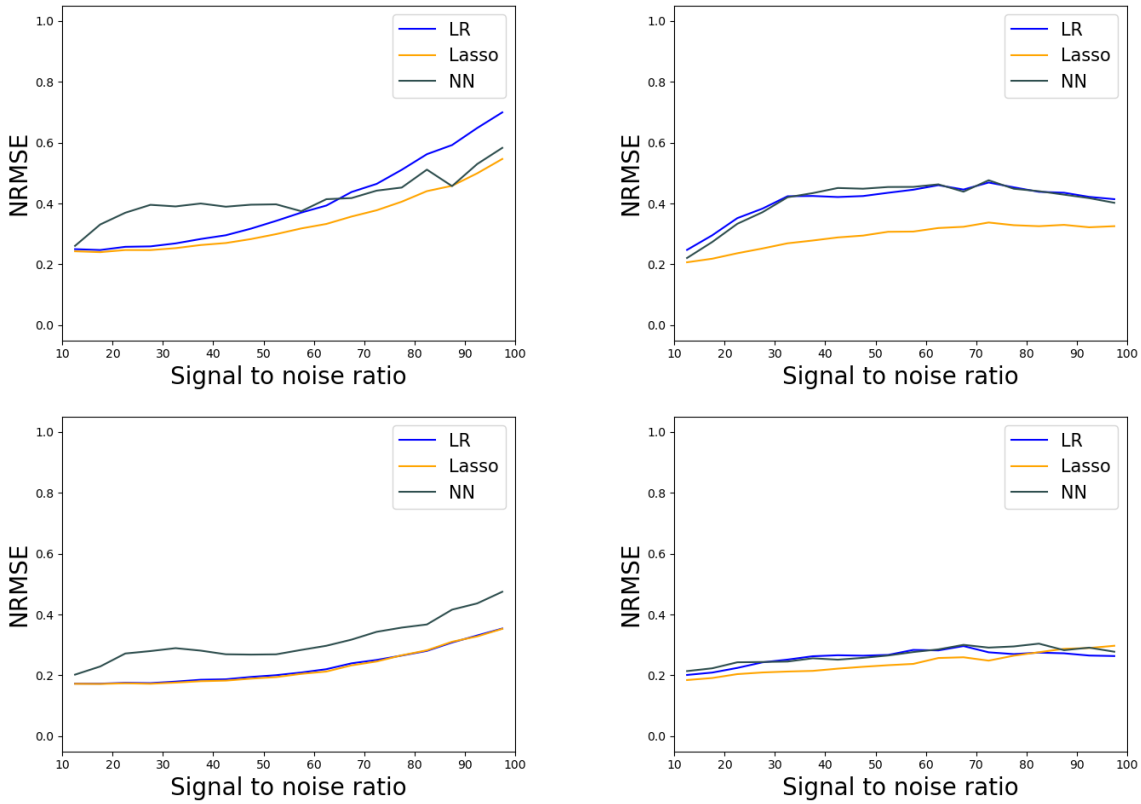

Figure 3.16: Normalized root mean square error of the predictions by the regression algorithms when leaving one model out as a function of the SNR. On the left are the results for PCA and on the right for UMAP. On top the results when predicting the angular velocity, on the bottom the progenitor mass.

Now moving on to the classifier algorithms, the same problem could occur where the algorithms perform better while not recognizing the physics behind the waveforms. There is one fundamental difference though. Even when reducing the data to only two dimensions already some clustering can be seen in the presence of prompt convection (figure 3.10), especially when reducing the data using PCA. Also when plotting the balanced accuracy of the algorithms for various properties, for both PCA and UMAP, the peak in performance isn't at the reduction of the data to 1 dimension (see for example figure 3.17). So for the classification algorithms, the number of dimensions will be chosen to be the best-performing one. These vary from 2 dimensions for neural networks to in some cases around 100 for support vector machines. The latter for which its kernel was switched into a polynomial kernel as the linear kernel was unsuitable for the non-linear distribution of models on the reduced space by UMAP (see figure 3.10). Other than the number of dimensions the data is reduced to, also a lot of effort was made into finding the ideal hyperparameters of the problem to determine whether machine learning can already be used with this limited set of supernova models. A short summary can be found in Appendix C. As there were only two MHD-models, it would be senseless to leave one of them out. The algorithms could just always predict a neutrino-driven explosion and produce seemingly good results. The explosion mechanism was thus not further researched. What was researched is the presence of rotation in the star to see whether the classification algorithms would perform better in determining whether the star was rotating or not.

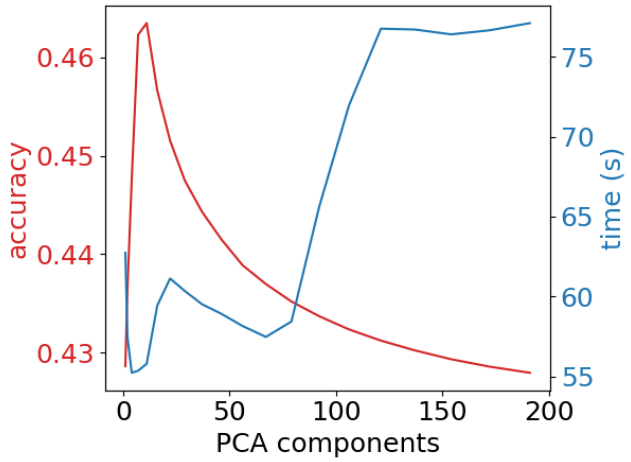

Figure 3.17: Time and balanced accuracy of the predictions for the presence of prompt convection as a function of the number of remaining dimensions using PCA and K nearest neighbors. The peak in performance of the algorithms are for dimensions higher than one.

The results of these fine-tuned parameters can be found on figure 3.18. The results look already a lot more promising than on figure 3.11. For every supernova characteristic a balanced accuracy above 0.50 could be achieved, meaning that the algorithm is better than just guessing the properties. But these results show that with the limited amount of supernova models to this date, machine learning cannot be used in practice to determine the properties of supernovae by their gravitational wave signatures. Again, a research with more supernova models is needed to determine its usefulness.

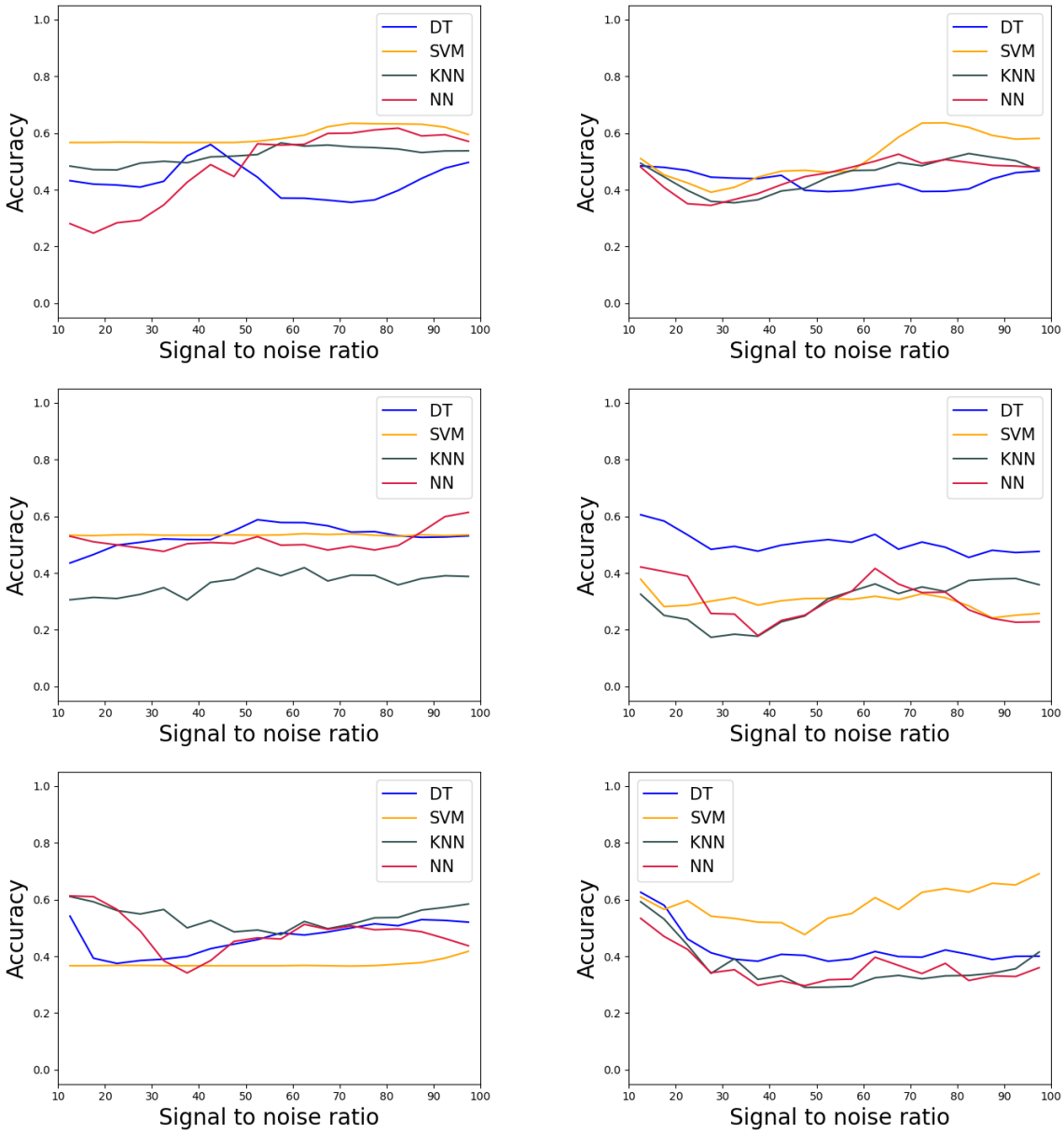

Figure 3.18: Balanced accuracy of the predictions by the classification algorithms when leaving one model out as a function of the SNR. On the left are the results for PCA and on the right for UMAP. From top to bottom, the performances for the predictions of the presence of prompt convection, SASI and rotation.

