Investigations of Hadronic vs Electromagnetic Cascade Identification at The PeV Energy Scale in KM3NeT

Master's Thesis in Physics

presented by

Yara Darras

July 10, 2023

Erlangen Centre for Astroparticle Physics
Department of Physics
Friedrich-Alexander-Universität Erlangen-Nürnberg

Supervisor: PD Dr. Thomas Eberl

Abstract

KM3NeT/ARCA is an underwater Cherenkov detector located 100 km offshore Portopalo di Capo Passero on the south-eastern coast of Sicily. Its main goal is the detection of high-energy neutrinos from astrophysical sources such as supernova remnants. Neutrino interactions with matter are detected as events of different topologies depending on the neutrino flavour and interaction type. The Glashow resonance is a particular type of neutrino interaction in which an electron antineutrino with an energy of about 6.3 PeV interacts with an electron producing a W boson which can decay through different channels. In this thesis, the use of deep learning techniques to distinguish between hadronic and leptonic decay modes of the W boson produced in the Glashow resonance is described.

Contents

Chapter 1

Introduction

Neutrinos are the second most abundant particles in the universe after photons of cosmic microwave background[1], although they rarely interact with matter. The probability to detect neutrinos is influenced by two factors: the cross section of the neutrino interaction, which increases with neutrino energy, and the neutrino fluxes whose properties depend on the source. Most of the known neutrino sources emit fluxes that decrease with the neutrino energy following a power-law dependence. Looking at the spectra of the different known neutrino sources, one can see that different sources emit neutrinos in different energy ranges, and that the higher the neutrino energies are, the lower the flux is. Detectors of high energy neutrinos need to compensate the low fluxes by enlarging the detection volume. Neutrino telescopes such as IceCube or KM3NeT/ARCA are designed to cover a volume of $1km^3$. The IceCube neutrino detector was built in the ice at the South Pole while the KM3NeT (km^3 neutrino telescope) is currently under construction in the Mediterranean Sea. These detectors implement the "Cherenkov" technique, which is the measurement of Cherenkov radiation induced by secondary charged particles resulting from neutrino interactions in the vicinity of the detector. Neutrino events in neutrino telescopes have different topologies depending on the neutrino interaction channel. The two main topologies are shower-like (where secondary charged particles are either hadrons or electrons) and track-like (with a muon as a secondary charged particle), see Figure1.1. In the energy range between $\sim$ 4 PeV and $\sim$ 10 PeV, the electron antineutrino cross section is dominated by the Glashow resonance. In this, an electron antineutrino interacts with an electron producing a W boson on the mass shell. The subsequent decay of the W boson can take place through different channels, and each of them could lead to a different event topology. Km3NeT/ARCA is designed for the detection of neutrinos in the TeV-PeV range and is therefore suited for the detection and study of

neutrino interactions around the Glashow resonance.

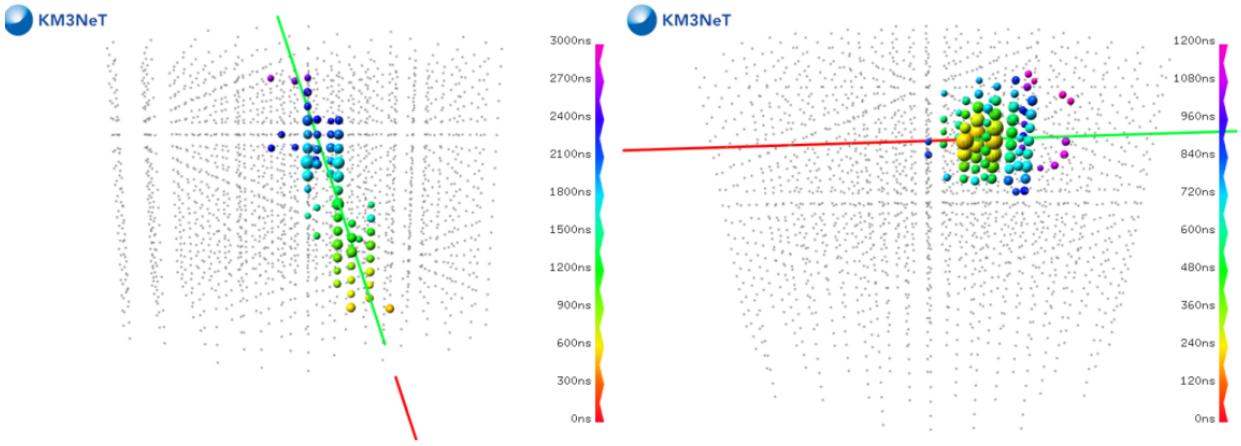


Figure 1.1 – Topologies of neutrino events in KM3NeT.

In this thesis, I investigate the capabilities of the KM3NeT to discriminate between Glashow events according to the decay of the W-boson, hadronic or leptonic cascades. The second chapter gives an overview of neutrino interactions, neutrino oscillations and their sources. The KM3NeT will be presented in detail in the third chapter. The fourth chapter explains the theory of the Glashow resonance. The fifth chapter introduces graph neural networks which are used to conduct the analysis.

Chapter 2

Neutrinos

The Standard Model of particle physics classifies elementary particles into fermions and bosons. Fermions are matter particles which are further classified into quarks and leptons. The leptons consist of charged leptons, electron, muon, tau, and their antiparticles, associated to each one of them an electrically neutral (anti)neutrino. Neutrinos interact with matter through weak processes. They also have a property that goes beyond the standard model, a non-zero mass, which in comparison with that of the charged leptons is very small. This property is needed to explain neutrino flavour oscillations phenomena.

2.1 Neutrino Interactions

Neutrinos interact through the weak force with nucleons, electrons, or nuclei via two interaction processes: charged current (CC) interaction, in which a $W^{\pm}$ boson is exchanged, neutral current (NC) interaction, mediated by the Z^0 boson. These neutrino interactions occur across various energy scales.

Figure 2.1 summarizes the CC neutrino cross section across a specific energy range. For energies below $\sim$ 1 GeV, the interactions are predominantly quasielastic scattering (QE). In this, neutrinos scatter off the entire nucleon and change into the corresponding charged lepton. If the interaction takes place through a NC, the interacting neutrino is present in the final state, and no charged lepton is produced. This is a consequence of charge and lepton number conservation. As for the target, the neutron is converted to a proton in a CC interaction. In the case of an antineutrino scattering, the proton is converted to a neutron. At energies above $\sim$ 1 GeV, neutrinos can excite the nucleon producing a baryon resonance state. This interaction

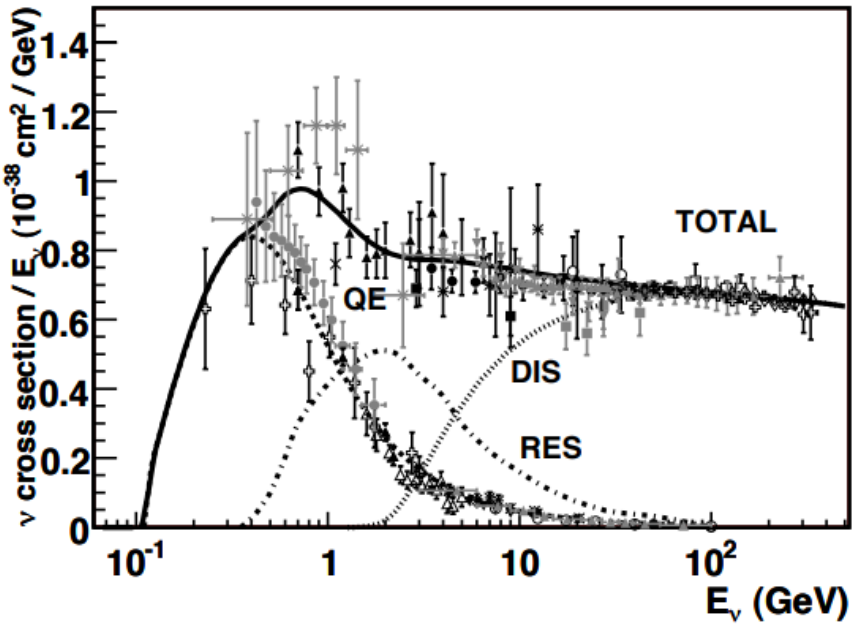


Figure 2.1 – Charged current neutrino interaction with nucleons cross section divided by neutrino energy as a function of energy[2]

is called Resonance production (RES), which then decays producing combinations of nucleons and mesons. Given enough energy, the neutrino can actually begin to resolve the internal structure of the target (interacts with the constituents of the nucleons), a process called deep inelastic scattering (DIS) which dominates at neutrino energies above $\sim$ 10 GeV. CC and NC processes are possible in this interaction, where hadrons are present in both cases as final state particles[2]. Figure 2.2 shows the cross section of CC antineutrino interactions with nucleons. At 6.32 PeV energies, the Glashow resonance interaction dominates over DIS. This interaction will be explained explicitly in chapter 4.

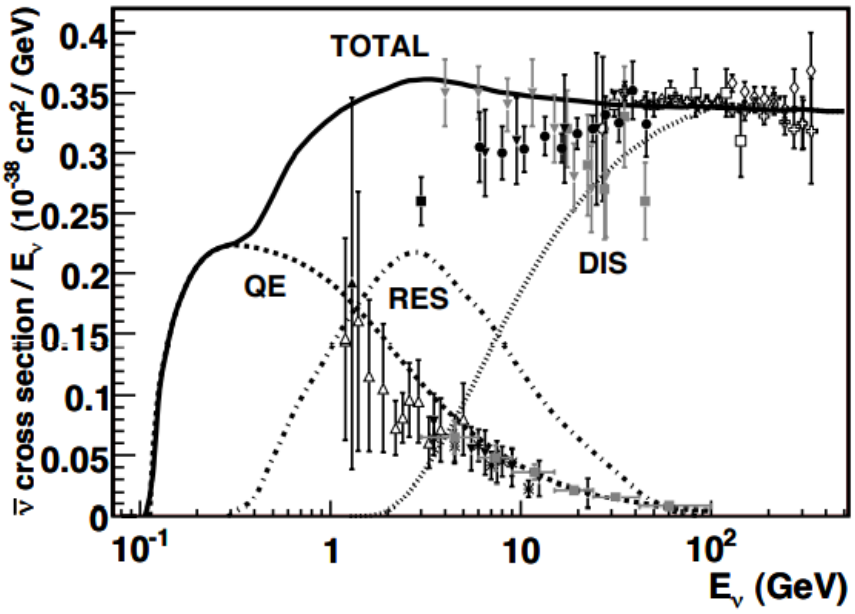


Figure 2.2 – Charged current antineutrino interaction with nucleons cross section divided by neutrino energy as a function of energy[2]

2.2 Neutrino Oscillations

in 1957 Pontecorvo proposed the idea of neutrino oscillations by stating that the physical state of neutrinos produced in weak interaction processes is a superposition of neutrino states with definite masses[3]. The mixing of the neutrino mass eigenstates in neutrino flavours is described by a matrix, called Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix. The idea of neutrino mixing leads to neutrino flavor changes (neutrino oscillations). In the case of three flavor neutrino mixing, these flavor and mass eigenstates are related by the PMNS 3×3 unitary mixing matrix (U)

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U_{\alpha i} |\nu_i\rangle \tag{2.1}$$

with $\alpha = e, \mu, \tau$. One of the parameterization of the PMNS unitary matrix U has three mixing angles $\theta_{12}, \theta_{23}, \theta_{13}$ and one phase describing the CP violation δ_{CP}[3].

$$U = \begin{pmatrix} c_{13}c_{12} & c_{13}s_{12} & s_{13}e^{-i\delta_{CP}} \\ -c_{23}s_{12} - s_{13}s_{23}c_{12}e^{i\delta_{CP}} & c_{23}c_{12} - s_{13}s_{23}s_{12}e^{i\delta_{CP}} & c_{13}s_{23} \\ s_{23}s_{12} - s_{13}c_{23}c_{12}e^{i\delta_{CP}} & -s_{23}c_{12} - s_{13}c_{23}s_{12}e^{i\delta_{CP}} & c_{13}c_{23} \end{pmatrix} \tag{2.2}$$

where $c_{ij} = \cos\theta_{ij}$, and $s_{ij} = \sin\theta_{ij}$.
The probability of the flavour change for oscillations in vacuum is[3]

$$P_{\nu\alpha\to\nu\beta}(L,E) = 4\sum_{i>j}(|U_{\alpha i}|^2|U_{\alpha j}|^2)\sin^2(\frac{\Delta m^2_{ij}}{4E}L) \tag{2.3}$$

where Δm^2_{ij} are the mass squared difference with i, j = 1, 2, 3.
Neutrino oscillations have been confirmed in several experiments, and the observation of them implies that neutrinos have masses. However, their masses are expected to lie somewhere in the range of only a few eV[4]. The sign of Δm^2_{12} has been determined to be positive via measurements of solar neutrinos, which are neutrinos originating from the Sun, in the Sudbury Neutrino Observatory (SNO) experiment while the sign of Δm^2_{31} is still unknown. Therefore, two distinct ways of ordering the neutrino masses are possible, the normal hierarchy with $m_1 < m_2 < m_3$, and the inverted hierarchy with $m_3 < m_1 < m_2$. A scheme of these orderings is shown in Figure 2.3.

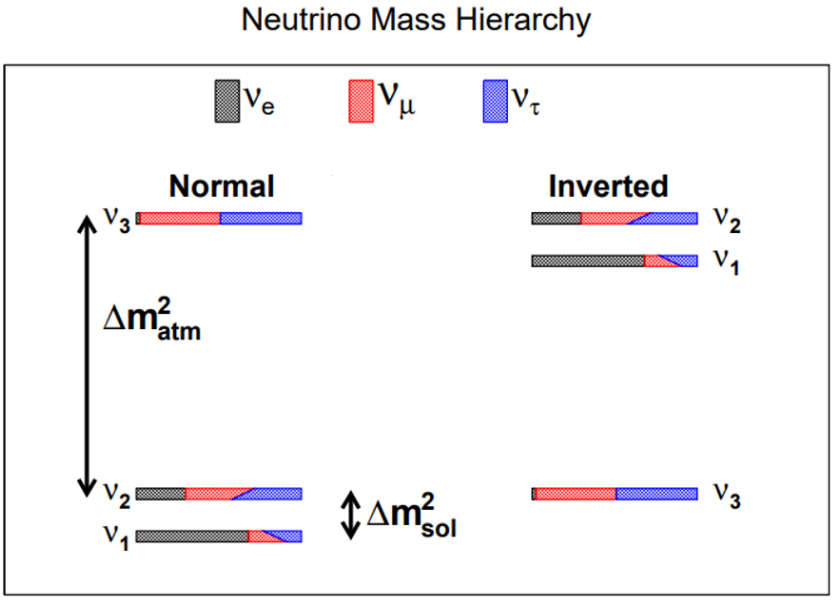


Figure 2.3 – Normal and inverted orderings of neutrino masses are shown, where each mass eigenstate has different flavor composition.
$\Delta m^2_{atm} \sim |\ \Delta m^2_{31}\ | \sim |\ \Delta m^2_{32}\ |$ and $\Delta m^2_{sol} \sim \Delta m^2_{12}$ stands for the atmospheric and the solar mass squared differences, respectively. [5].