3.2 Distance project

Before the results of the detection efficiencies of different supernova models with ET are discussed, a test was done to see the difference in the performance of BayesWave for a different amount of Monte Carlo iteration steps as it lead to small inaccuracies for large SNRs before. To investigate that, 2000 waveforms of the Radice s9 model were generated and sent through BayesWave twice. Once with 400.000 iteration steps and the other time with 4.000.000 iteration steps. The detection efficiency as a function of the distance for both cases or seen on figure 3.19.
One can see that for near, or in other words: high SNR, supernova events the detection efficiency drops when one uses a lower amount of Monte Carlo iteration steps. Also for far-away events, there seem to be some small inaccuracies as one does not suspect the signal to be visible after a certain threshold. A higher amount of iteration steps partially resolves this issue. Other than that the results of both cases are similar. There is not a noteworthy difference in the distance on the 90, 50 or 10 % detection efficiency mark. This combined with the running time going from days to weeks, it was chosen to use 400.000 iteration steps for BayesWave from now on. Only when the graph was very obviously wrong at very high or low SNRs, these specific distances would be rerun with 4.000.000 iteration steps.

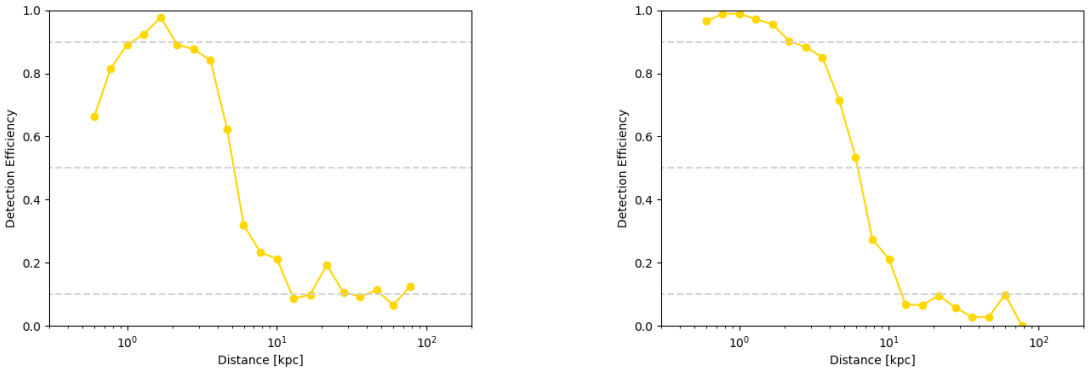

Figure 3.19: Detection efficiency measurement of the Radice s9 model for 400.000(left) and 4.000.000(right) iteration steps. At a high signal-to-noise ratio, one indeed needs more resources to fully converge to a steady distribution.

For every supernova model, the 2000 waveforms were divided into 20 evenly spaced bins on a logarithmic scale. The detection efficiencies of these models as a function of their distance are shown on figure 3.20. This detection efficiency can be interpreted as the probability that BayesWave would be able to detect that model at a certain distance from the detector.

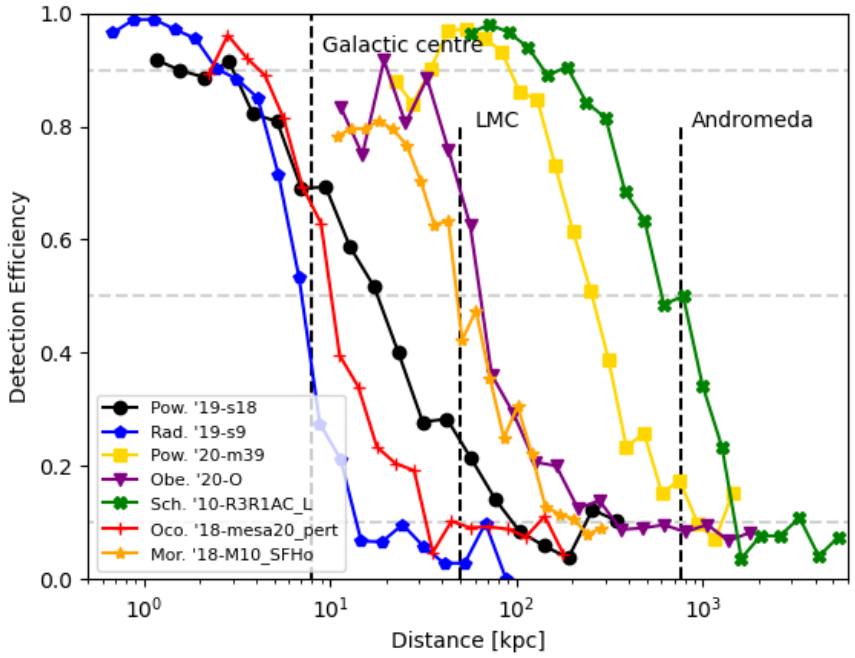

Figure 3.20: Detection efficiency for seven supernova models as a function of the distance from the source using BayesWave as the searching algorithm. The gray dashed lines are for reference the 10%, 50% and 90% detection efficiencies. Also the distances of the Galactic center, the Large Magellanic Cloud and Andromeda to us are plotted.

One can see that the most energetic model can be seen of the furthest distance, just as previous studies have shown with LIGO O1-O2 [3] and O4-O5 [51] this is a model with an MHD-driven explosion mechanism. What really is encouraging is that the most energetic supernovae could in principle even be detected from the Andromeda galaxy. In comparison with LIGO, also supernovae with moderate energies can be seen from the Large Magellanic Cloud, which hosted SN 1987A[8]. Furthermore is every model visible from the galactic center. This means that the vast majority of core-collapse supernovae in the Milky Way would be detectable as the density and star formation rate there is significantly higher than anywhere else in the galaxy [4, 20]. Most supernovae are thus expected to occur in the galaxy's center.
The detection efficiency of the Morozova models did not come close to reaching 100% detection efficiency at high signal-to-noise ratios with the cWB algorithm. Instead, its maximal detection efficiency lies around 65%. BayesWave can improve this as even with lower resources it was able to get a maximal detection efficiency of around 80%
A non-linear least squares fit was performed to truly compare the performance of ET compared to LIGO on the matter. As can be seen from figure 3.20, the detection efficiency curve has the form of an inverted sigmoid function:

$$f(x) = \frac{a}{1 + e^{c(x-b)}} \tag{3.2.1}$$

[8]SN 1987A was the first and only supernova thus far with not only detected electromagnetic observations but also a neutrino signal was received [7]

Using that fit, the values for the 10, 50 and 90% detection efficiency distances were calculated. These are shown in table 3. On figure 3.21, these values are then compared with LIGO O3 and O5 from Szczepańczyk et al. [51] and Abbott et al. [4], which were determined using the cWB search algorithm. The Obergaulinger model wasn't included in the former so no LIGO O5 information is known. The same is true for the Morozova and Scheidegger models with LIGO O3.

Waveform family	Waveform Identifier	ET det. range [kpc]		
		90 %	50 %	10 %
Morozova et al. 2018 [32]	M10_SFHo	NaN	63.6	186
Obergaulinger et al. 2020 [37]	O	NaN	78.1	219
O'Connor et al. 2018 [36]	mesa20_pert	3.6	10.3	35.4
Powell and Müller et al. 2019 [41]	s18	2.0	18.2	112
Powell and Müller et al. 2020 [42]	m39	71.1	280	764
Radice et al. 2019 [44]	s9	3.2	7.1	14.5
Scheidegger et al. 2010 [47]	R3E1AC_L	178	666	2130

Table 3: Values for the 10, 50 and 90% detection efficiency distances for the seven supernova models. These values were calculated by doing a non-linear least squares fit on an inverted sigmoid function.

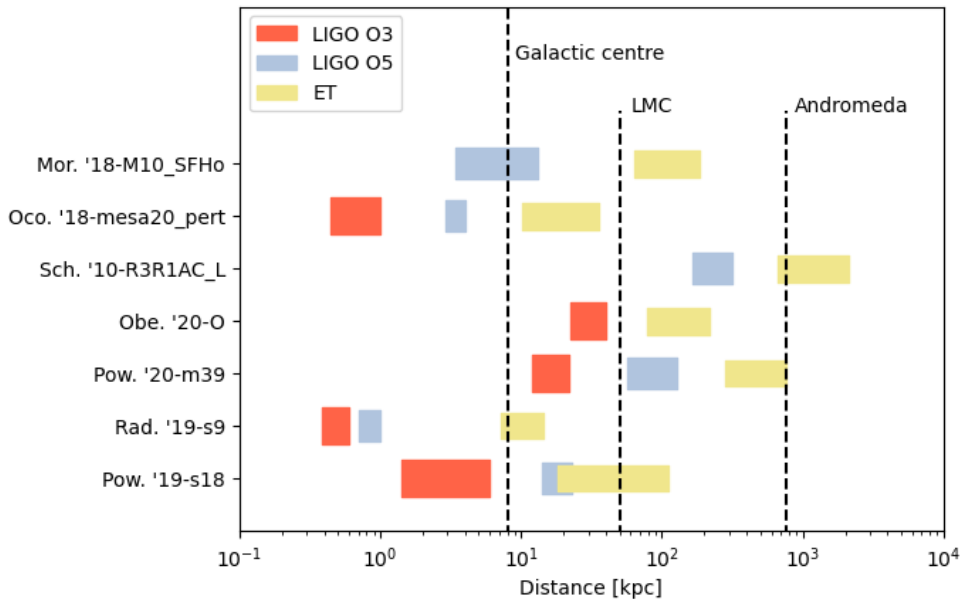

Figure 3.21: Comparison in detection efficiency between the Einstein Telescope and O3 and O5 of LIGO. The edges of the boxes are respectively the values where 90% and 50% of waveforms were seen as a signal. For most supernova models, the detection efficiencies improve by around one order of magnitude with respect to the current gravitational wave observatories.

Across all supernova models, the performance of ET is significantly better than LIGO up to around an order of magnitude on average. Again, even the lightest progenitors produce a supernova that is visible from the galactic center. Other than that is the range between the 10 and 50% detection efficiencies using BayesWave larger than with coherent WaveBurst. This effect is the largest in the Powell s18 model. A study on the detection efficiencies of using solely cWB should be done to determine the best search algorithm. Lastly, the results from the distance project will be part of the Blue Book of the Einstein Telescope Observational Science Board, which outlines the scientific rational behind building the telescope.

4 Future work

In this section, some potential improvements on the results or future work are presented. The first three subsections involve improvements on the machine learning algorithms. The last two involve future work related to the distance project.

4.1 New models

As mentioned earlier, the applicability of machine learning in practice can only be proven when more supernova models are accessible. At the time of writing this thesis, 16 different waveform families are made, but many more are necessary to fill the parameter space sufficiently. This way also the potential bias in the training set could be prevented. Especially MHD-driven models and models with a high mass and/or high angular velocity. Ideally, also every equation of state and approximations to general relativity and neutrinos should be represented with enough models. The choices of these also influence the outgoing gravitational wave. Lastly, real detections could be used as training data and would give helpful insights. For that, we will have to wait until the next galactic supernova of which Betelgeuse seems to be a good candidate. According to Saio et al. [48], she reached the carbon-fusing stage and could go supernova within decades.

4.2 Other training data

As the models are all weighted to have the same SNR, no difference in the magnitude of the power spectral density is expected. This is crucial information that is lost to train the machine learning algorithms. Fang and Maeda [21] showed that the kinetic energy of supernovae is related to the progenitor mass with a power law. It is possible that the same is true for the emitted gravitational wave energy. So this information about the magnitude of the power spectrum seems to be really important. A better way to predict the progenitor mass might thus be by not scaling to SNR but by training and testing the data as a function of the distance from the detector. Note that then, the performance of the algorithms as a function of the distance is detector-depended.
For other characteristics, such as the presence of prompt convection or SASI, the emitted gravitational wave energy might be of less importance. One way of improving these algorithms is by training them within their specific SNR-bin, as that way all the waveforms have around the same order of magnitude in the power spectrum. Whether the algorithms are trained with or without SNR-weighting, one could expect that the performance of the algorithms would improve when they are trained with only data within the same SNR- or distance-bin instead of training across the whole range of SNRs or distances.

4.3 Multilabel predictions

Another way to possibly improve the performance of the machine learning algorithm is multilabel classification or regression. This effectively means determining multiple characteristics of the supernovae at the same time. This way, the machine learning algorithms could search for links across different properties. We already know that some links exist. As mentioned in the introduction, MHD-driven supernovae are only possible with high enough rotational velocity. The machine learning algorithms might even be able to find new links that are not known yet.

4.4 Detection efficiency with other search algorithms

BayesWave is not the only algorithm capable of dealing with gravitational wave bursts. Another popular algorithm is coherent WaveBurst (cWB). When writing the Blue Book of ET, it would be very interesting to see how much the results from cWB differ from the obtained results using BayesWave. A similar study using cwB, like the one conducted in this thesis, is suggested and is already underway.

4.5 Determination rate of detectable supernovae

With the information about the detection limit of different supernova models, the chances of detecting a certain supernova model in the Milky Way or Local Group can be calculated using models of the density and star formation rate of both. This has been done by Abbott et al. [4] for LIGO O3. It's interesting to see how ET would compare to those results. Ultimately also the rate of detectable supernovae with ET could be calculated from this knowledge.

5 Conclusion

In previous research, machine learning has proven its worth in determining the properties of supernovae from solely gravitational wave data. Here, the effect of the slightly different morphologies of the sensitivity curves of the LVK-network and ET are shown to be negligible on the performance of the machine learning algorithms compared to the systematic errors from the latter. It was shown that the performances of the machine learning algorithms heavily rely on the chosen hyperparameters. With a good choice of hyperparameters, the machine learning classification algorithms all perform almost perfectly for SNR above 30 on known waveforms. Also, the regression algorithms perform significantly better with a more tuned set of hyperparameters. Neural networks only appear to be a valuable addition in predicting continuous characteristics such as progenitor mass and angular velocity.

In practice, we don't know whether a supernova event would emit the same waveform as one from the supernova model set. To test the applicability in practice, one supernova model was excluded from training and only used as test data. The performances of the machine learning algorithms, even after tuning the hyperparameters, weren't particularly positive. It was concluded that more supernova models are necessary for training data to determine whether machine learning really is useful in practice.

The detection limits in observing different supernova models are around one magnitude further with ET than LIGO using BayesWave. Even the lowest energy supernovae are now visible up to the galactic center, where the chances of supernovae are the highest. The most energetic supernovae could in principle even be visible from Andromeda.

References

[1] J. Aasi et al. "GRAVITATIONAL WAVES FROM KNOWN PULSARS: RESULTS FROM THE INITIAL DETECTOR ERA". In: *The Astrophysical Journal* 785.2 (Apr. 2014), p. 119. DOI: 10.1088/0004-637x/785/2/119. URL: https://doi.org/10.1088%2F0004-637x%2F785%2F2%2F119.

[2] B. P. Abbott et al. "Observation of Gravitational Waves from a Binary Black Hole Merger". In: *Physical Review Letters* 116.6 (Feb. 2016). DOI: 10.1103/physrevlett.116.061102.

[3] B. P. Abbott et al. "Optically targeted search for gravitational waves emitted by core-collapse supernovae during the first and second observing runs of advanced LIGO and advanced Virgo". In: *Physical Review D* 101.8 (Apr. 2020). DOI: 10.1103/physrevd.101.084002.

[4] R. Abbott et al. "All-sky search for short gravitational-wave bursts in the third Advanced LIGO and Advanced Virgo run". In: *Physical Review D* 104.12 (Dec. 2021). DOI: 10.1103/physrevd.104.122004.

[5] R. Abbott et al. "Search for intermediate-mass black hole binaries in the third observing run of Advanced LIGO and Advanced Virgo". In: *Astronomy &amp Astrophysics* 659 (Mar. 2022), A84. DOI: 10.1051/0004-6361/202141452.

[6] K. Van Akoleyen. *Relativiteitstheorie*. 2020.

[7] E.N. Alexeyev et al. "Detection of the neutrino signal from SN 1987A in the LMC using the INR Baksan underground scintillation telescope". In: *Physics Letters B* 205.2 (1988), pp. 209–214. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(88)91651-6.

[8] H Andresen et al. "Gravitational waves from 3D core-collapse supernova models: The impact of moderate progenitor rotation". In: *Monthly Notices of the Royal Astronomical Society* 486.2 (Apr. 2019), pp. 2238–2253. ISSN: 0035-8711. DOI: 10.1093/mnras/stz990.

[9] H. Andresen et al. "Gravitational wave signals from 3D neutrino hydrodynamics simulations of core-collapse supernovae". In: *Monthly Notices of the Royal Astronomical Society* 468.2032–2051 (2017). DOI: 10.1093/mnras/stx618.

[10] Z. Arzoumanian et al. "The NANOGrav 12.5 yr Data Set: Search for an Isotropic Stochastic Gravitational-wave Background". In: *The Astrophysical Journal Letters* 905.2 (Dec. 2020), p. L34. DOI: 10.3847/2041-8213/abd401.

[11] H. A. Bethe and J. R. Wilson. "Revival of a stalled supernova shock by neutrino heating". In: 295 (Aug. 1985), pp. 14–23. DOI: 10.1086/163343.

[12] G. S. Bisnovatyi-Kogan. "The Explosion of a Rotating Star As a Supernova Mechanism." In: 14 (Feb. 1971), p. 652.

[13] J. M. Blondin, A. Mezzacappa, and C. DeMarino. "Stability of Standing Accretion Shocks, with an Eye toward Core-Collapse Supernovae". In: 584.2 (Feb. 2003), pp. 971–980. DOI: 10.1086/345812.

[14] S. W. Bruenn and A. Mezzacappa. "Prompt Convection in Core Collapse Supernovae". In: 433 (Sept. 1994), p. L45. DOI: 10.1086/187544.

[15] M. Bugli, J. Guilet, and M. Obergaulinger. "Three-dimensional core-collapse supernovae with complex magnetic structures – I. Explosion dynamics". In: *Monthly Notices of the Royal Astronomical Society* 507.1 (July 2021), pp. 443–454. ISSN: 0035-8711. DOI: 10.1093/mnras/stab2161.

[16] N. Christensen. "Stochastic gravitational wave backgrounds". In: *Reports on Progress in Physics* 82.1 (Nov. 2018), p. 016903. DOI: 10.1088/1361-6633/aae6b5.

[17] N. J. Cornish and T. B. Littenberg. "Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches". In: *Classical and Quantum Gravity* 32.13 (June 2015), p. 135012. DOI: 10.1088/0264-9381/32/13/135012.

[18] G. Dálya et al. *Constraining Supernova Physics through Gravitational-Wave Observations.* 2023. arXiv: 2302.11480 [astro-ph.HE].

[19] H. Dimmelmeier et al. "Gravitational wave burst signal from core collapse of rotating stars". In: *Physical Review D* 78.6 (Sept. 2008). DOI: 10.1103/physrevd.78.064056.

[20] D. Elia et al. "The Star Formation Rate of the Milky Way as Seen by Herschel". In: *The Astrophysical Journal* 941.2 (Dec. 2022), p. 162. DOI: 10.3847/1538-4357/aca27d.

[21] Q. Fang and K. Maeda. *Inferring the progenitor mass-kinetic energy relation of stripped-envelope core-collapse supernovae from nebular spectroscopy.* 2023.

[22] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning.* 2017.

[23] M.D. Kistler. W. C. Haxton and H. Yüksel. "TOMOGRAPHY OF MASSIVE STARS FROM CORE COLLAPSE TO SUPERNOVA SHOCK BREAKOUT". In: *Astrophys J.* 778(1).:81 (2013). DOI: 10.1088/0004-637X/778/1/81.

[24] A. Heger, S. E. Woosley, and H. C. Spruit. "Presupernova Evolution of Differentially Rotating Massive Stars Including Magnetic Fields". In: *The Astrophysical Journal* 626.1 (June 2005), pp. 350–363. DOI: 10.1086/429868.