Neutrino oscillations take place in matter as well. KM3NeT/ORCA will be one of the experiments that will determine the neutrino mass ordering in the next years.

2.3 Neutrino Sources

Figure 2.4 is taken from [6], and shows the neutrino fluxes at Earth from different sources. The neutrino energies range between $\mu eV - EeV$, and the flux spans over 52 orders of magnitude. This section focuses on high energy astrophysical neutrinos and atmospheric neutrinos, since they are the two main sources of signal in the KM3NeT which is used to conduct the analysis.

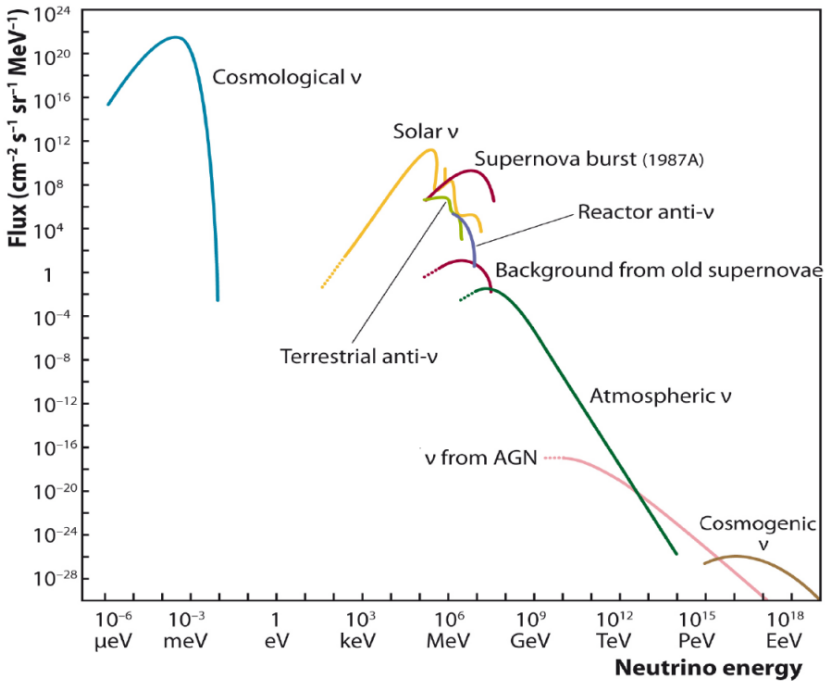


Figure 2.4 – Known neutrino sources and their fluxes as a function of neutrino energies

2.3.1 Atmospheric Neutrinos

Atmospheric neutrinos are produced by cosmic rays interacting with the atmosphere of the Earth. There are two types of atmospheric neutrinos, the conventional neutrinos which are produced when the secondary charged pions or kaons decay. Muons are produced in addition to these neutrinos, that may also decay and produce other neutrinos and antineutrinos. The expected ratio of $\nu_e : \nu_\mu = 1 : 2$, see equations (2.4, 2.6). Equation 2.5 shows one of the most probable decay channels of the kaons.

$$\pi^{\pm} \rightarrow \mu^{\pm} + \nu_\mu(\bar{\nu_\mu}) \tag{2.4}$$

$$K^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu}(\bar{\nu_{\mu}}) \tag{2.5}$$

$$\mu^{\pm} \rightarrow e^{\pm} + \nu_{e}(\bar{\nu_{e}}) + \nu_{\mu}(\bar{\nu_{\mu}}) \tag{2.6}$$

The second type is prompt neutrinos which are associated with the decay of heavy hadrons.

Neutrino oscillations depend on the path length traveled by neutrinos and neutrinos energies. In the case of two-flavor mixing of neutrinos, one of which is ν_{μ}, the probability that a ν_{μ} traveling a distance L is detected as a ν_{μ} is

$$P_{\nu_{\mu}\rightarrow\nu_{\mu}}(L,E) = 1 - \sin^2(2\theta_m)\sin^2(1.27\Delta m^2_{12}\frac{E_{eV}}{L_{km}}) \tag{2.7}$$

where Δm^2_{12} is the squared mass differences of the two related mass eigenstates and θ_m is the mixing angle between these two states. Therefore neutrino oscillations lead to a deviation from the expected flux of a specific neutrino flavor as L/E varies[7].

Atmospheric neutrino flux depends on the zenith angle, where the flux is larger closer to the horizon. This is due to the increased path length traveled by secondary muons, resulting in a higher decay probability for muons. Atmospheric neutrinos are either down-going neutrinos i.e., coming from the atmosphere above, or up-going neutrinos, which come through the Earth and experience oscillations in matter.

As muons traverse the medium, they lose energy through different kinds of interactions with the surrounding matter depending on the energy of the muon. At energies below $\sim$ 100 GeV, the loss is dominated by ionization processes, which is proportional to the density of the medium (g/cm^3) times the distance traveled in centimeters. This leads to an energy loss of about 2 MeV per g/cm^2. The atmospheric muons pose a significant background to the measurement of neutrinos in KM3NeT, which is being built deep in the Mediterranean Sea to reduce this background. The vertical depth of the atmosphere is about 1000 g/cm^2. Therefore, muons will lose $\sim$ 2 GeV due to ionization before reaching the ground. At sea level, atmospheric muons arrive with an average flux of about 1 muon per square centimeter per minute[8].

2.3.2 Astrophysical Neutrinos

High-energy neutrinos produced in Galactic and extragalactic astrophysical sources are known as astrophysical neutrinos. Neutrinos can escape dense region of astrophysical sources, due to their low cross sections. They are also unaffected by intervening electromagnetic fields along their paths towards the Earth, since they have no electric charge[9].

The measured flux of astrophysical neutrinos is described by a power law energy distribution ($\propto E^{\gamma}$, with some spectral index γ) ranging between $\sim$10 TeV- PeV energies, but its origin remains unresolved[10].

The IceCube neutrino telescope first observed a diffuse flux (an isotropic diffuse flux generated by extragalactic sources) of high energy astrophysical neutrinos in 2013. Since then, the measurement of a diffuse extragalactic neutrino flux has been confirmed in several detection channels such as cascades and tracks. The atmospheric neutrinos together with atmospheric muons form the background in the search for astrophysical neutrinos[11]. These backgrounds have distinctive spectral, directional, and flavor characteristics that allow their separation from an astrophysical signal. This will be explained in more detail in section 3.2.

The main postulated production models for astrophysical neutrinos are[12]:

1. protons from cosmic accelerators collisions with the ambient matter ($\pi^{\pm}$ mode)

$$p + N \rightarrow \pi^{\pm}, \pi^{0} + X \tag{2.8}$$

$$\pi^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu}(\bar{\nu_{\mu}}) \tag{2.9}$$

$$\mu^{\pm} \rightarrow e^{\pm} + \nu_{e}(\bar{\nu_{e}}) + \nu_{\mu}(\bar{\nu_{\mu}}) \tag{2.10}$$

   The initial neutrino flavour ratio $\nu_e : \nu_\mu : \nu_\tau$=1:2:0, which then oscillates over a very long distance into approximately $\nu_e : \nu_\mu : \nu_\tau$=1:1:1 at Earth.

2. protons from cosmic accelerators collisions with radiation fields (π^{+} mode)

$$p + \gamma \rightarrow \Delta^{+} \rightarrow \pi^{0}(\pi^{+}) + p(n) \tag{2.11}$$

$$\pi^{+} \rightarrow \mu^{+} + \nu_{\mu} \tag{2.12}$$

$$\mu^{+} \rightarrow e^{+} + \nu_{e} + \bar{\nu_{\mu}} \tag{2.13}$$

In this case, we have different production rates of neutrino and antineutrino, with initial flavour ratios of $\nu_e : \nu_\mu : \nu_\tau$=1:1:0 and $\bar{\nu_e} : \bar{\nu_\mu} : \bar{\nu_\tau}$=0:1:0 which then become at Earth $\nu_e : \nu_\mu : \nu_\tau$=14:11:11 and $\bar{\nu_e} : \bar{\nu_\mu} : \bar{\nu_\tau}$=4:7:7.

The presence of magnetic fields leads to the deflection of muons before they decay. Which then changes the initial flavour ratio for each scenario. In the case of damped $\pi^{\pm}$ mode the neutrino ratio for $\nu_e : \nu_\mu : \nu_\tau$ changes from 0:1:0 to 4:7:7 at Earth. In the damped π^{+} mode antineutrinos disappear completely.

Chapter 3

KM3NeT

Muons produced in the muon neutrino CC interactions propagate long distances without experiencing energy loses, and the hit pattern along their path follows a track-like signature. On the other hand, the mean free path for electrons and taus is too short compared to the typical distance between optical modules. While electrons experience radiative energy losses leading to photons which in turn produce electron-positron pairs, tau leptons decay. In both cases, a cascade of particles is produced, which lead to a more homogeneous light distribution in the detector. Hadrons produced in the NC neutrino interactions also produce showers. The different event topologies are summarised in Figure 3.1.

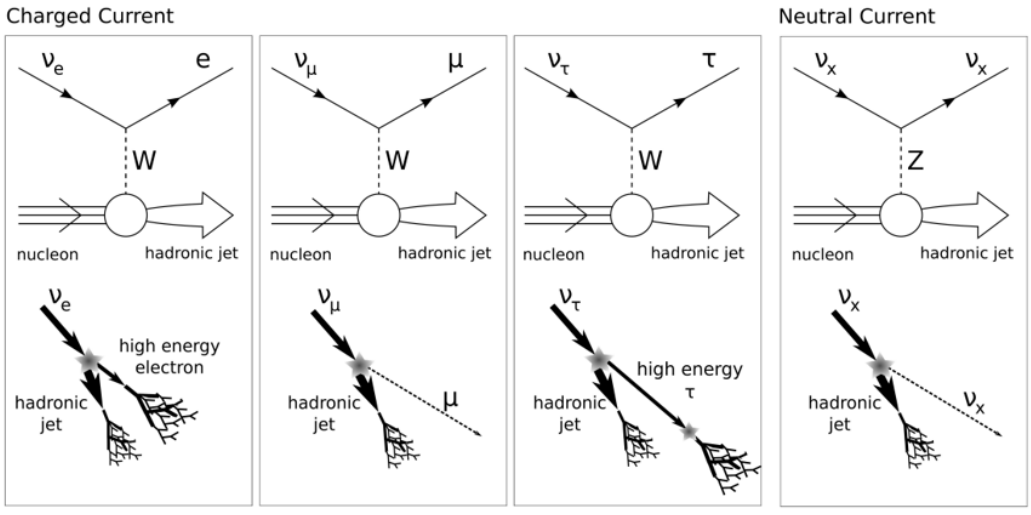


Figure 3.1 – The topologies of different neutrino events[13].

The KM3NeT collaboration is currently building two water Cherenkov neutrino detectors in the Mediterranean Sea. Each detector consists of a three-dimensional array of light sensor modules distributed over large volumes. Such sensor modules are also referred to as Digital Optical Modules (DOMs). These are arranged into vertical structures called detection units (DU) or strings, each of which contains 18 equally spaced DOMs. Each string is kept in a vertical position with the help of a buoy. The ARCA (Astroparticle Research with Cosmics in the Abyss) detector is currently being built 100 km off-shore of Portopalo di Capo Passero on the south-eastern coast of Sicily, Italy at a depth of 3500m. The volume of KM3NeT/ARCA is $\sim 1km^3$ and it is optimized for the measurement of neutrinos from astrophysical sources with energies between $\sim$ 100 GeV and PeV. ARCA will consist of 2 so-called building blocks, each of which will consist of 115 strings. The ORCA (Oscillation Research with Cosmics in the Abyss) detector is being deployed 40 km from the southern coast of France off-shore the city of Toulon[14]. It will consist of a single building block of 115 strings and it's purpose is to measure neutrino oscillation parameters as well as to determine the neutrino mass ordering. For this, ORCA will use atmospheric neutrinos in the range between $\sim$ 1GeV and 100 GeV. The geometries of ARCA and ORCA have been optimised for the detection of neutrinos in different energy ranges. The DOMs are placed closer to each other in ORCA than in ARCA. The characteristic distances for each detector are summarised in table 3.1.

Blocks	String spacing (m)	DOM spacing (m)	Depth (m)	Instrumented mass (Mton)	Building blocks
ORCA	23	9	2470	8	1
ARCA	90	36	3400	500*2	2

Table 3.1 – The characteristic of both ARCA and ORCA detectors

3.1 Detection Principle

The KM3NeT optical modules detect the Cherenkov radiation induced by the charged particles produced in neutrino interactions. The charged particle moving inside a polarizable medium with molecules excites the molecules, which then de-excite by emitting electromagnetic radiation. These waves move out spherically at the phase velocity of the medium. For particles with speed less than the speed of light in the medium, the radiated waves bunch up in the direction of motion without overlapping. However, if the particle moves faster than the light speed in the medium, the emitted waves interfere and add up constructively leading to a coherent radiation which

is the Cherenkov radiation emitted at an angle with respect to the particle direction. This angle is called Cherenkov angle, which is calculated by

$$\cos(\theta_{Ch}) = \frac{1}{\beta n} \tag{3.1}$$

where n is the refractive index of the medium and β is the velocity of the particle with respect to the speed of light in vacuum. Due to the Cherenkov angle, a light cone is defined around the track of the moving charged particle as shown in Figure 3.2.

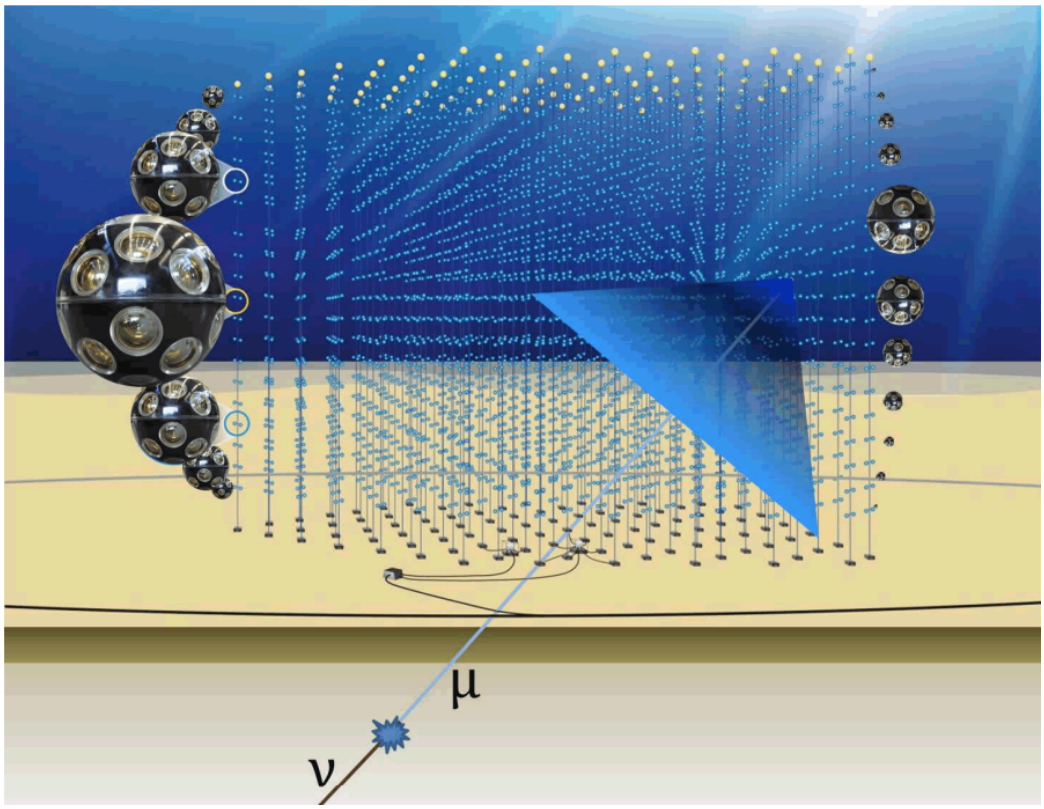


Figure 3.2 – Visualization of the Km3net building block with 115 strings, each of which contains 18 DOMs.

For particles with $\beta \approx 1$ propagating in sea water, the Cherenkov angle is $\theta_{Ch} \approx 42°$. Photons reaching the PMTs produce the so-called hits. A hit contains the position and time of the photon detection. When a set of causally connected hits is detected, the detector triggers an event and the hit positions and times are fitted to models for particle propagation that allow to infer the direction of the incoming neutrino. On the other hand, the amount of hits allow to estimate the neutrino energy.

3.2 Signal and Background in KM3NeT

As previously mentioned, the main scientific objective of KM3NeT/ARCA is the detection of high-energy neutrinos of cosmic origin, which then could point back to the astrophysical objects in which cosmic rays are accelerated. KM3NeT/ARCA is sensitive to energy range of $\sim$ 100 GeV-PeV. The signal produced by astrophysical neutrinos in this detector can be masked by background sources which are atmospheric neutrinos and atmospheric muons. The atmospheric muon background, which is a background source in both ORCA and ARCA detectors, can be significantly reduced by considering only up-going events since only neutrinos can traverse the Earth without being absorbed. Due to this, the KM3NeT/ARCA is more sensitive to Galactic sources (where most potential Galactic sources are in the Southern sky) than the IceCube detector located at the South Pole in the energy range where the signal is expected (a few TeV to a few tens of 10 TeV)[14].

The discrimination between atmospheric and astrophysical neutrinos could be achieved by analysing the spectral distribution of the measured sample. Where at higher energies, a diffuse cosmic neutrino emission is likely to produce an excess of events. This is due to softer atmospheric neutrino spectrum ($E^{-3.7}$) compared to the spectrum expected from cosmic neutrinos (typically E^{-2}). The astrophysical neutrinos can also be identified by searching for coincidences between its signal and the signal from other astrophysical messengers[15].

The optical background sources in KM3NeT are: bioluminescence, which is the emission of light by bioluminescent organisms present in the deep-sea, and hits from ^{40}K decays, where ^{40}K is a radioactive isotope that decays into ^{40}Ca. This process then emits electrons with very high energy, which can emit Cherenkov radiation that is detected by the PMTs[16].

3.3 Simulation

KM3NeT performs scientific analyses by comparing the distributions of recorded events to reference distributions which correspond to different scientific scenarios. Such reference distributions contain information about the physics model under test, as well as about all the detector performance details. It is therefore not trivial to obtain them, and instead of deriving them analytically, they are obtained from complex simulations that represent as much as possible the data taking conditions. The simulation (Monte Carlo MC) chain is split into the three following main parts:

1. Event generation:

   In which the generation of particle fluxes (neutrino signal and the atmospheric muon background) incident on the can is done with the consideration of the detector geometry and deep underwater environment. The can is defined as the instrumented volume (volume of a building block) extended by ~ 3 photon absorption lengths, which is also known as Cherenkov sensitive volume. Astrophysical and atmospheric fluxes of (anti-)neutrinos of all three flavours are simulated where neutrinos are propagated through the Earth and their interactions are generated in rock and sea water. For reactions outside the can, long-range interaction products (muons) are propagated to the can.

   Both NC and CC reactions are simulated. The deep inelastic scattering (DIS) and the Glashow resonance interactions are simulated since they are relevant in the energy regime at which the KM3NeT operates. The propagation of muons in rock and water is performed taking into account both radiative and ionization energy losses, as well as the decay probability of the particle. The average distance travelled by a particle before interacting or decaying is called the range, which depends on the energy of the particle. Figure 3.3 shows this range for all charged leptons in both rock and water. If the atmospheric muons have an energy in the TeV range or above at sea level, they can penetrate the detector. Therefore, they highly contribute to the background which is simulated using MUPAGE[17]. In KM3Net/ARCA gSeaGen is used to generate high energy neutrino interactions[14].

   The particles produced by event generation correspond to neutrino interactions which energy spectrum follows a power law with a spectral index γ defined by the user. The chosen input parameters determine

the number of simulated interactions in a given solid angle Ω, and volume V .i.e., This generated sample doesn't necessarily represent a realistic neutrino flux. To obtain a sample of particles that represents the arrival to the Earth of a certain neutrino flux, weights to each of the events need to be assigned to finally obtain the correct distribution in energy, time, volume and solid angle. This weight is calculated by

$$W = \frac{V_{gen} I_E I_\Omega E^\gamma \rho N_A \sigma(E) P_{Earth}(E_\nu, \theta) t_{gen}}{N_{gen}} \tag{3.2}$$

where $I_E = \int_{E_{max}}^{E_{min}} E^\gamma \, dE$, $I_\Omega = \int_{\theta_{max}}^{\theta_{min}} E^\gamma \, d\Omega$, N_A is the Avogadro number, ρ is the mass density per unit volume, $\sigma(E)$ is the cross section of the neutrino-nucleon interactions, P_{Earth} is the probability that neutrinos propagate through the Earth, t_{gen} is the (arbitrary) time represented by the simulation, and N_{gen} is the number of events generated in the simulation.

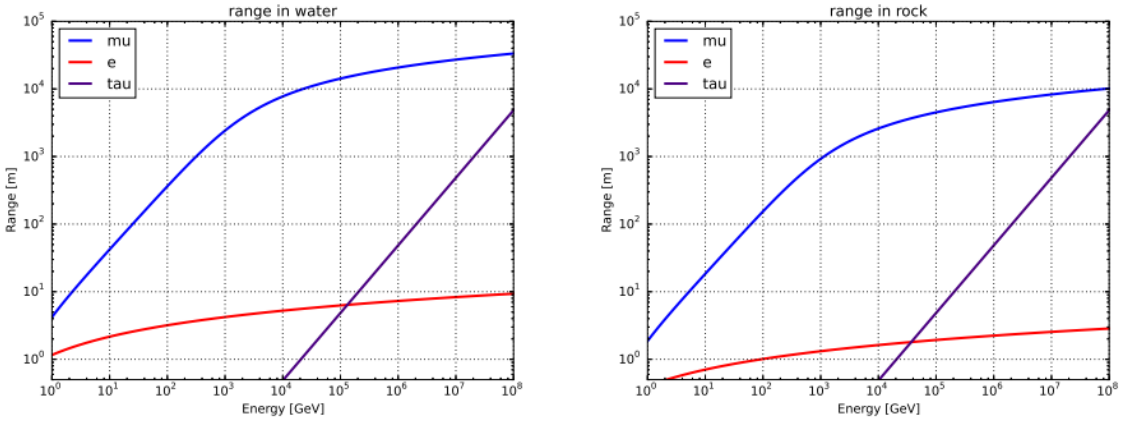


Figure 3.3 – Ranges of all three charged leptons in rock and water[15].

2. Light generation:

   The simulation contains information about the positions of all PMTs in a three dimensional space. The Cherenkov radiation emitted along the propagation of the particle is simulated, which then generates hits in PMTs. This is done for the low energy case using the KM3Sim light simulator[18]. For the high energy case, it would require too much computational power. Instead, the probability of a photon arriving at each of the PMTs in the detector is calculated for any particle traveling through the can using JSirene light simulator[19].

3. Detector response simulation:

   KM3NeT PMT hits are recorded using the start time and the duration of the signal above a predefined threshold called time-over-threshold (ToT). The PMT analog signal above this threshold is digitised into a hit.

   The expected background signal from bioluminescence, dark current and ^{40}K decay present in the sea water is simulated by the Trigger Efficiency code which adds background photons to the output of the previous step. Trigger Efficiency also simulates the detector response to the generated light and produces events that look like the observed ones by simulating the effects of the electronics and applying the different triggers[14].

All of this is shown in Figure 3.4. Finally, the events are reconstructed. This reconstruction is achieved by separate algorithms optimized for each event type, Aashowerfit for reconstruction of shower events, and JGandalf for reconstruction of track events. Reconstruction is not a part of the simulation and can be applied on the data in the same way as on MC in the processing chain.

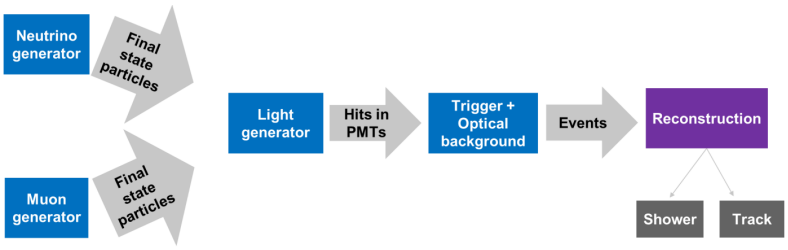


Figure 3.4 – Simulation chain scheme

Chapter 4

Glashow Resonance

Glashow resonance was predicted by Sheldon Glashow in 1959[20]. The theory states that an electron antineutrino interacts with an electron producing a W boson on the mass shell, which can then decay either hadronically or into leptons. See Figure 4.1.

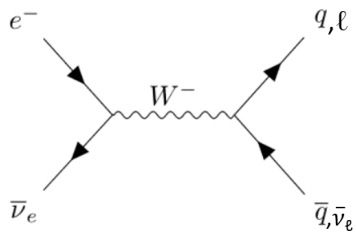


Figure 4.1 – Feynamn diagram of the Glashow resonance interaction

The total cross section for the Glashow Resonance is

$$\sigma_{tot}^{GR}(E_\nu) = (3 + 6.3415) * \frac{G_F^2 M_W^4 (m_e^2 - m_\mu^2 + 2m_e E_\nu)^2}{3m_e E_\nu (\Gamma_W^2 M_W^2 + 2\pi (M_W^2 - 2m_e E_\nu)^2)} \tag{4.1}$$

where G_F is Fermi constant, m_e is electron mass, m_μ is muon mass, M_W is the W boson mass, and Γ_W is W boson width. The factor 3 comes from approximately equal cross sections for the leptonic decays. The factor 6.3415 corresponds to the relation between hadronic and leptonic branching ratio of the W boson decay.

As previously stated, the neutrino-electron cross section is usually smaller than the neutrino-nucleon by orders of magnitude. However, the scattering of $\bar{\nu_e}$ with electrons is enhanced in the energy range of 4 PeV-10 PeV, where

it exceeds that of DIS by more than two orders of magnitude. This is shown in Figure 4.2.

Assuming the electron is at rest, the antineutrino energy at which the cross section of the Glashow resonance reaches its maximum value is

$$E_\nu = \frac{M_W^2 - m_e^2}{2m_e} \approx 6.32 PeV \tag{4.2}$$

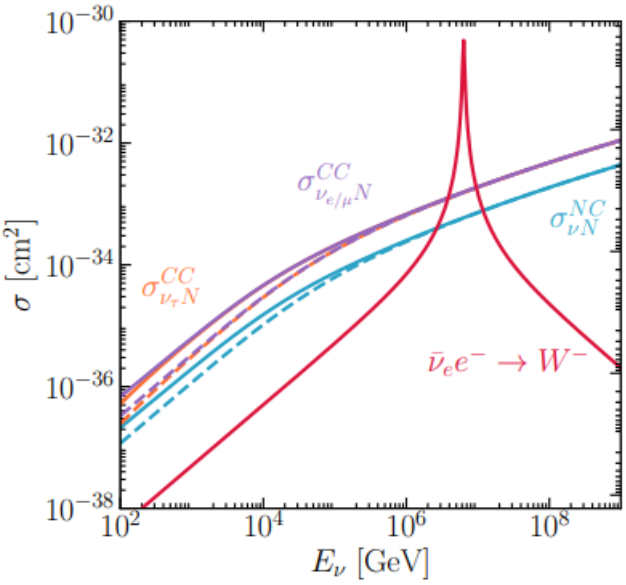


Figure 4.2 – Cross sections of neutrino scattering on a nucleon or an electron. The red curve represents the Glashow resonance interaction. The solid curves are for neutrinos and dashed ones for antineutrinos. The blue curves signify the NC deep inelastic scattering interaction for all neutrino flavors. The CC DIS interactions for ν_μ/ν_e is indicated by purple curves and for ν_τ by orange curves [21].

Neutrinos from atmospheric sources tend to have lower fluxes in comparison to astrophysical neutrinos in the energy range where the Glashow resonance dominates. These astrophysical neutrinos provide the energy needed to produce a W^- boson. The decay of the W boson produces either hadronic showers with deposited energy around 6.3 PeV (branching ratio of $\sim 67\%$), or leptons directly with a lower deposited energy (branching ratio of$\sim 11\%$

for each flavor). In the case of electron presence in the final state, an electromagnetic shower is produced. Meanwhile, the produced muon will survive as a long track, and the tau with a lifetime of $2.9 * 10^{-13}$ s decays. A leptonic event is defined as an event where the W boson produced by the Glashow resonance interaction decays into either an electron, a muon or a tau along with their corresponding antineutrinos. In this type of event, neutrinos escape with most of the energy, unlike the hadronic scenario.