[25] H.-T. Janka. "Explosion Mechanisms of Core-Collapse Supernovae". In: *Annual Review of Nuclear and Particle Science* 62(1).:407–451 (2012). DOI: 10.1146/annurev-nucl-102711-094901.

[26] T. Kuroda, K. Kotake, and T. Takiwaki. "A NEW GRAVITATIONAL-WAVE SIGNATURE FROM STANDING ACCRETION SHOCK INSTABILITY IN SUPERNOVAE". In: *The Astrophysical Journal Letters* 829.1 (Sept. 2016), p. L14. DOI: 10.3847/2041-8205/829/1/L14.

[27] T. Kuroda et al. "Correlated Signatures of Gravitational-wave and Neutrino Emission in Three-dimensional General-relativistic Core-collapse Supernova Simulations". In: *The Astrophysical Journal* 851.1 (Dec. 2017), p. 62. DOI: 10.3847/1538-4357/aa988d.

[28] M. Maggiore. *Gravitational Waves Volume 1: Theory and Experiments.* Oxford University Press, 2008. ISBN: 978-0-19-857074-5.

[29] M. Maggiore. *Gravitational Waves Volume 2: Astrophysics and Cosmology.* Oxford University Press, 2018. ISBN: 978-0-19-857089-9.

[30] D. V. Martynov et al. "Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy". In: *Physical Review D* 93.11 (June 2016). DOI: 10.1103/physrevd.93.112004.

[31] L. McInnes, J. Healy, and J. Melville. *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. 2020. arXiv: 1802.03426 [stat.ML].

[32] V. Morozova et al. "The Gravitational Wave Signal from Core-collapse Supernovae". In: *The Astrophysical Journal* 861.1 (June 2018), p. 10. DOI: 10.3847/1538-4357/aac5f1.

[33] E. Muller et al. "Toward Gravitational Wave Signals from Realistic Core-Collapse Supernova Models". In: *The Astrophysical Journal* 603.1 (Mar. 2004), pp. 221–230. DOI: 10.1086/381360.

[34] E. Müller, H.-Th. Janka, and A. Wongwathanarat. "Parametrized 3D models of neutrino-driven supernova explosions". In: *Astronomy &amp Astrophysics* 537 (Jan. 2012), A63. DOI: 10.1051/0004-6361/201117611.

[35] J. W. Murphy, C. D. Ott, and A. Burrows. "A MODEL FOR GRAVITATIONAL WAVE EMISSION FROM NEUTRINO-DRIVEN CORE-COLLAPSE SUPERNOVAE". In: *The Astrophysical Journal* 707.2 (Dec. 2009), pp. 1173–1190. DOI: 10.1088/0004-637x/707/2/1173.

[36] E. P. O'Connor and S. M. Couch. "Exploring Fundamentally Three-dimensional Phenomena in High-fidelity Simulations of Core-collapse Supernovae". In: *The Astrophysical Journal* 865.2 (Sept. 2018), p. 81. DOI: 10.3847/1538-4357/aadcf7.

[37] M. Obergaulinger and M. Á . Aloy. "Magnetorotational core collapse of possible gamma-ray burst progenitors – IV. A wider range of progenitors". In: *Monthly Notices of the Royal Astronomical Society* 512.2 (Mar. 2022), pp. 2489–2507. DOI: 10.1093/mnras/stac613.

[38] C. D. Ott et al. "GENERAL-RELATIVISTIC SIMULATIONS OF THREE-DIMENSIONAL CORE-COLLAPSE SUPERNOVAE". In: *The Astrophysical Journal* 768.2 (Apr. 2013), p. 115. DOI: 10.1088/0004-637X/768/2/115.

[39] K.-C. Pan et al. "Stellar Mass Black Hole Formation and Multimessenger Signals from Three-dimensional Rotating Core-collapse Supernova Simulations". In: *The Astrophysical Journal* 914.2 (June 2021), p. 140. DOI: 10.3847/1538-4357/abfb05.

[40] A. L. Piro and Eric Pfahl. "Fragmentation of Collapsar Disks and the Production of Gravitational Waves". In: *The Astrophysical Journal* 658.2 (Apr. 2007), p. 1173. DOI: 10.1086/511672.

[41] J. Powell and B. Müller. "Gravitational wave emission from 3D explosion models of core-collapse supernovae with low and normal explosion energies". In: *Monthly Notices of the Royal Astronomical Society* 487.1 (May 2019), pp. 1178–1190. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1304.

[42] J. Powell and B. Müller. "Three-dimensional core-collapse supernova simulations of massive and rotating progenitors". In: *Monthly Notices of the Royal Astronomical Society* 494.4 (Apr. 2020), pp. 4665–4675. ISSN: 0035-8711. DOI: 10.1093/mnras/staa1048.

[43] J. Powell, B. Müller, and A. Heger. "The final core collapse of pulsational pair instability supernovae". In: *Monthly Notices of the Royal Astronomical Society* 503.2 (Mar. 2021), pp. 2108–2122. ISSN: 0035-8711. DOI: 10.1093/mnras/stab614.

[44] D. Radice et al. "Characterizing the Gravitational Wave Signal from Core-collapse Supernovae". In: *Astrophys J.* 876.L9 (2019). DOI: 10.3847/2041-8213/ab191a.

[45] S. N. Raza. *Reconstructing gravitational waves from core-collapse supernovae in Advanced LIGO-Virgo*. 2021. DOI: http://dx.doi.org/10.14288/1.0403714.

[46] S. N. Raza et al. "Prospects for reconstructing the gravitational-wave signals from core-collapse supernovae with Advanced LIGO-Virgo and the BayesWave algorithm". In: *Phys. Rev. D* 106 (6 Sept. 2022), p. 063014. DOI: 10.1103/PhysRevD.106.063014.

[47] S. Scheidegger et al. "The influence of model parameters on the prediction of gravitational wave signals from stellar core collapse". In: *A&A* 514 (2010), A51. DOI: 10.1051/0004-6361/200913220.

[48] H. Saio et al. *The evolutionary stage of Betelgeuse inferred from its pulsation periods*. 2023. arXiv: 2306.00287 [astro-ph.SR].

[49] P. R. Saulson. *Fundamentals of Interferometric Gravitational Wave Detectors*. 1994. DOI: 10.1142/2410.

[50] E. Sheldon. "An Introduction to Nuclear Astrophysics, by R.N. Boyd". In: *Contemporary Physics* 50.6 (2009), pp. 662–663. DOI: 10.1080/00107510902917079.

[51] M. J. Szczepańczyk et al. "Detecting and reconstructing gravitational waves from the next galactic core-collapse supernova in the advanced detector era". In: *Physical Review D* 104.10 (Nov. 2021). DOI: 10.1103/physrevd.104.102002.

[52] T. Takiwaki and K. Kotake. "GRAVITATIONAL WAVE SIGNATURES OF MAGNETOHYDRO DRIVEN CORE-COLLAPSE SUPERNOVA EXPLOSIONS". In: *The Astrophysical Journal* 743.1 (Nov. 2011), p. 30. DOI: 10.1088/0004-637x/743/1/30.

[53] A. Burrows. D. Vartanyan. "Core-collapse supernova explosion theory". In: *Nature* 589.29–39 (2021). DOI: 10.1038/s41586-020-03059-w.

[54] P. Frank Winkler and G. Gupta. "THE SN 1006 REMNANT: OPTICAL PROPER MOTIONS, DEEP IMAGING, DISTANCE, AND BRIGHTNESS AT MAXIMUM". In: *Astrophys J.* 585.324–335 (2002). DOI: 10.1086/345985.

[55] S. Zha et al. "Gravitational-Wave Signature of a First-Order Quantum Chromodynamics Phase Transition in Core-Collapse Supernovae". In: *Phys. Rev. Lett.* 125 (5 July 2020), p. 051102. DOI: 10.1103/PhysRevLett.125.051102.

A Detectie van Supernovae en hun eigenschappen met de Einstein Telescoop

Eerdere studies hebben aangetoond dat machine learning veel potentie heeft om de eigenschappen van supernova's te bepalen uit het gravitationele golf-signaal. In deze thesis werd onderzocht in welke mate deze oorspronkelijke resultaten verbeterd konden worden door een geschikte set hyperparameters te kiezen. Als training data werden 14 supernova modellen gebruikt waarvan elk 1000 supernova's gesimuleerd werden op een willekeurige locatie aan de hemel. Op deze golven werd dan ruis toegevoegd welke overeenkomt met de sensitiviteit van de Einstein Telescoop. Ten slotte werden deze signalen gereconstrueerd door BayesWave. Er werd aangetoond dat de verschillende morfologieën van LIGO en ET geen verschil maakten in de prestatie van de machine learning algoritmes. Verder werd bewezen dat de keuze van de set hyperparameters een grote invloed heeft op de prestatie van de algoritmes. Ten slotte werd er geconcludeerd dat men meer verschillende supernova modellen nodig heeft om deze methode in de praktijk toe te kunnen passen.

Uit de verbeterde sensitiviteit van ET in vergelijking met LIGO, kan men verwachten dat supernova's detecteerbaar zijn vanop een grotere afstand met ET. De detectie limiet van zeven verschillende supernova modellen werd bepaald door 2000 supernova's te simuleren van elk op een willekeurige positie aan de hemel en met een log-uniforme distributie in de afstand tot de detector. BayesWave classificeerde de signalen dan als signaal, glitch of ruis. Zelfs de laagst energetische supernova's bleken detecteerbaar vanaf het centrum van de Melkweg, waar de kans op supernova's is het grootst. De hoogst energetische supernova's kunnen in principe gedetecteerd worden vanuit de Andromeda galaxie.