The event rate expected from the Glashow resonance in the KM3NeT/ARCA detector within one year is calculated from

$$Rate = \Phi \sigma n d \tag{4.3}$$

where Φ is the integrated anti-neutrino flux in a given time period, σ is the cross section of $\bar{\nu_e}$, n is the number density of the medium, and d is the path length of an interacting particle. According to [22], the result is 0.42 event/year.

IceCube has detected one event with a visible energy of $6.05 \pm 0.72 PeV$. Based on its energy and direction, it is classified as an astrophysical neutrino at the 5σ confidence level. Features consistent with the production of secondary muons in the particle shower indicate the hadronic decay of a resonant W boson[23].

Chapter 5

Analysing KM3NeT/ARCA Data Using Graph Neural Networks

The reconstruction of the properties of a particle based on the data that was measured by the KM3NeT can be achieved using deep learning techniques. This prediction is an approximation of the real value of the property. The deep Learning techniques applied in this thesis are based on artificial neural networks (ANN), more specifically graph neural networks (GNN), which is a class of artificial neural networks where the processed data are represented by graphs. This chapter will give a brief overview of artificial neural networks, shortly present graph neural networks, and explain how to analyze KM3NeT/ARCA data using deep learning techniques.

Figure 5.1 summarizes the implementation of GNNs for analysing KM3NeT/ARCA data to perform event identification. KM3NeT/ARCA MC events are represented by graphs, which are then the inputs of the network (GNN). The datasets are split into two sets, training and validation. One would train only over the training set, then the model has learned. After that, the performance of the model is checked with the validation set, where for this model it is done after every epoch. The number of epochs is a hyperparameter that defines how many times the model will loop over the entire training dataset. This training is done over a number of epochs to ensure that the model converges (i.e., model's loss or error moving towards a minimum), where the loss and the accuracy are calculated for both training and validation files.

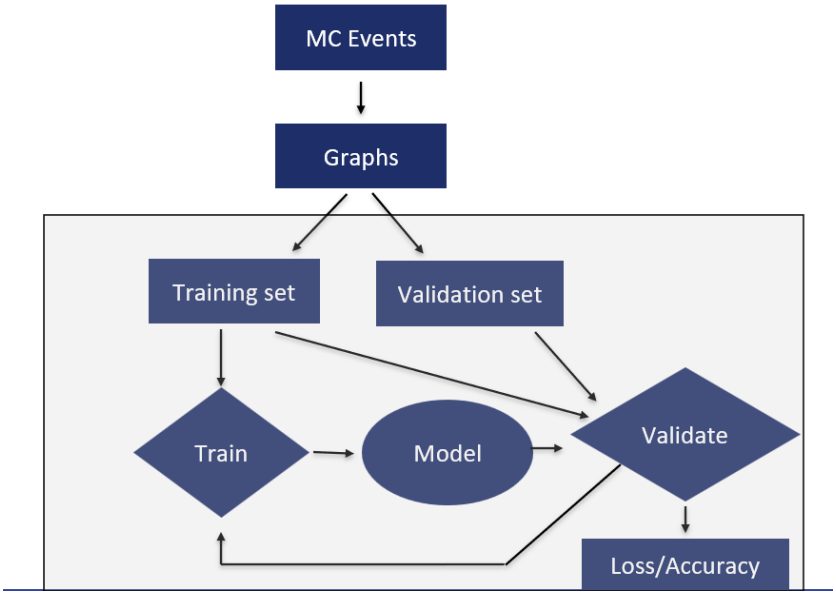


Figure 5.1 – Implementation of GNNs for Analysing KM3NeT Data

5.1 Artificial Neural Networks

An artificial neural network consists of connected artificial neurons (nodes) which can be arranged in layers. The stacking of multiple hidden layers (which are the layers between the input and the output) can enhance the network's ability to reproduce desired values, this method is known as deep learning. Each neuron usually applies the following operation to its N inputs[24]

$$f(x_1, .., x_{1=N}) = \Theta(\sum_{j=1}^{N} w_j x_j + b) \tag{5.1}$$

which is a sum over every input (x_j), each weighted with a real number w_j, then adding the bias b (a constant), and after that we apply the activation function Θ. This is an arbitrary function (non-linear function) applied to the inputs, which affects the ability of the network to approximate target functions[25], where the target function is an approximation function that is defined by its free parameters, the weights w and biases b of all the layers in the network. These free parameters need to be chosen to ensure that the output values are as close as possible to the real values of the desired properties, such closeness is defined in terms of some specific metric. To achieve this, we adjust the weights and the biases of the network repeatedly until the difference between the actual output of the network and the desired output reaches a minimum, this procedure is called back-propagation[26].

The difference between the outputs is defined by a cost function, which quantifies the error between predicted and actual outputs. The free parameters are changed in the direction of the negative gradient of the cost function, which is called gradient descent method (see equation 5.2). A cost function therefore measures how well the network approximates the desired output. The purpose of the training process is finding a set of parameters that minimizes the cost function for the whole training dataset

$$p \to p - \eta \frac{\partial C}{\partial p} \tag{5.2}$$

where p are the free parameters (w, b), η is the learning rate, which is the step size of the descent, and C is the cost function.

The sigmoid function is an example of an activation function (see equation 5.3). It is a differentiable function that ranges between 0-1. For very large or very small x values, the derivative of this function becomes very small (very close to zero) leading to gradient vanishing in deep networks. This is a disadvantage of the function because the derivative of the cost function with respect to a certain parameter (p), that is calculated through the chain rule shown in equation 5.4, depends on the derivative of the output of the network (F) with respect to that same parameter (p). Therefore, the parameters update (see equation 5.2) will be very small, which in turn slows the training of deep networks.

$$\Theta(x) = \frac{1}{1 + e^{-x}} \tag{5.3}$$

$$\frac{\partial C}{\partial p} = \frac{\partial C}{\partial F} \frac{\partial F}{\partial p} \tag{5.4}$$

Another activation function is the rectified linear unit (ReLU), which is the most used activation function nowadays. This function f(x) is zero when $x < 0$, and f(x) is equal to x for $x \geq 0$. But the issue with this function is that at negative values it becomes zero. The neuron would not receive updates to its weights anymore, and would be stuck in this state. Thus, decreasing the ability of the model to train properly. To overcome this problem, some modifications have been done to the ReLU, like leaky ReLU. See [27] for more details.

Neurons from different layers can be connected in multiple ways. When the input to a specific layer consists only of neurons from the previous layer, it is called a feed-forward network[24]. Figure 5.2 shows a multilayer feed-forward neural network.

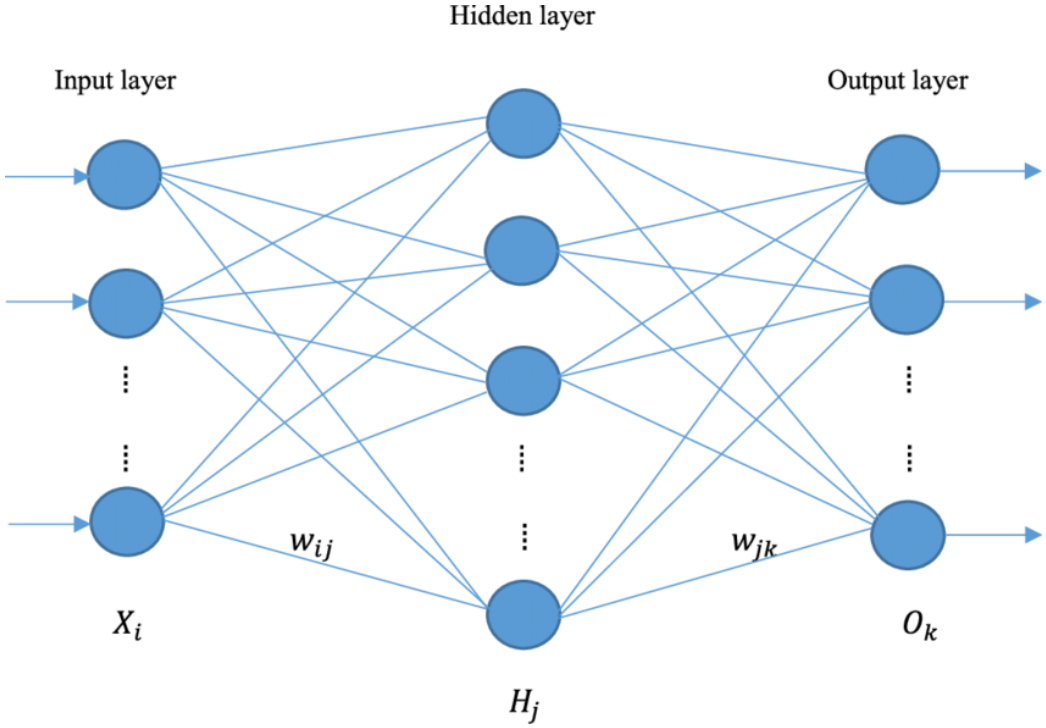


Figure 5.2 – A multilayer feed-forward neural network[28], where X_i are the inputs, w_{ij} are the weights, H_j are the hidden layers, and O_k are the outputs.

The quality of the network's prediction can be assessed on a small sample, called validation dataset. The choice of cost (loss) function is task specific. In the case of a classification task (which is done in this thesis), the categorical cross entropy loss function is chosen, which is defined as

$$C(\vec{y}^{\,true}, \vec{y}^{\,predict}) = -\sum_{i=1}^{n} \vec{y}_i^{\,true} ln(\vec{y}_i^{\,predict}) \tag{5.5}$$

where for n different categories, $\vec{y}^{\,true}$ is the desired output, and $\vec{y}^{\,pred}$is the prediction of the network. The categorical cross entropy loss function is often used together with the softmax activation function , which is a generalized sigmoid function for multiple neurons and is used exclusively in the last layer of classification networks. The outputs of softmax function are between zero and one[24].

5.2 Graph Neural Networks (GNNs)

Graphs are a type of data structure which defines a set of objects (nodes) and their relationships (edges)[29]. Graph neural networks working principle is based on recursive neighborhood aggregation (or message passing procedure), in which the new feature vector of a certain node is computed by aggregating feature vectors of its neighbors. Then after k (which is the chosen number of neighbors) iterations of aggregation, the node is represented by a vector which contains the structural information within its neighborhood of radius k. The representation of the whole graph can then be obtained through pooling, which is an operation that summarizes values of each channel of the input data independently into a single one. It plays a crucial role in reducing the size and complexity of deep learning models while preserving important features and relationships in the input data [30].

Graph neural networks are chosen for the architecture of the network for KM3NeT because the KM3NeT data closely resembles point clouds (discrete set of data points in space)[31]. To encode the information of an event as a graph, each hit is represented by a node, where each node has its features and is connected to its k nearest neighbours in terms of a chosen distance, see Figure 5.3. The features of the node are the coordinates of the hit it represents (time and x, y, z-position of the PMTs, and PMTs pointing direction). The hits of an event are then fed to classification and reconstruction algorithms.

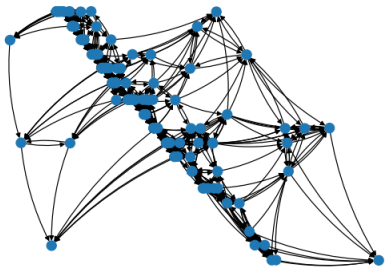


Figure 5.3 – Graph representation of a random event, where each node represents a hit, and each node is connected to its eight nearest neighbors by edges in terms of the euclidean distance in x, y, z and t. The arrows are pointing away from the neighbors, and towards the central node[24].

The model architecture (the design of the model) of the GNN used in this thesis is based on the ParticleNet model[32]. As seen in Figure 5.4a, it consists of three edge convolutional blocks (cf. below for an explanation) followed by a channel-wise global average pooling operation, which aggregates the learned features over all nodes. After that, a fully connected layer with 256 units (the number of nodes in a dense layer) and the ReLU activation function are added. Another fully connected layer but with two units is applied, then a softmax function, which generates an output for the binary classification task is used. The original model was changed for the work in this thesis by removing the dropout in the first layer after the pooling.

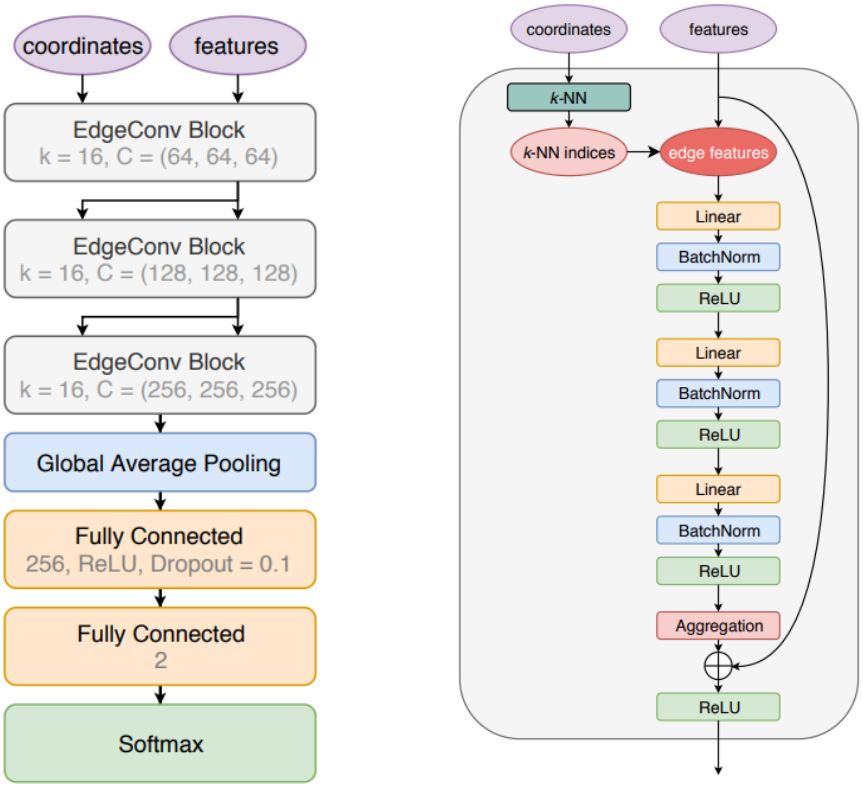
The first edge convolutional block (EdgeConv block) of ParticleNet architecture finds the k nearest neighboring nodes for each node, where it uses the coordinates of each node to compute the distances. Then, edge features are constructed from the features input using the indices of k nearest neighboring nodes. These features are time and x, y, z-positions of the PMTs, and PMTs pointing directions. Given two nodes $\vec{n_i}$, $\vec{n_j}$ with a number of features, the edge features $\vec{e_{ij}}$ going from node i to node j are defined as

$$\vec{e_{ij}} = (\vec{n_i}, \vec{n_j} - \vec{n_i}) \tag{5.6}$$

Each edge features vector is then fed to a multilayer perceptron (MLP), which consists of three layers with a constant number of neurons per layer. The MLP returns an update vector as output, the weights of this MLP are shared among all nodes in the graph. The number of neurons in the last layer of the MLP determines the number of output features. The updated node features are calculated by taking a mean over the update vectors. This MLP is then followed by a batch normalization, and a ReLU. The batch size is a hyperparameter defined as the the number of training samples to go through before updating the model's internal parameter (weights and biases), where training data are sampled in batches. During training, the parameters of a certain layer keep changing, which leads to a change in the output distributions. The subsequent layers use this output as their input, and therefore must keep adapting to this change. This causes the network training to slow down. The purpose of batch normalization is to speed up the training process, which is achieved by setting the mean of each feature over the whole batch to zero and the variance to one[33].