A.1 Introductie

Wanneer de brandstof van sterren opgebruikt is, is er geen uitwaartse druk meer die de zwaartekracht kan tegenhouden, waardoor de ster ineen stort. Wanneer de ster zwaar genoeg was, is de ontaardingsdruk van de elektronen in de kern niet genoeg om deze instorting initieel tegen te houden en stort de ster verder in tot op het moment dat de ontaardingsdruk van de nucleonen een uitwaartse druk produceren. Deze plotse extra druk zorgt voor een schokgolf in de ster, waardoor de ster ontploft.
De energie van deze schokgolf blijkt echter niet genoeg te zijn om de explosie helemaal te voeden. Er moeten dus nog andere drijfveren zijn om deze explosie tot stand te brengen. Wetenschappers hebben meerdere potentiële effecten voorop gesteld om de supernova explosie te verklaren, maar geen enkele van deze kon bevestigd of ontkracht worden met observaties in het elektromagnetische spectrum. Elektromagnetische golven worden immers waar vrijgelaten uit de dense ster uren na de start van de supernova explosie, waardoor er via zo'n observaties geen informatie te verkrijgen is over de explosie zelf. Andere manieren om hierover toch informatie te vergaren is via neutrino's of gravitationele golven. Zij kunnen wel quasi vrij uit de ster ontsnappen vanaf het moment van de supernova. Op dit moment hebben de gravitationele golf-detectoren zoals LIGO, VIRGO en KAGRA nog geen supernova's kunnen waarnemen. Toekomstige detectors zoals de Einstein Telescoop kunnen hier verandering in brengen.

A.2 Methodes

Om te onderzoeken hoe machine learning de eigenschappen van supernova explosies kan voorspellen moeten we eerst een training dataset vooropstellen. Hiervoor werden voor 14 verschillende supernova families elk 1000 supernova's gesimuleerd. Op deze gravitationele golven werd dan ruis toegevoegd welke met de sensitiviteit van de Einstein Telescoop overeenkomt. Dit werd op zo'n manier gedaan dat de signalen elk een SNR hebben van tussen 10 en 100. BayesWave probeert dan deze gravitationele golven zo goed als mogelijk te reconstrueren. Dit werd gedaan om de resultaten zo dicht als mogelijk bij de realiteit te houden. Ten slotte werd als training data het gereconstrueerde vermogensspectrum gebruikt.

Deze data werd dan geschaald naar een gemiddelde van 0 en een variantie van 1. Verder werd de data gereduceerd in dimensies op twee verschillende manieren. Eenmaal met PCA en eenmaal met het non-lineaire UMAP. 90 procent van deze data werd dan gebruikt als training data en de rest als test data. In het onderzoek naar de bruikbaarheid in de realiteit werd 1 supernova familie gebruikt als test data (1000 signalen dus) en de rest als training data. Hiermee werden dan de machine learning algoritmes getraind. Er werd gekozen voor een variatie in simpelere en moeilijkere algoritmes om de data te trainen. De classificatie algoritmes die werden gebruikt zijn: Decision trees, support vector machines, k-nearest neighbor. De regressie algoritmes die gebruikt werden zijn: lineaire regressie en LASSO. Bovenop deze algoritmes werd ook het resultaat van neurale netwerken getest voor zowel de classificatie als regressie.

In het onderzoek naar de detectie limiet van supernova's met de Einstein Telescoop werden zeven verschillende supernova modellen onderzocht. Deze waren een variatie van laag- en hoogenergetische modellen. Van elk van deze modellen werden 2000 supernova's gesimuleerd met een willekeurige locatie aan de hemel en met een log-uniforme distributie van de afstand tot de detector. Deze signalen werden dan door BayesWave beoordeeld of deze inderdaad een signaal waren of een glitch of ruis.

A.3 Resultaten

Het effect van de verschillende morfologieën van LIGO en ET op de prestaties van de machine learning algoritmes zijn verwaarloosbaar gebleken tegenover andere systematische effecten. Verder was het duidelijk dat de resultaten van de algoritmes veel verbeterd konden worden door een goede keuze te maken van de gebruikte hyperparameters. Met de goede keuze van hyperparameters bleken alle algoritmes bijna perfect te presteren bij een SNR van boven de 30. Neurale netwerken bleken enkel een waardevolle bijdrage te leveren bij de continue eigenschappen zoals de massa van de oorspronkelijke ster en de angulaire snelheid. Er werd ook aangetoond dat machine learning in deze vorm niet bruikbaar is in de praktijk. Het grootste probleem is het gelimiteerde aantal aan bestaande supernova modellen. De bestaande modellen vullen de parameter ruimte allesbehalve volledig en zijn ook bevooroordeeld.

De detectie limieten van de verschillende supernova modellen waren ongeveer een magnitude beter met ET in vergelijking met LIGO. Zelfs de modellen met de laagste gravitationele golf-energie konden gedetecteerd worden vanuit het centrum van de Melkweg waar de kans op een supernova het hoogste is. De meest energetische supernova's zijn zelfs detecteerbaar vanuit de Andromeda galaxie. De bevindingen van dit deel van de thesis zullen gebruikt worden in het "Blue Book" van de Einstein Telescoop.

A.4 Toekomstig werk

Naast de implementatie van meer modellen zijn er ook nog andere manieren om de resultaten van de algoritmes te verbeteren. Een manier is door de algoritmes de kans te geven om informatie van de grootte van het vermogen te gebruiken in hun predicties. Hier werden alle supernova's geschaald naar SNR waardoor deze informatie verloren gaat. Voor sommige eigenschappen zoals de oorspronkelijke massa van de ster kan het beter zijn om de algoritmes te trainen op modellen van ongeveer dezelfde afstand als de geobserveerde supernova. Verder kan het ook nuttig zijn om de algoritmes meerdere eigenschappen te laten voorspellen op hetzelfde moment, omdat sommige eigenschappen gelinkt kunnen zijn aan elkaar.

In deze thesis werd de detectie limiet van de verschillende modellen bepaald met BayesWave. Er zijn ook andere algoritmes die in staat zijn om dit werk te doen wat voor andere resultaten kan zorgen. Een onderzoek met cWB is aangewezen en al onderweg. Ten slotte kan het ook interessant zijn om nu de mate van detecteerbare supernova's te voorspellen. Hiervoor moet er een accuraat model gemaakt worden van de kans op een supernova van de hele Lokale Groep.

A.5 Conclusie

In deze thesis werd de toepasbaarheid van machine learning in het bepalen van de eigenschappen van supernova's verder onderzocht. Er werd geconcludeerd dat correcte keuzes van hyperparameters een groot effect heeft en dat meer supernova modellen nodig zijn om deze methode in de realiteit te gebruiken.
Verder werd er aangetoond dat supernova's vanop een veel grotere afstand detecteerbaar zijn met ET in vergelijking met LIGO. Het feit dat vrijwel elke supernova detecteerbaar zou zijn vanuit het centrum van de Melkweg toont aan dat men met ET veel meer supernova's potentieel kan waarnemen.

B Supernova models

Andresen et al., 2017 [9] (s11)

The s11 model is a non-rotating model with the lightest progenitor mass of Andresen et al. [9] with 11.2 solar masses and one of the lightest model in general. It is a neutrino-driven supernova that is dominated by convection. This convection is not prompt convection though. Also, no SASI activity can be observed. Finally, non-resonant g-modes can be seen in the signal.

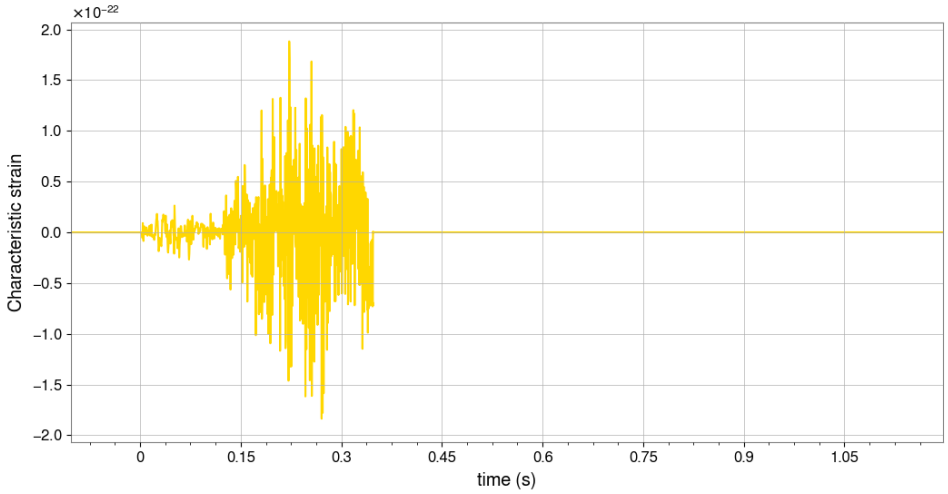

Figure B.1: Example waveform of the Andresen s11 model.

Andresen et al., 2019 [8] (m15fr)

Andresen et al. [8] made three simulations with a different progenitor angular velocity range, all with a progenitor mass of 15 solar masses. Here the fastest rotating one, which has a angular velocity of 0.5 rad/s, will be used. As a result of the fast rotational velocity, strong SASI and resonant g-modes are visible. Just as the previous model it is neutrino-driven.

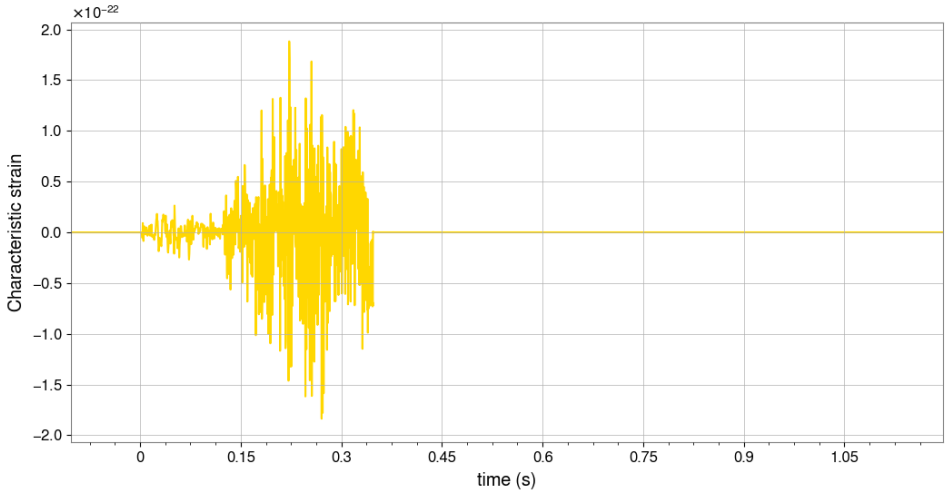

Figure B.2: Example waveform of the Andresen m15fr model.