The EdgeConv operation consists of three layers, where each one of them contains the multilayer described above, the batch normalization, and the ReLU. An EdgeConv block is defined by two hyperparameters, the number

of neighbors k, and the number of channels C = (C1, C2, C3), where C1, C2, C3 are the number of nodes in each linear transformation layer. All of this is shown in Figure 5.4a, where k=16 for all three blocks, and C for the EdgeConv blocks are (64, 64, 64), (128, 128, 128), and (256, 256, 256), respectively[32]. In parallel to the EdgeConv operation, the input features are passed directly to the end of the EdgeConv. These features are then added to updated features, which are the outputs of the three layers in the EdgeConv.



a The architectures of the ParticleNet

b The Edge Conv block structure

Figure 5.4

5.3 Deep Learning Software in KM3NeT

An open software was developed to facilitate the use of deep learning techniques in KM3NeT. This software is split into two Python packages:

OrcaSong [1], which is used for pre-processing KM3NeT files into a format suitable for machine learning use. This is done in two main steps:

• First, generation of hdf5 DL files: originally we have root files from which we extract hdf5 files, then we produce hdf5 DL files. DL files consist of two datasets, the samples x (in our case the full hit information for graphs generation), and the labels y contains labels and other event information. The configuration files define the kind of information in these datasets.

• Second, splitting the files into training and validation files: this is done according to the configuration file used. The validation set ,which is used to evaluate the performance of models with different hyperparameter values, is smaller than the training set. The training and validation files are then concatenated and shuffled to make sure that the order of events in the files is as random as possible.

OrcaNet [2], which does the training of the networks. At first, we construct the model from the configuration files, then we train the model. Three configuration files can be used to set up OrcaNet:

1. `list.toml`: Contains paths of training and validation DL (shuffled) files which were created in the previous step by OrcaSong.

2. `config.toml`: This file includes hyperparameters of the network, which will be used during the training, like the learning rate or the batchsize.

3. `model.toml`: Defines the architecture of the model. As mentioned in the previous section, the ParticleNet model is used.

The training is done over a number of epochs. After that, the prediction of the network can be saved to an hdf5 file, where the network with the lowest validation loss across the entire training history will be loaded automatically[24].

[1] https://git.km3net.de/ml/OrcaSong
[2] https://git.km3net.de/ml/OrcaNet

Chapter 6

Analysis and Results

In this thesis, KM3NeT/ARCA files are used to try to discriminate between Glashow events according to the decay of the W boson, hadronic or leptonic channels. This chapter reports on the details of how the analysis was conducted and shows the results of this analysis.

6.1 Analysis

As mentioned in chapter 5, the analysis is conducted using deep learning techniques. The deep learning software that was developed in KM3NeT consists of the OrcaSong and OrcaNet python packages that are used to extract the files, build and train the network which then would be able to predict the desired quantities (the type of the events). All of this is summarized in Figure 6.1.

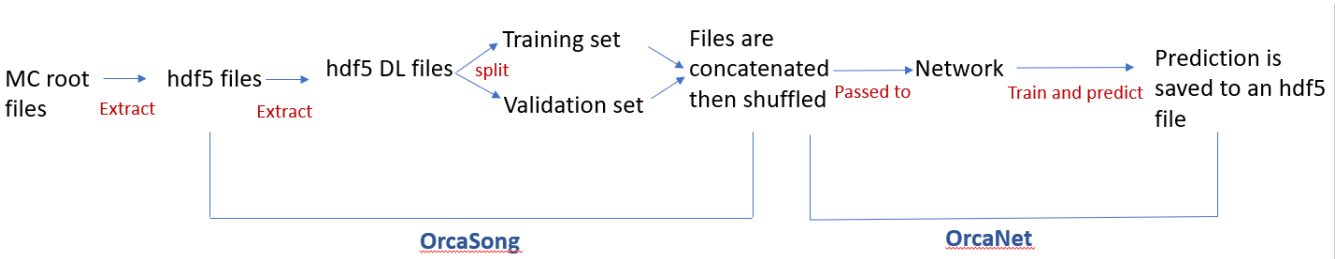


Figure 6.1 – The steps that need to be done to finally have a network that can classify the events, starting from the root files up to saving the prediction of the network.

The extraction of hdf5 files from the KM3NeT root files is done on the trigger or reconstruction level since at these levels hits have been triggered. Therefore, the data can be represented by graphs, which are then the inputs of our network. The output files (hdf5) are saved to the indicated output path.

OrcaSong is then used to extract hdf5 DL files from these hdf5 files. This is done by searching for hdf5 files in the input directory, and then running orcasong on each of them producing an hdf5 DL file. The output files are stored in the indicated output path. The OrcaSong configuration file is passed as one of the command line inputs. This configuration file specifies the mode, which is "graph" in our case, since GNNs are being used, and the extractor which indicates the function that extracts information from an event (information contained in x and y) in the h5 files. These functions are defined in OrcaSong [1](OrcaSong/orcasong/extractors/ neutrino_chain.py). For our analysis, an additional function "get neutrino mc info extr glashow (input_file)" was defined. This function classifies Glashow resonance events according to the decay of the W boson, hadronic or leptonic indicating the decay mode for each one of them. In MC truth the particles produced by each event are defined in an array where the first particle is the primary neutrino, the second particle is the target and so on. For Glashow resonance events the 5th particle is the W boson which then decays. If the 6th particle which is the result of this decay is an electron, a muon or a tau then the event is classified as leptonic event, otherwise it is identified as a hadronic event. The OrcaSong configuration file can also contain additional information such as the maximum number of hits (=2000). This is defined as the limit for the number of hits that can be supplied for each event, which is due to computational limitation where reducing the number of hits per graph reduces the memory burden. If the number of hits in events exceeds the maximum, they are randomly removed until the threshold is reached, where the number of hits that need to be removed on average depends on the number of hits in each event. In order to reduce the number of hits, a time cut is applied. In our case, hits outside the time window of [-1000,5500] ns are removed since according to previous studies, most of signal hits coming from Glashow resonance decay modes are inside this interval. This time window is also defined in the configuration file.

[1]https://git.km3net.de/ml/OrcaSong

The hdf5 DL files are split into training and validation files. This is achieved by defining another configuration file which splits the datasets, and then running the command `orcasong make_data_split config.toml`. After that, the training and validation files are concatenated and shuffled to ensure that the order of events is as random as possible. The shuffling is only necessary for training files since the order of events in the validation files, where the weights are constant, is not important.

OrcaNet is used to build and train the network, then the prediction of the fully trained network can be saved to an hdf5 file. In order to start the training of a new model, an OrcaNet directory is created in which the three following configuration files are defined:

1. `list.toml`: Contains paths of training and validation DL (shuffled) files created in the previous step by OrcaSong.

2. `config.toml`: Defines the hyperparameters of the network. For the model used in this study, batchsize = 32 and learning rate = [0.025,0.02], which were optimized in previous studies. In Classification tasks the label is required to be in one-hot encoding, which is a technique used to represent categorical data as vectors of zero elements except for one element, which is set to 1. For example, the classes {1, 2, 3} need to be encoded as vectors (1, 0, 0), (0, 1, 0), (0, 0, 1). This is done in OrcaNet by using user-defined functions known as modifiers. The one-hot encoding function is already implemented in OrcaNet and can be used by specifying it as a string in the config.toml file, where the corresponding string is "ClassificationLabels"[24]. The classes are class1 = 0 corresponding to hadronic decays, and class2 = 1 which corresponds to leptonic decays.

3. `model.toml`: Defines the architecture of the model which is based here on the ParticleNet model. For this network, the EdgeConv block is defined by k nearest neighbors=10 for all three blocks, and the number of channels C for the EdgeConv blocks, which is defined in section 5.2, is (64, 64, 64), (128, 128, 128), and (256, 256, 256), respectively. Batch normalization and pooling are applied.

   The optimizer that is used in our model is 'adam' optimizer, which aims to speed up the convergence of network training by varying the step size, where it adapts to the current situation of convergence, and allows different free parameters to have different effective step sizes. The cost function applied here is categorical cross entropy since it's a classification task.

Running the command line tool "`orcanet train name of directory`" will by default train for an unlimited number of epochs, unless the number of epochs is indicated (`- -to_epoch number of epochs`). If the training is interrupted it can resume from the last auto-generated checkpoint by running the "`orcanet train name of directory`" command again. The prediction of the network then can be saved to an hdf5 file by running the command "`orcanet predict name of directory`", then the network with the lowest validation loss across the entire training history will be loaded automatically.

6.2 Results

In the analysis, 48 root files were used with a total number of events of 142908 (weighted total number of events is $\sim$ 42046). The energy distribution of these events, which is the energy of primary antineutrinos for the Glashow resonance interaction, ranges between $10^5 - 10^8$ GeV and is shown in Figure 6.2. One can see that both event classes have almost equal energy distributions, and the Glashow resonance is visible around 6.3 PeV (as expected), this resonance leads to asymmetric distributions.

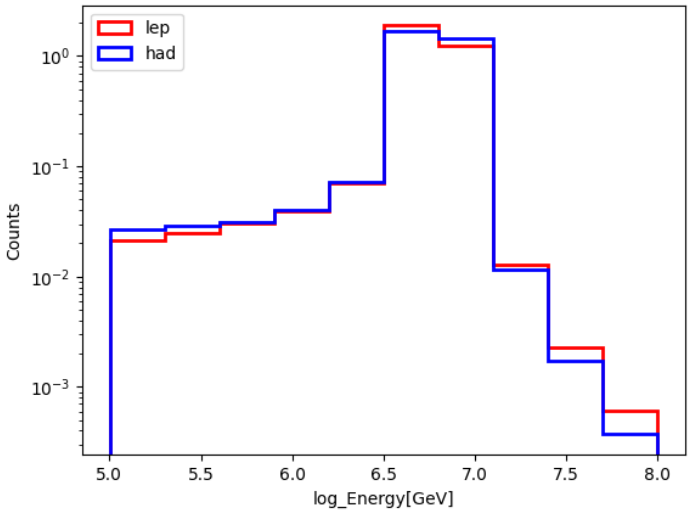


Figure 6.2 – The energy for both leptonic and hadronic events.

These files are then split into training files (80%) and validation files (20%). The network is trained, and its prediction is saved. The loss of this network for both training and validation, which assesses how well the network is working, is shown in Figure 6.3. This network was trained for 45 epochs and validation was applied after every epoch. As seen in this Figure, the value of the loss is decreasing towards a minimum loss as the network trains for more epochs. After 36 epochs the loss value is hardly changing indicating that the network converges. The minimum training set size of GNNs is not a fixed value, it depends on many factors like the complexity of the problem and the chosen algorithm. However, it is concluded that this is sufficient data for a meaningful training based on Figure 6.3, where there is no underfitting or overfitting observed (where they might indicate that the size of the training dataset that is used is not enough).

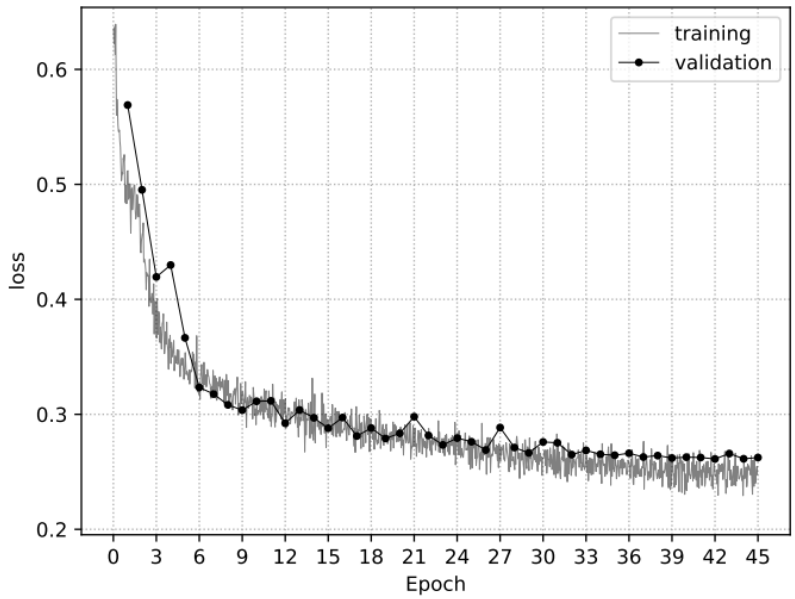


Figure 6.3 – The loss for both training and validation of a network trained for 45 epochs. The validation is done after every epoch.

The performance of the network can also be assessed by quantities that are not differentiable with respect to the model's weights, these quantities are known as metrics in Deep Learning. Accuracy is a metric that is used in binary classification tasks, which in our case is defined as the number of correctly classified events with respect to the total number of events. If the prediction of the network is the same as the true label of the event, it is defined as true positive (TP). In case they don't match, then it is defined as false positive (FP). In an analogous way, one can define true negatives (TN) and false negatives (FN). The accuracy can be calculated by

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6.1}$$

Figure 6.4 shows the accuracy of the network, where it increases for higher number of epochs. Finally, it reaches a value of ~ 0.9 (high accuracy) for 45 epochs.