Bugli et al., 2021 [15] (H)

This is the first non-neutrino-driven model in the list. It is an MHD-driven explosion with a very high angular velocity of 10 rad/s. Its progenitor mass is 35 solar masses and the signal does not contain any signs of convection or SASI.

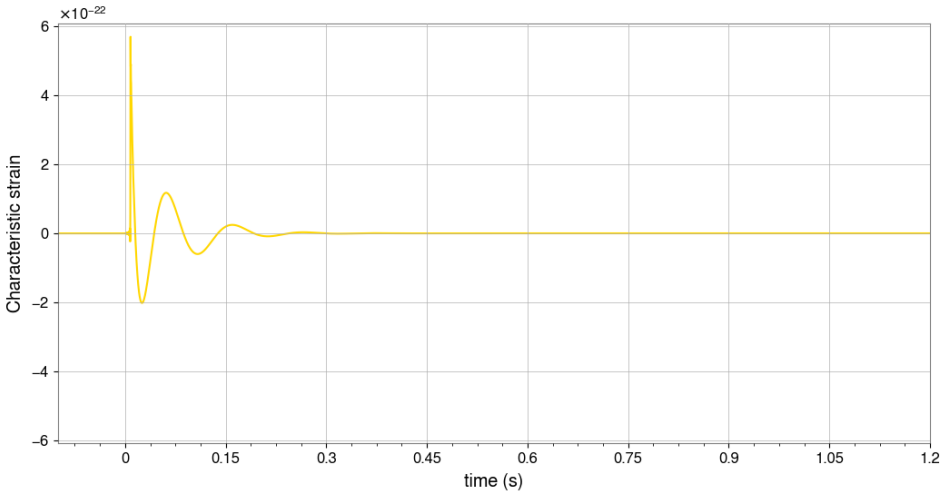

Figure B.3: Example waveform of the Bugli H model.

Kuroda et al., 2016 [26] (TM1)

This model has the same mass as the Andresen m15fr model, but it differs from it in two ways. The first one is that this model does not have any progenitor angular velocity and the second is a different equation of state. Kuroda et al. [26] have shown that the softer[9] the equation of state, the more pronounced the influence of SASI is in the gravitational wave signal. The TM1 model possesses a strong SASI and prompt convection contribution and is neutrino-driven. Also some g-modes were observed.

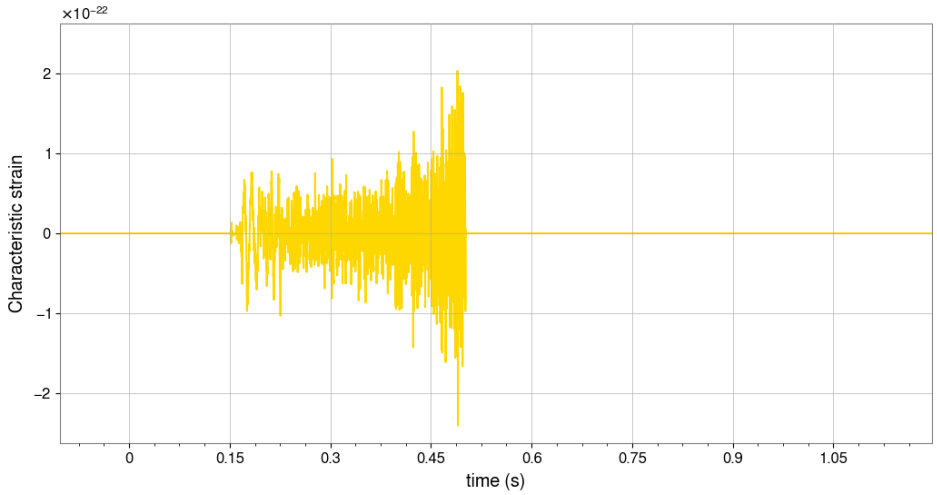

Figure B.4: Example waveform of the Kuroda TM1 model.

Kuroda et al., 2017 [27] (s11)

The same story yields for this model. It has a mass of 11.2 solar masses but a different equation of state than the Andresen s11 model. Its equation of state is softer and thus now SASI is visible. Other than that also prompt convection and g-modes were visible. Lastly it is also a neutrino-driven supernova model.

[9] A softer equation of state means that the pressure rises slowly with density.

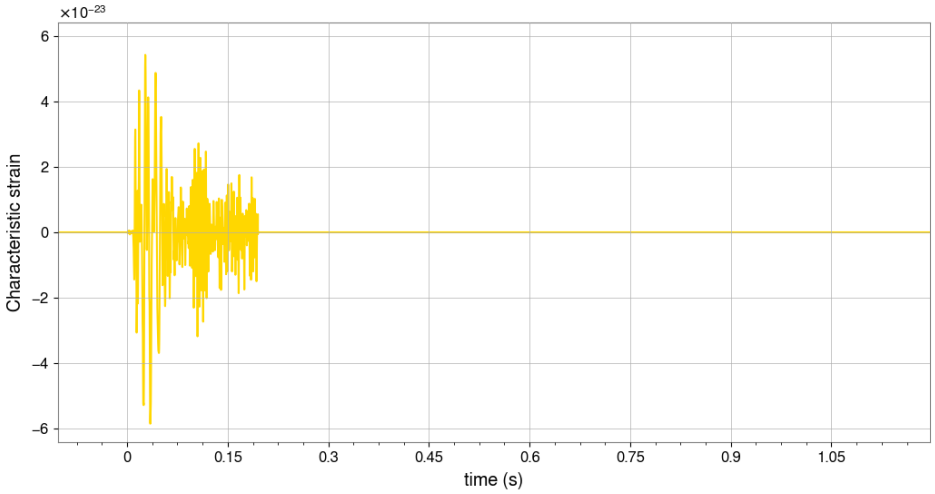

Figure B.5: Example waveform of the Kuroda s11 model.

Morozova et al., 2018 [32] (M10_SFHo)

Morozova et al. [32] explored the dependence on the gravitational wave signature of the equation of state and neutrino microphysics. They do so by simulating supernovae of different progenitor masses. The chosen model has a progenitor mass of 10 solar masses and a weak contribution of SASI. The non-rotating neutrino-driven model also shows signs of g-, f-, and p-modes. Its prompt convection contribution is negligible in the gravitational wave signal.

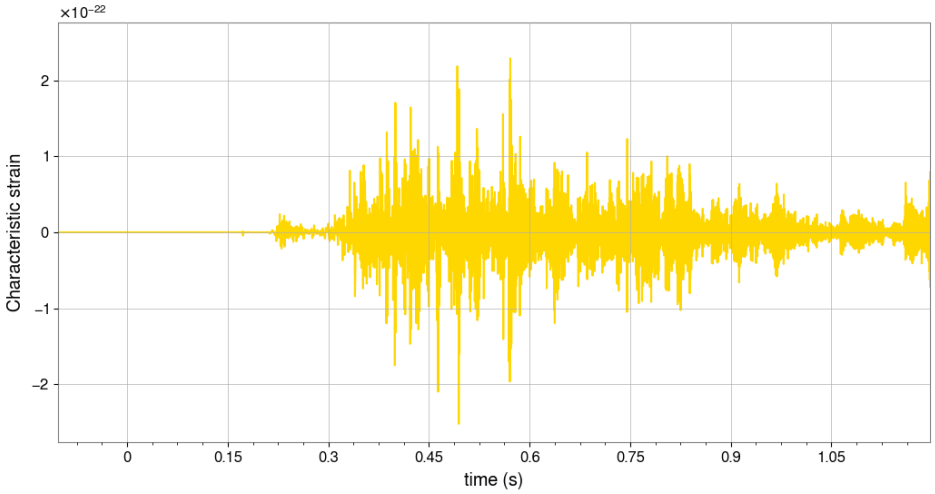

Figure B.6: Example waveform of the Morozova M10_SFHo model.

Müller et al., 2012 [34] (N20-2)

Also this model is a neutrino-driven model. It's a non-rotating model with a progenitor mass of 20 solar masses. The gravitational wave signal is dominated by low-frequency matter movement or in other words, SASI and convection is very dominant in the signal. Lastly, g-modes are also observable in the waveform.

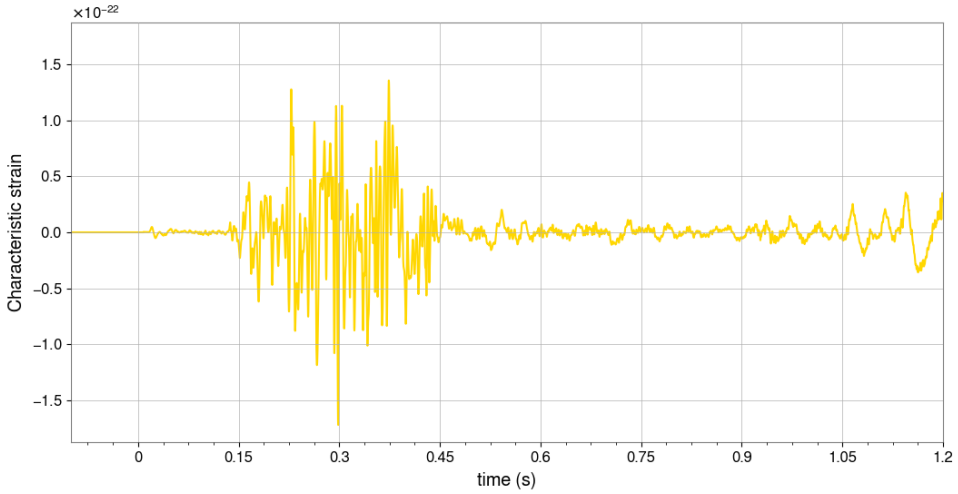

Figure B.7: Example waveform of the Müller N20-2 model.

Obergaulinger et al., 2020 [37] (O)

The Obergaulinger O model is one of the most high-energetic ones in this thesis. Its explosion is driven by strong magnetic fields and rapid rotation of 380 km/s. Its progenitor has a mass of 35 solar masses and possesses the right conditions for the bounce signal to be observable.

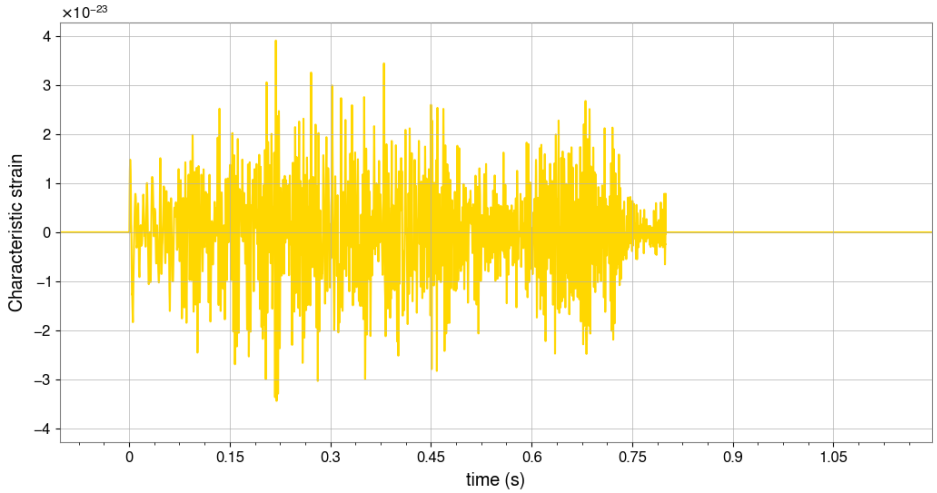

Figure B.8: Example waveform of the Obergaulinger O model.

O'Connor et al., 2018 [36] (Mesa20_pert)

The Mesa20_pert model is a neutrino-driven model with no initial angular velocity. However, some velocity perturbations are added on top of the silicon and oxygen fusing shell during the simulation. The gravitational wave signal is dominated by SASI activity and g-modes. No clear convection signals were found.

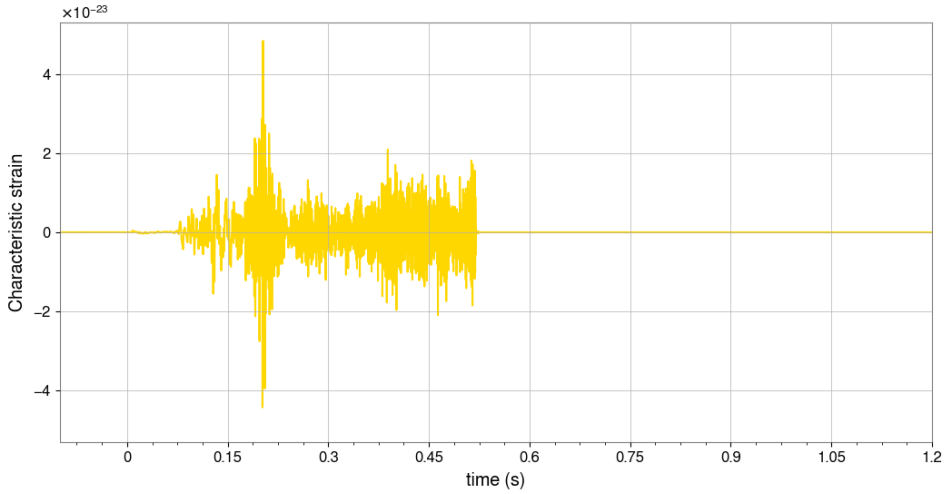

Figure B.9: Example waveform of the O'Connor Mesa20_pert model.

Ott et al., 2013 [38] (s27-fheat1.00)

Ott et al. research the post-core bounce phase of neutrino-driven supernovae. The specific model used here has a progenitor mass of 27 solar masses and has no angular velocity. The gravitational wave signal is dominated by prompt convection, but also SASI and neutrino-driven convection are visible. Finally also g-modes are visible in the waveform.

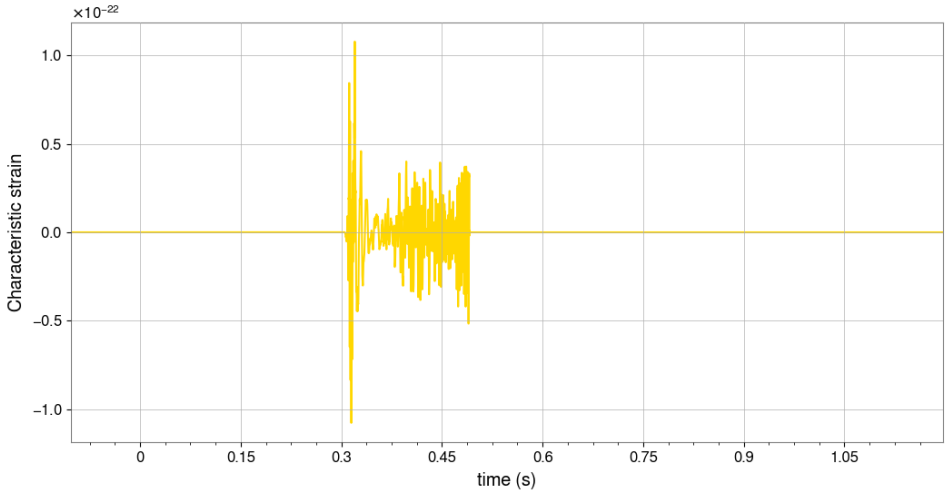

Figure B.10: Example waveform of the Ott s27-fheat1.00 model.

Pan et al., 2020 [39] (s40_fr, s40_sr)

Pan et al. [39] simulated one of the models with the highest progenitor mass of 40 solar masses with different radial velocities. In this thesis two models were chosen, one fast rotating one with a angular velocity of 1 rad/s and one with slower a angular velocity of 0.5 rad/s. Other than that, they both have signs of SASI, prompt convection and g-modes.

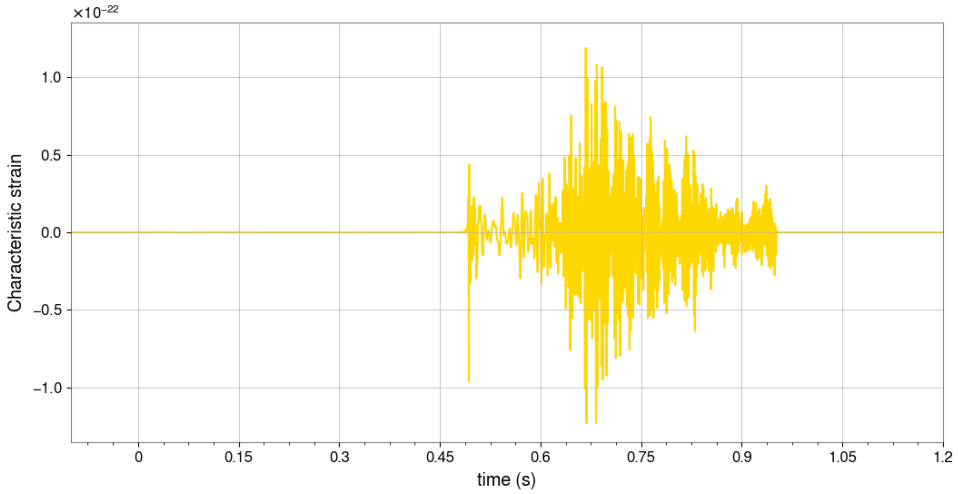

Figure B.11: Example waveform of the Pan s40_fr model.

Powell and Müller, 2019 [41] (s18)

Powell and Müller [41] analyse different models and perform simulations in different evolution phases. The progenitor mass of the chosen model is 18 solar masses and no angular velocity is present. The neutrino-driven model is primarily dominated by g-modes. No SASI and convection can be observed.

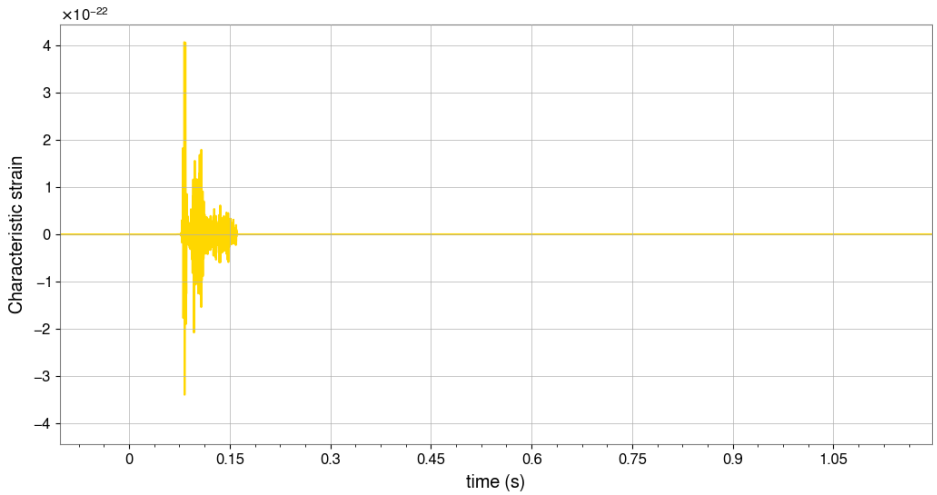

Figure B.12: Example waveform of the Powell and Müller s18 model.

Powell and Müller, 2020 [42] (m39)

The chosen model from Powell and Müller is the m39 model and is a rapidly rotating Wolf-Rayet star with an initial mass of 39 solar masses. It's a neutrino-driven model with a angular velocity of 600 km/s or 0.542 rad/s. Due to its fast rotation it is a very energetic model and is claimed to be detectable up to 2 Mpc with ET. It's gravitational wave signature also includes important contributions of SASI, prompt convection and f- and g-modes.