The prediction of the network gives the result shown in Figure 6.5, which is the leptonic score for both hadronic and leptonic events. A threshold (or a cut) is chosen for this binary classification above which the prediction is defined as positive, where positive and negative terms indicate how the sample was labeled by the classifier (hadronic events=0 and leptonic events=1).

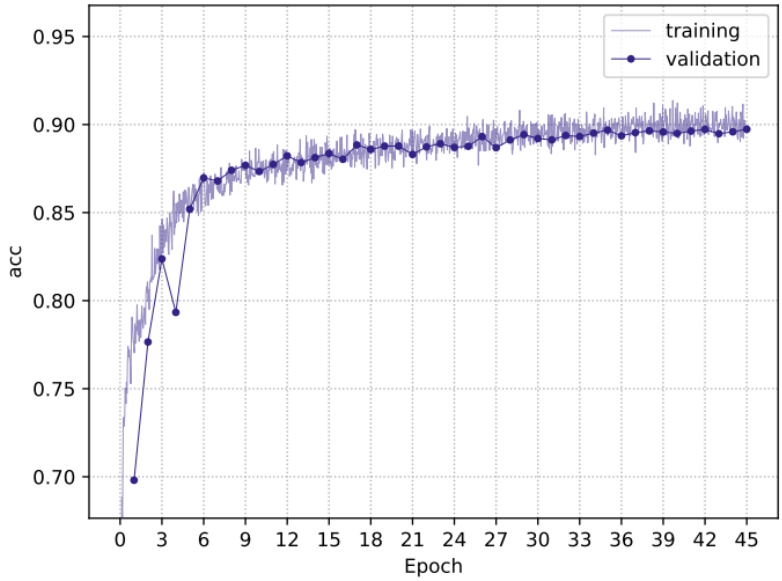


Figure 6.4 – The accuracy for both training and validation of a network trained for 45 epochs. The validation is done after every epoch.

For a specific cut (e.g. leptonic score= 0.4), the prediction of the network is interpreted as a leptonic event for prediction values larger than this cut, or as a hadronic event otherwise.

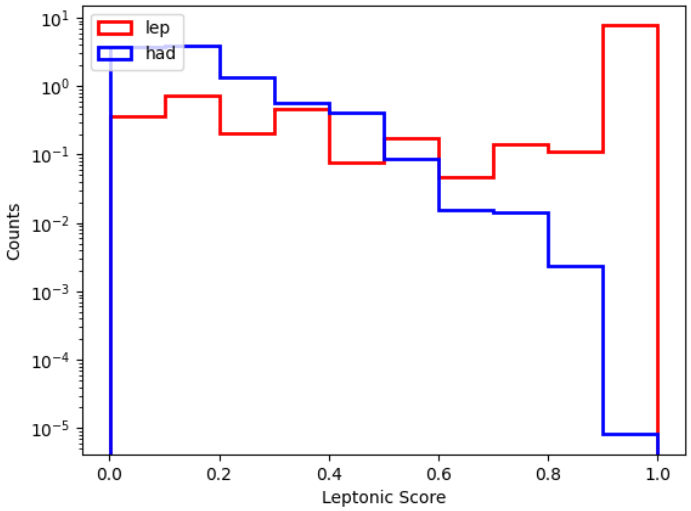


Figure 6.5 – The leptonic score for hadronic and leptonic events.

In Figure 6.5, the leptonic score for hadronic events mostly has low values while it is approximately homogeneous for leptonic events (i.e., leptonic events have sufficient events across all leptonic score values). Nevertheless, the discrimination between the two event classes is possible.

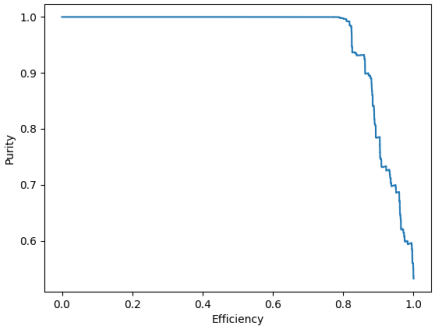
Two other metrics can be quantified for each cut from Figure 6.5: efficiency, which is the fraction of correctly classified events, and purity is defined as the fraction of classified events being correctly classified. Mathematically, efficiency and purity are calculated by

$$Efficiency = \frac{TP}{TP + FN} \tag{6.2}$$

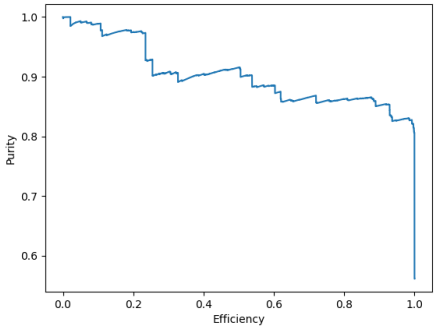
$$Purity = \frac{TP}{TP + FP} \tag{6.3}$$

Efficiency is also defined as the probability that a positively labeled sample will be predicted as positive, and purity is the probability that a positively predicted sample is in fact a true positive.

Both purity and efficiency depend on the cuts. Figures 6.6 a and b show this dependency of the purity and the efficiency for both leptonic and hadronic events, respectively. There is a tradeoff between purity and efficiency. For leptonic events, at efficiency higher than ~ 0.8 the purity starts decreasing while hadronic events of purity values less than ~ 0.8 the efficiency is 1.



a Leponic Events.



b Hadronic Events.

Figure 6.6 – Purity vs efficiency.

The leptonic score for all four possible decay modes of the W boson is shown in Figure 6.7. The total number of events is 142908 which consists of 11407 Glashow resonance events where the W boson decays into an electron and an electron antineutrino, 50207 Glashow resonance events in which the W boson decays into a muon and a muon antineutrino, 13684 Glashow resonance events with the W boson decaying into a tau and a tau antineutrino, and 67610 events for the hadronic channel of the W boson decay. The leptonic events add up to 75298 events, which differs only slightly from the number of hadronic events. The number of events in which the W boson decays into $\mu + \bar{\nu}_\mu$ is higher than the number of each of the remaining leptonic events.

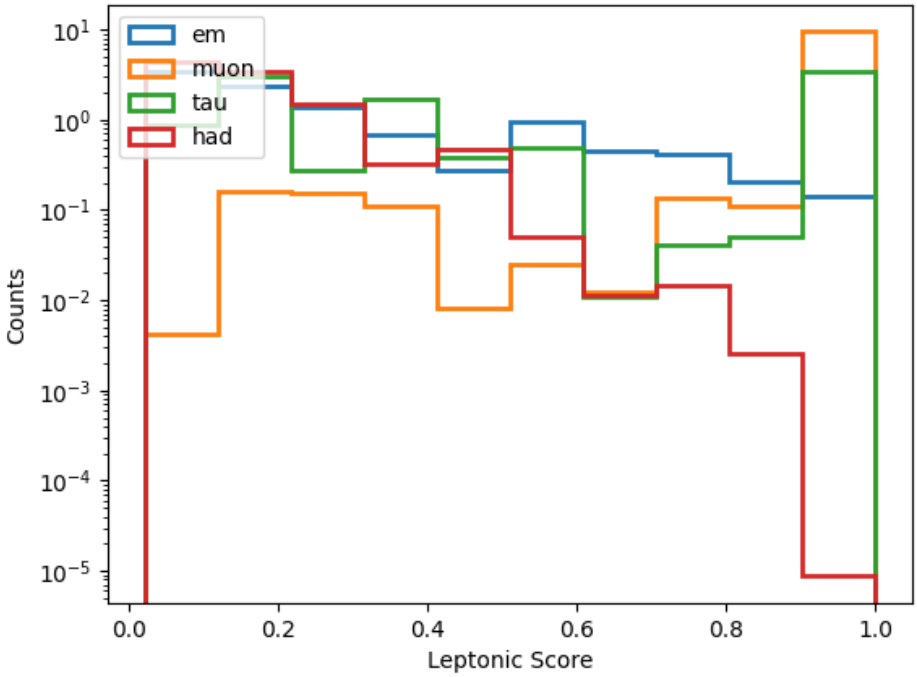


Figure 6.7 – The leptonic score for all events, where em indicates electromagnetic showers, muon indicates events of W boson decays into a muon and a muon antineutrino, tau stands for decay modes of W boson into a tau, and had is for hadronic showers.

6.3 Running Checks on The Network

The results show that the network can discriminate between the two event classes. However, three checks were done on the network to ensure it is working properly.

6.3.1 Check 1

The correlation coefficients between observables, which include primary observables (the number of triggered hits and the number of hits) and derived observables (reconstruction algorithms variables), as well as the leptonic score were calculated to see if the discrimination between both classes of events could be visible without the help of a GNN. This correlation coefficient is defined as

$$Corr = \frac{\Sigma(x - m_x)(y - m_y)}{\sqrt{\Sigma(x - m_x)^2 \Sigma(y - m_y)^2}} \tag{6.4}$$

where m_x is the mean of x and m_y is the mean of y.

The correlation coefficient can have a value between -1 and 1. This coefficient defines the strength and direction of the relationship between variables. For positive correlation, if one variable changes the other one changes in the same direction. If there is no relationship between the variables, the coefficient's values is 0 while negative values of correlation (anti-correlated) indicate that when one variable changes, the other variables change in the opposite direction.

Figure 6.8 shows the correlations between the observables with the addition of leptonic score. The red color has a value of 1, white indicates correlation value of zero, and blue is for -1 correlation value.

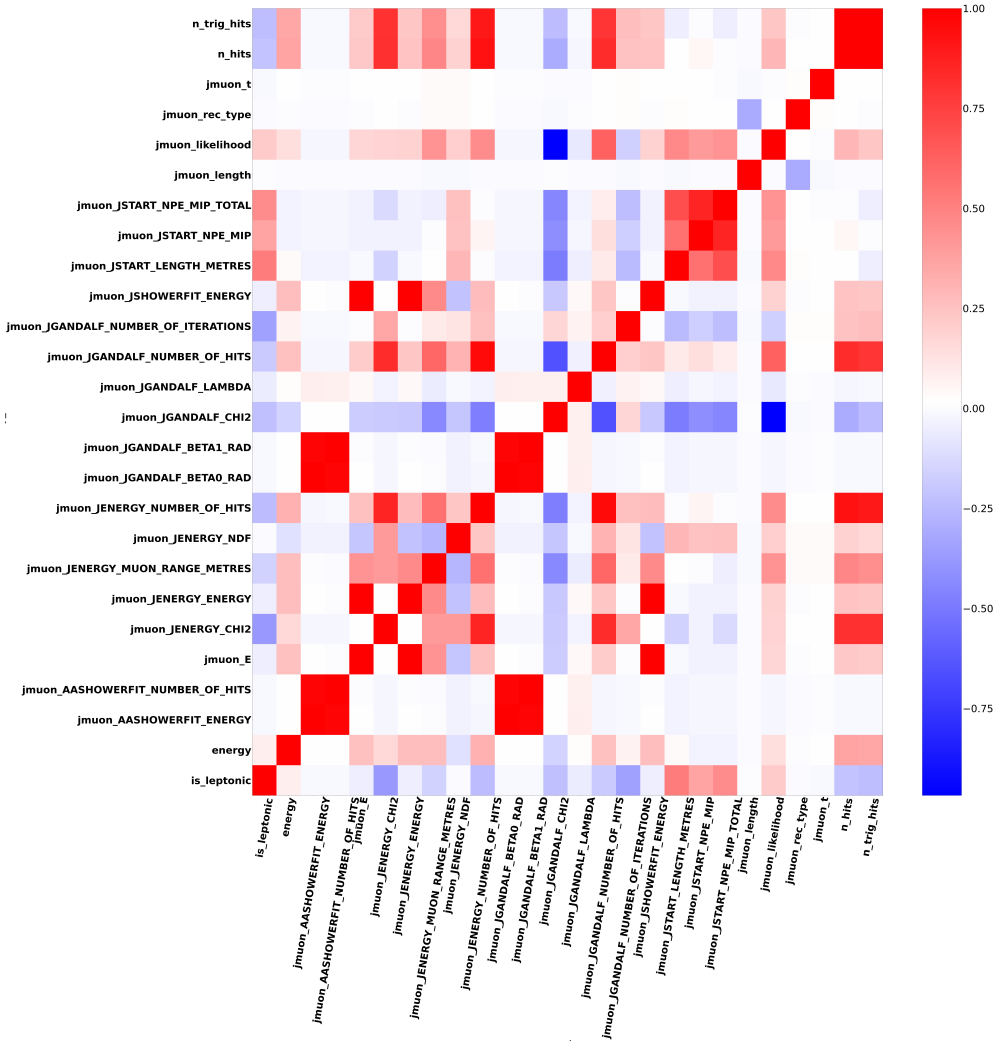
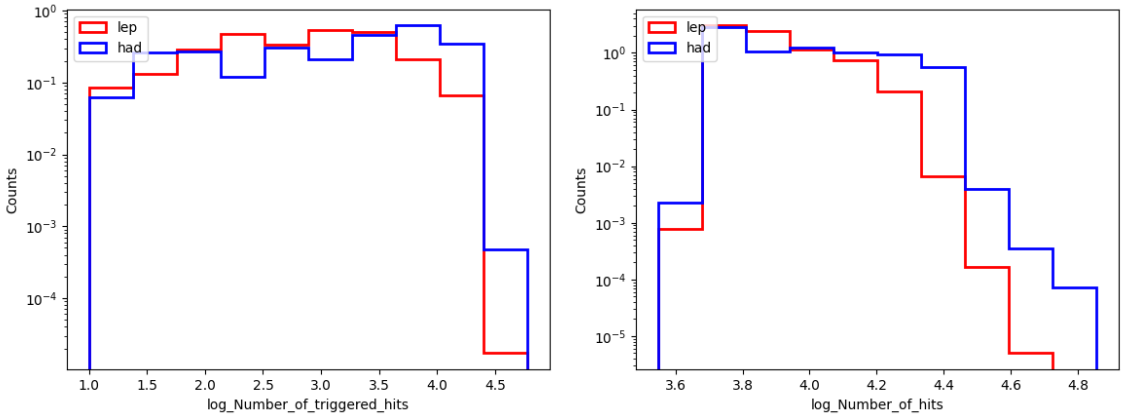


Figure 6.8 – The correlation between observables with the addition of leptonic score.

The most common color in Figure 6.8 is white indicating that there is no relationship between most of the observables. Figure 6.8 shows that some observables anti-correlate and correlate with the leptonic score (referred to in the Figure as is_leptonic which is the first bin of the matrix). 1D histograms for these observables were generated to investigate more.