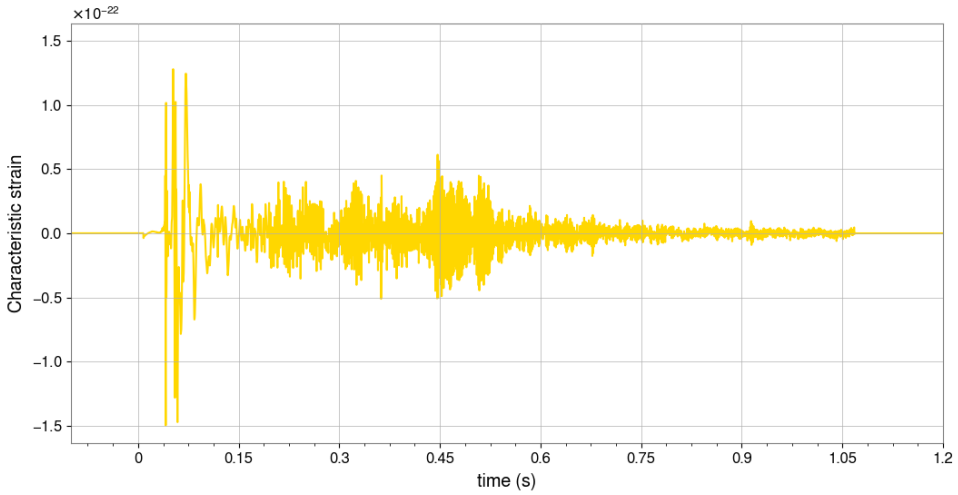

Figure B.13: Example waveform of the Powell and Müller m39 model.

Powell et al., 2021 [43] (Z100_SFHx)

Powell et al. [43] were able to explode a star with a blistering progenitor mass of 100 solar masses. The neutrino-driven model has no angular velocity and shows SASI activity with some Rayleigh-Taylor convection on top. No signs of prompt convection were observed though.

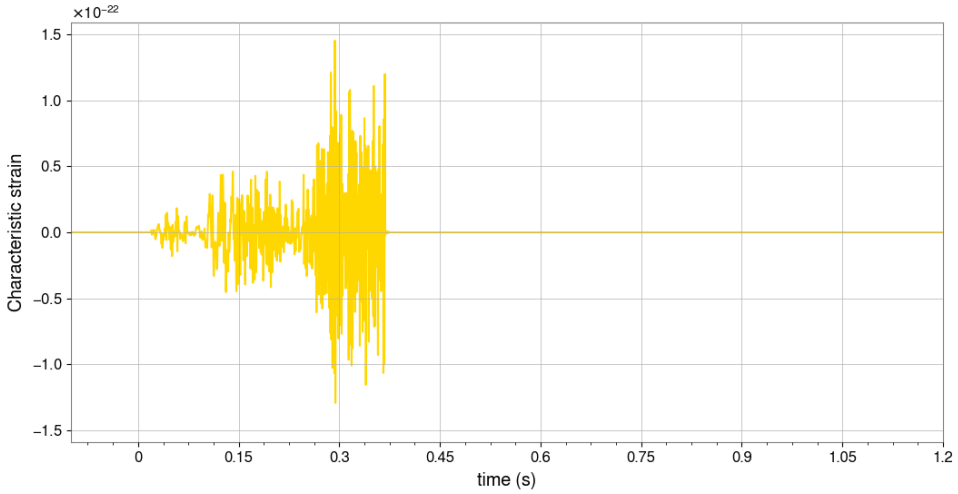

Figure B.14: Example waveform of the Powell Z100_SFHx model.

Radice et al., 2019 [44] (s9, s60)

Radice et al. [44] studied the dependence of the progenitor mass on the gravitational wave signature. They simulated 10 supernova events with each different progenitor masses between 9 and 60 solar masses and no angular velocity. The lightest and heaviest of the two were chosen to be included in this thesis. Both of the models are dominated by f- and g-modes and show signs of prompt convection.

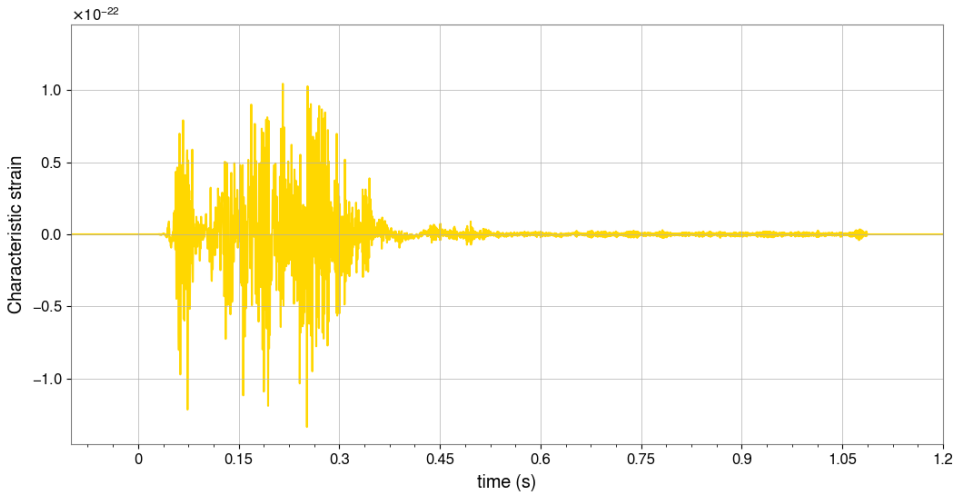

Figure B.15: Example waveform of the Radice s9 model.

Scheidegger et al., 2010 [47] (R4E1FC_L, R3E1AC_L)

Scheidegger et al. [47] performed a systematic study of MHD-driven models with various equations of state, rotational velocity and magnetic field. The R4E1FC_L and R3E1AC_L have rotational velocities of respectively 12.6 and 9.4 rad/s. These models are the most energetic ones of the whole list. Their devastating initial conditions mean that the models exhibit clear bounce signals in their waveforms. Also convection and prompt convection can be observed.

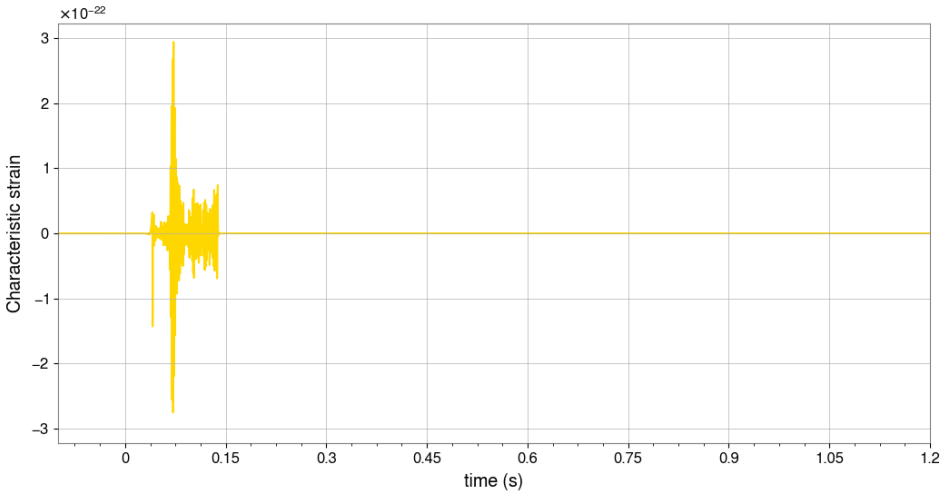

Figure B.16: Example waveform of the Scheidegger R4E1FC_L model.

C Hyperparameters

In the following table, the hyperparameters of the machine learning algorithms used in the last part of the project are presented. Note that only the ones with the biggest effect on the performance are shown. After the table, a short overview of the meaning of these parameters will be explained.

DT			
Feature	Dimensionality reduction	# components	max leaf nodes
prompt convection	PCA	2	4
	UMAP	2	2
SASI	PCA	2	4
	UMAP	2	2
rotating	PCA	2	4
	UMAP	2	2
polynomial SVM			
Feature	Dimensionality reduction	# components	degree polynomial fit
prompt convection	PCA	100	7
	UMAP	4	7
SASI	PCA	100	7
	UMAP	4	7
rotation	PCA	100	7
	UMAP	7	7
KNN			
Feature	Dimensionality reduction	# components	# neighbors
prompt convection	PCA	10	51
	UMAP	2	51
SASI	PCA	10	51
	UMAP	2	51
rotation	PCA	14	51
	UMAP	2	51
Neural Networks			
Feature	Dimensionality reduction	# components	Hidden layer size
prompt convection	PCA	2	(10, 10)
	UMAP	2	(10, 10)
SASI	PCA	2	(10, 10)
	UMAP	2	(10, 10)
rotation	PCA	2	(10, 10)
	UMAP	2	(10, 10)

Table 4: Used hyperparameters for the classification algorithms

Linear regression			
Feature	Dimensionality reduction	# components	
mass	PCA	100	
	UMAP	100	
angular velocity	PCA	100	
	UMAP	100	
LASSO			
Feature	Dimensionality reduction	# components	α
mass	PCA	100	0.002
	UMAP	100	0.002
angular velocity	PCA	100	0.002
	UMAP	100	0.002
Neural Networks			
Feature	Dimensionality reduction	# components	Hidden layer size
mass	PCA	100	(10, 10)
	UMAP	100	(10, 10)
angular velocity	PCA	100	(10, 10)
	UMAP	100	(10, 10)

Table 5: Used hyperparameters for the regression algorithms

- The *number of components* are the dimensions the data is transformed into.
- The *max leaf nodes* is the maximal amount of cells the decision tree is allowed to make. This is purposely set to a low value so that the algorithm would be able to make cells across different models.
- The *degree of polynomial fit* is the maximal degree the polynomial can have to use as a kernel for the support vector machines. A higher number means a more accurate cut but more chance of overfitting.
- The *amount of neighbors* is the literally the amount of nearest neighbors the algorithm has to use to predict the class of the data. Purposly set to an odd number to prevent a draw.
- The *hidden layer size* is the amount of neurons every hidden layer of the neural network has. When an array is given, it means that their are multiple hidden layers.
- α is the penalty term used with LASSO.