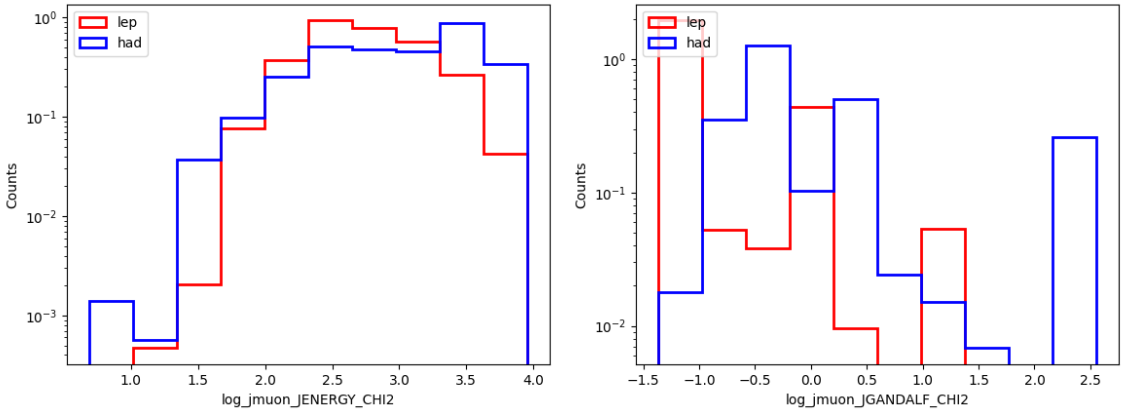
Figures 6.9a and 6.9b show the distributions of triggered hits and total number of hits for both classes. One can see that the distribution of triggered hits is almost uniform in both cases while for the total number of hits, the distributions are right skewed with slight different skew between both classes.



a **b**

Figure 6.9 – The number of both triggered hits and hits for leptonic and hadronic events.

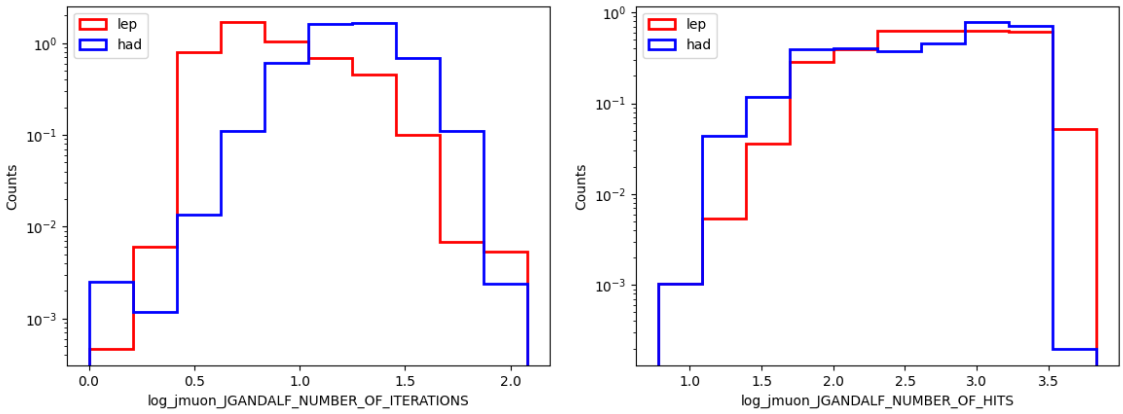
The jmuon_JENERGY_CHI2 parameter distribution for both hadronic and leptonic events is shown in Figure 6.10a, which appears to be left skewed. Figure 6.10b shows the jmuon_JGANDALF_CHI2 parameter distributions which seem to be different. However, it is hard to conclude whether these variables could help to discriminate between the two event classes.



a **b**

Figure 6.10 – 1D histograms of some derived observables for leptonic and hadronic events.

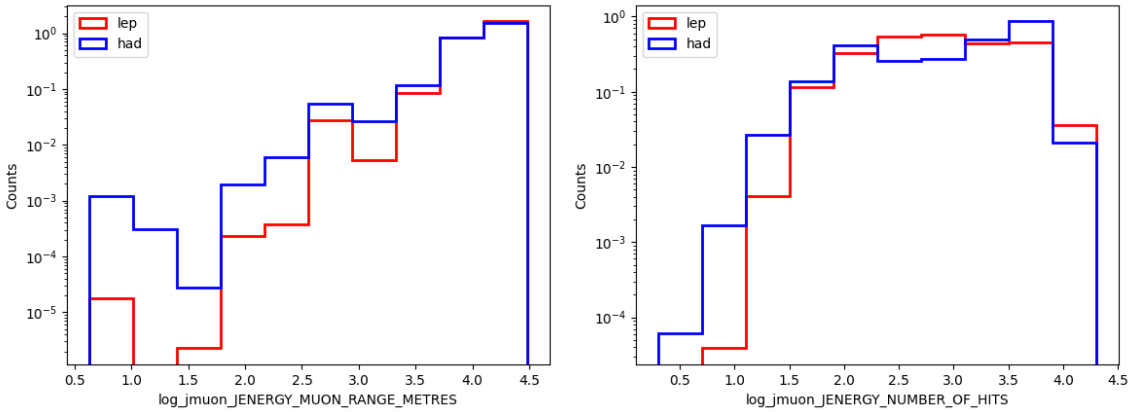
Figures 6.11a and 6.11b show the distributions for variables related to the muon reconstruction algorithms. For hadronic events in Figure 6.11a, the distribution is left skewed. On the other hand, the distribution is right skewed for leptonic events. The distributions in Figure 6.11b don't show any appreciable differences between leptonic and hadronic events.



a **b**

Figure 6.11 – 1D histograms of some derived observables for leptonic and hadronic events.

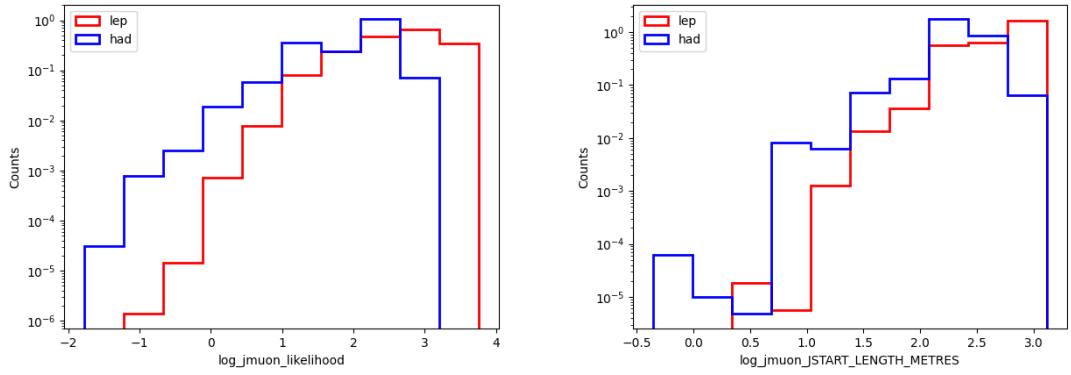
The distribution of the parameter shown in Figure 6.12a is left skewed. Figure 6.12b shows no differences between the distributions of different event classes.



a **b**

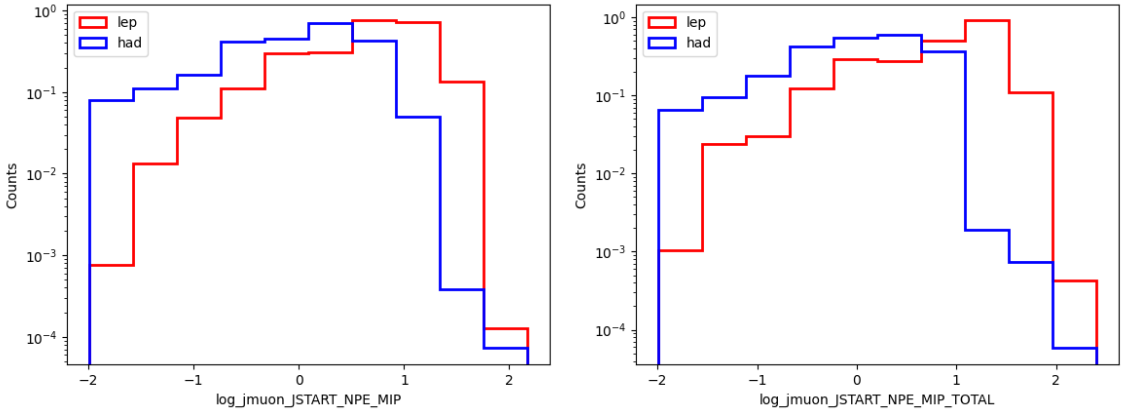
Figure 6.12 – 1D histograms of some derived observables for leptonic and hadronic events.

Figures 6.13a, 6.13b, and 6.14 for different derived observables distributions show a shift between the two distributions of the event classes. The shift is not so evident for Figure 6.13b, but one can still see a difference between the two classes.



a **b**

Figure 6.13 – 1D histograms of some derived observables for leptonic and hadronic events.



a **b**

Figure 6.14 – 1D histograms of some derived observables for leptonic and hadronic events.

An additional variable was calculated for both MC and reconstructed events, the sphericity. It is an event shape variable that describes how spherical is the event. An isotropic event has a sphericity=1 while a 2-jet event corresponds to sphericity=0.

The sphericity tensor is calculated in terms of the four-momentum vectors of the particle $p_i = (\vec{p}_i, E_i)$ with i=1,2,...,n, where n is the number of particles in the event[34]. The sphericity tensor is

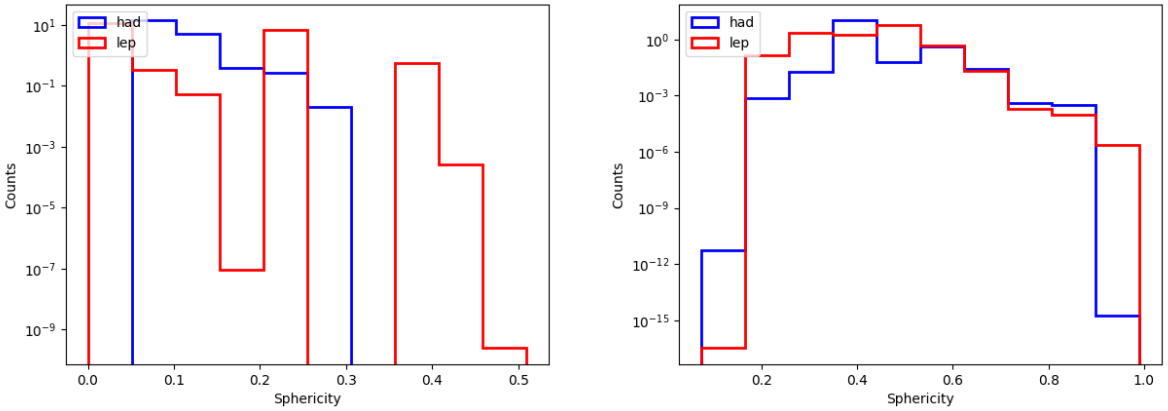
$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i \mid \vec{p}_i \mid^2} \tag{6.5}$$

where $\alpha, \beta = 1, 2, 3$ are the x, y and z components. By diagonalizing $S^{\alpha\beta}$, three eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are found, where $\lambda_1 + \lambda_2 + \lambda_3 = 1$. The sphericity of the event is then defined as

$$S = \frac{3}{2}(\lambda_2 + \lambda_3) \tag{6.6}$$

so that $0 \leq S \leq 1$.

The sphericity is calculated for both event classes to see if it can be used to discriminate between hadronic and leptonic events, where hadronic showers with 2-jets have sphericity values close to zero while electromagnetic showers sphericity values are close to 1 since they are expected to be more isotropic.



a The sphericity. **b** The reconstructed sphericity.

Figure 6.15 – The Sphericity for both leptonic and hadronic events.

The momentum for this sphericity in Figure 6.15a is calculated from the x, y, and z directions of the mc_tracks of the events.

Figure 6.15a shows the distribution of the sphericity which shows a difference between both classes. As for the reconstructed sphericity, the momentum is calculated from the difference between the x, y, and z positions of both hits and reconstructed tracks of the events. This method of calculating the reconstructed sphericiy is taken from previous studies done on the tau neutrinos. The distribution of the reconstructed sphericity is shown in Figure 6.15b. One can see the sphericity is distributed over the full range of 0 and 1 for both classes of the events and there are no strong differences between the two event classes.

To summarise, there's not a single variable that could be used to achieve the discrimination power observed in the results of the GNN. Nevertheless, there were appreciable differences between the distributions of leptonic events and hadronic events for a few variables which are mostly related to the muon reconstruction algorithms. This seems to indicate that one of the event classes is more similar to muon events than the other. To confirm this, a second check was performed which is explained in section 6.3.2.

6.3.2 Check 2

The branching ratios of each of the decay modes mentioned in section 5 were calculated. The resulting unweighted, weighted branching ratios of the decay channels of the W boson are shown in Figure 6.16 respectively, which are then compared with the expected branching ratios. As seen in this Figure, The resulting branching ratios differ from the expected ones.

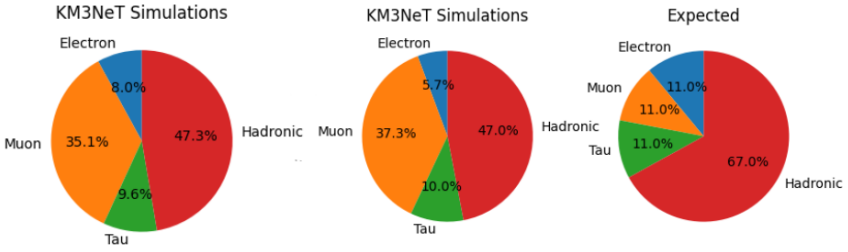


Figure 6.16 – The branching ratios (BR) of the decay channels of the W boson. The KM3NeT Simulations unweighted and weighted branching ratios are calculated from the root files, where the left one is the unweighted, and the one in the middle is the weighted.

Most of the contribution in the resulting branching ratios of the leptonic channels ($\sim$ 52.7 %) is from muons, as previously stated. The large number of produced muons could be the reason why the network was able to discriminate between leptonic and hadronic events since muons have a track like signature which is different (easier to distinguish) from the shower like ones expected from other events.

To investigate whether the number of these track events affect the performance of the network, the network was trained again with the exclusion of events where the W boson decays into $\mu + \bar{\nu}_\mu$. The results are shown in Figure 6.17. This Figure indicates that the network is still able to discriminate between the two different event classes, where the leptonic score for hadronic events mostly has low values while for leptonic events, it is approximately homogeneous but has few more events in the low leptonic score region in comparison with the leptonic score where all decay modes of the leptonic channel of W boson were included. However, if the distribution is flat beyond 0.5, but the increasing statistical uncertainty does not allow us to draw this conclusion. Then this would indicate that the discrimination is

not possible, which leads to the fact that the network is potentially a track vs shower classifier.

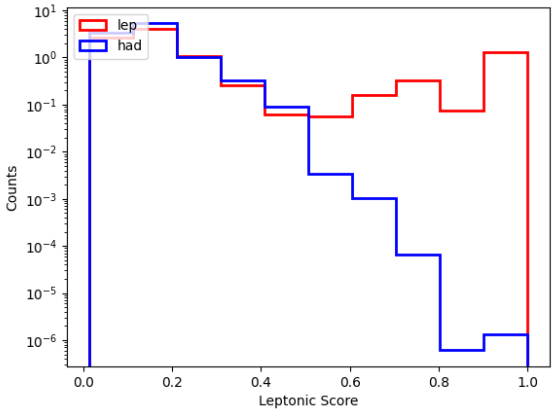


Figure 6.17 – The leptonic score for hadronic and leptonic events (without events where the W boson decays into $\mu + \bar{\nu}_\mu$).

6.3.3 Check 3

The last check was training the network again but this time with the exclusion of hadronic events to distinguish between leptonic and leptonic events, where the total number of events used was 9316. Figure 6.18 shows that the network was unable to distinguish between them, and therefore is working as expected.

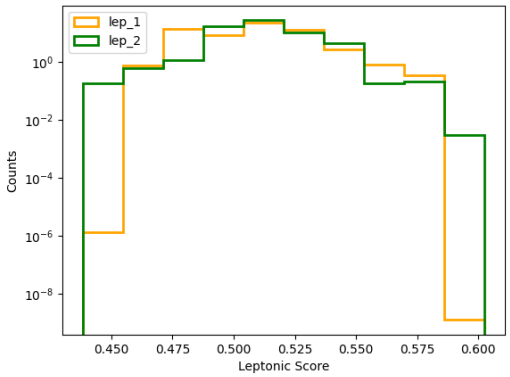
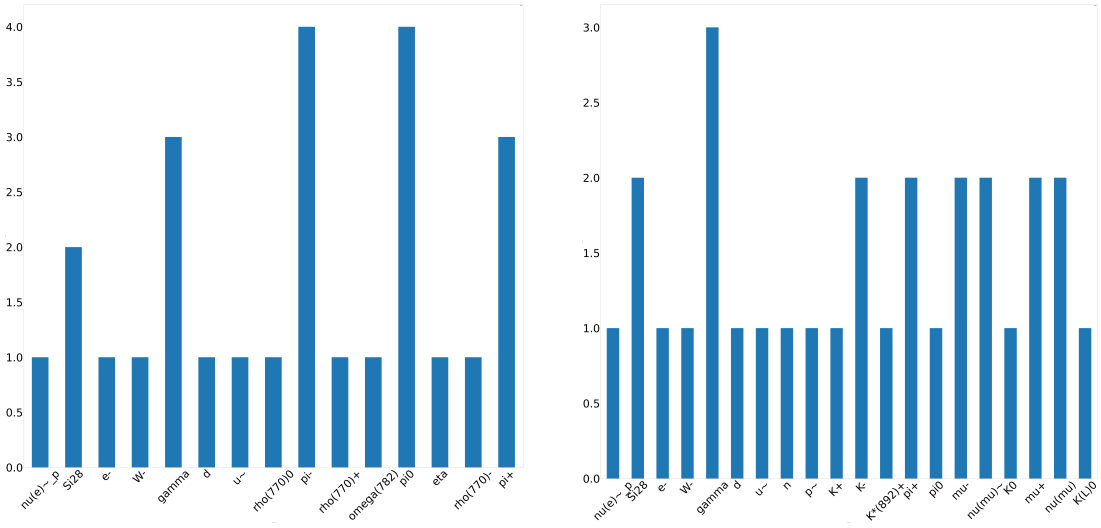


Figure 6.18 – The leptonic score of leptonic and leptonic events.

6.4 Other Investigations

The decays of pions and kaons are not included in this simulation, where these decays produce early hits by muons. Unfortunately, there was no time to prepare these simulations for training. However, a small set of simulations were produced where it was possible to include these decays. This is shown in Figure 6.19, where the number of particles for a specific event were counted for both decays disabled and decays enabled cases. As seen in this figure, the enabled case has more particles due to the additional particles produced from these decays including muons. However, the number of particles produced by the decays are not as excpected. For example, in the case where decays are disabled, there are four π^-, which then decay into two $\mu + \bar{\nu}_\mu$. The number of particles doesn't add up, which needs to be further investigated. Despite the fact that the first simulation used in our network didn't enable these decays, the network was able to discriminate between the two event classes.



a Decays disabled. **b** Decays enabled.

Figure 6.19 – The number of particles for a specific event, for both the enabled and disabled cases. The x-axis shows the particles and y-axis is the number of particles.

In addition, the branching ratio values produced by the current simulation are wrong. Figure 6.20 shows branching ratio values, where the branching ratio of the hadronic events in this new simulation (the one in the middle) is higher than that of the old simulation (the one on the left). The branching ratios of the new simulation are still a little bit different from that of the expected ones as well. Further investigations should be done in this regard.

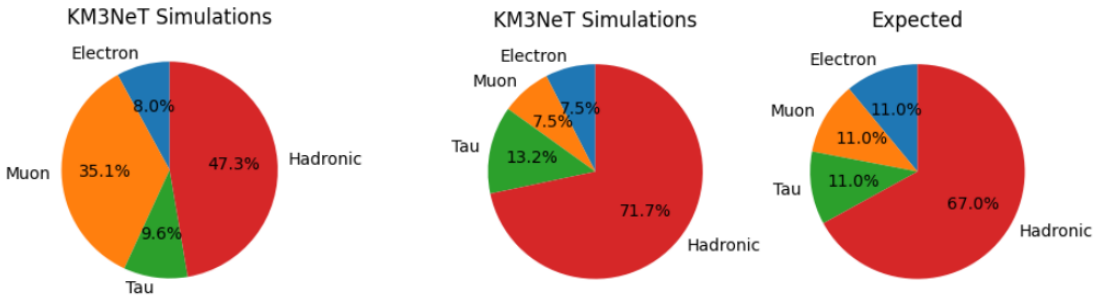


Figure 6.20 – The branching ratios of W boson decays. The old simulation branching ratios are the ones on the left. In the middle, the new simulation branching ratios are shown, and the expected branching ratios are shown on the right.

Chapter 7

Summary and Outlook

In this thesis, investigations of the capabilities of KM3NeT/ARCA to discriminate between Glashow events according to the decay of the W-boson, hadronic or leptonic cascades, were conducted using deep learning techniques. The data were represented by graphs which are then used as an input to a network (GNN). The results showed that the network was able to discriminate between the two event classes. The network was checked using three approaches to ensure that it was working properly. The first check showed that by looking into the correlations between observables and the leptonic score, there's no evidence that a single observable allows to discriminate between both classes of events. Nevertheless, a few of them showed small differences which combined could explain the results of the GNN. The second check in which we excluded the events where the W boson decays into $\mu + \bar{\nu}_\mu$, showed that the network can still discriminate between leptonic and hadronic events. In the last check, the events where the W boson decays into hadrons were excluded. This was done to see if the network would be able to distinguish between leptonic and leptonic events, which was not possible, thus indicating that the network is working as expected. In the end, none of the checks indicated that the network should be working otherwise.

The next step would be to produce new simulations that include the decays of pions and kaons, train the network using these simulations, and see if the network would be able to discriminate. Also, the wrong branching ratio values for the decay modes of the W boson and the number of particles after enabling the decays need further investigations.

Bibliography

[1] A. D. Dolgov. Cosmology and neutrino properties. *Physics of Atomic Nuclei*, 71(12):2152–2164, dec 2008.

[2] J. A. Formaggio and G. P. Zeller. From eV to EeV: Neutrino cross sections across energy scales. *Reviews of Modern Physics*, 84(3):1307–1341, sep 2012.

[3] M. Sajjad Athar, A. Fatima, S. K. Singh. Neutrinos and their interactions with matter. *The European Physical Journal Special Topics*, 2022.

[4] Th. M. Nieuwenhuizen . Do non-relativistic neutrinos constitute the dark matter? *Europhys.Lett.86:59001*, 2008.

[5] X. Qian, P. Vogel. Neutrino mass hierarchy. *PPNP, 83, 1-30*, 2015.

[6] Ulrich F. Katz, Christian Spiering. High-energy neutrino astrophysics: Status and perspectives. 2011.

[7] T.K. Gaisser, M. Honda. Flux of atmospheric neutrinos. *Ann.Rev.Nucl.Part.Sci.52:153-199*, 2002.

[8] J.L. Autran, D. Munteanu, T. Saad Saoud, and S. Moindjie. Characterization of atmospheric muons at sea level using a cosmic ray telescope. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 903:77–84, 2018.

[9] Naoko Kurahashi, Kohta Murase, Marcos Santander. High-energy extragalactic neutrino astrophysics. *Ann.Rev.Nucl.Part.Sci. 72 (2022) 365*, 2022.

[10] IceCube Collaboration. Evidence for neutrino emission from the nearby active galaxy ngc 1068. *Science 378, 6619, 538-543*, 2022.

[11] Erik Ganster, Richard Naab, Zelong Zhang (for the IceCube Collaboration). A combined fit of the diffuse neutrino spectrum using icecube muon tracks and cascades. *Proceedings of Science*, 395, 2022.

[12] V. Barger, Lingjun Fu, J.G. Learned, D. Marfatia, S. Pakvasa, T.J. Weiler. Glashow resonance as a window into cosmic neutrino sources. *Phys. Rev. D 90, 121301*, 2014.

[13] P. Abreu et al. Search for ultrahigh energy neutrinos in highly inclined events at the Pierre Auger Observatory. *Phys. Rev. D 84, 122005*, 2012.

[14] the KM3NeT Collaboration. Letter of Intent for KM3NeT 2.0. *Journal of Physics G: Nuclear and Particle Physics, 43 (8), 084001*, 2016.

[15] Rodrigo Gracia Ruiz. *Search for populations of unresolved sources of high energy neutrinos with the ANTARES neutrino telescope.* Thesis, Université Sorbonne Paris Cité, November 2016.

[16] Ellen Riefel. Bioluminescence in the KM3NeT neutrino telescope. Bachelor's thesis. Leiden University, The Netherlands. 2017.

[17] G. Carminati and M. Bazzotti and A. Margiotta and M. Spurio. Atmospheric MUons from PArametric formulas: a fast GEnerator for neutrino telescopes (MUPAGE). *Computer Physics Communications*, 179(12):915–923, 2008.

[18] Mona Dentler. Investigation of the one-particle approximation in the ANTARES simulation package KM3, Bachelor's thesis, Erlangen Centre for Astroparticle Physics Friedrich-Alexander-Universität Erlangen-Nürnberg. 2012.

[19] M. de Jong. Jsirene: A program to simulate the detector response. Jpp documentation. url: https://sftp.km3net.de/documentation/Jpp/v15.0.0-rc.2/JSirene.PDF.

[20] Sheldon L. Glashow. Resonant scattering of antineutrinos. *Phys. Rev. 118, 316*, 1960.

[21] Guo-yuan Huang, Qinrui Liu. Hunting the Glashow resonance with PeV neutrino telescopes. *JCAP03(2020)005*, 2019.

[22] Jerzy Mańczak. Sensitivity study for the Glashow resonance detection at KM3NeT. *MSc. thesis. University of Warsaw Faculty of Physics*, 2018.

[23] IceCube Collaboration. Detection of a particle shower at the Glashow resonance with icecube. *Nature 591, 220-224*, 2021.

[24] Stefan Reck. *Investigating systematics for KM3NeT/ORCA using unsupervised Deep Learning*. Thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU).

[25] Johannes Lederer. Activation functions in artificial neural networks: A systematic overview. *ArXiv*, abs/2101.09957, 2021.

[26] David E. Rumelhart, Geoffrey E. Hinton Ronald J. Williams. Learning representations by back-propagating errors. *Nature 323, 533–536*, 1986.

[27] Bing Xu, Naiyan Wang, Tianqi Chen, and Mu Li. Empirical evaluation of rectified activations in convolutional network. *ArXiv*, abs/1505.00853, 2015.

[28] X. Zhong, David Lee Enke. Predicting the daily return direction of the stock market using hybrid machine learning algorithms. *Financial Innovation, vol. 5, no. 1, SpringerOpen*, 2019.

[29] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, Maosong Sun. Graph neural networks: A review of methods and applications. 2018.

[30] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? *ArXiv*, abs/1810.00826, 2018.

[31] S. Reck, D. Guderian, G. Vermariën, A. Domi and on behalf of the KM3NeT collaboration. Graph neural networks for reconstruction and classification in KM3NeT. *Journal of Instrumentation, Volume 16*, 2021.

[32] Huilin Qu, Loukas Gouskos. ParticleNet: Jet Tagging via Particle Clouds. *Phys. Rev. D 101, 056019*, 2019.

[33] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning*, 2015.

[34] Torbjorn Sjostrand. PYTHIA 5.7 and JETSET 7.4 Physics and Manual. 1995.

[35] J. A. Formaggio, G. P. Zeller. From ev to eev: Neutrino cross sections across energy scales. *Rev. Mod. Phys. 84, 1307*, 2013.

[36] Anna M. Suliga. Diffuse supernova neutrino background. *Handbook of Nuclear Physics, Springer Nature Singapore*, 2022.

[37] IceCube Collaboration. Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector. *Science 342, 1242856*, 2013.

[38] Amit Loewy, Shmuel Nussinov, and Sheldon Lee Glashow. The Effect of Doppler Broadening on the 6.3 PeV W^- Resonance in $\bar{\nu}_e e^-$ Collisions. *arXiv: High Energy Physics - Phenomenology*, 2014.

[39] Guo-yuan Huang, Manfred Lindner, Nele Volmer. Inferring astrophysical neutrino sources from the Glashow resonance. *arXiv e-prints*, 2023.

[40] Hans Ulrich Schmidt. *Meßelektronik in der Kernphysik.* Teubner-Studienbücher. B. G. Teubner Stuttgart, 1 edition, 1986.

Acknowledgements

Thanks to everyone who supported me during this thesis. A special thanks goes to

- Rodrigo Gracia-Ruiz

Thank you very much!

Statement of Authorship

I hereby confirm that I have written this thesis independently and only with the help of the indicated sources and aids.

Erlangen, July 10, 2023

Yara Darras