Following the Light

Novel event reconstruction techniques for neutrino oscillation analyses in KM3NeT/ORCA

Brían Ó Fearraigh

Following the Light

Novel event reconstruction techniques for neutrino oscillation analyses in KM3NeT/ORCA

Brían Ó Fearraigh
Universiteit van Amsterdam

Following the Light – Novel event reconstruction techniques for neutrino oscillation analyses in KM3NeT/ORCA

Cover: My heartfelt gratitude goes to my friend (Dr.) Jessica Erkal for her wonderful cover illustration for this thesis.

This research work has been funded by the University of Amsterdam/Universiteit van Amsterdam, and has been carried out in the National Institute for Subatomic Physics (Nikhef) in Amsterdam, the Netherlands.

ISBN: 978-94-6496-019-8

Printed by: Gildeprint - The Netherlands

Following the Light

Novel event reconstruction techniques for neutrino oscillation analyses in KM3NeT/ORCA

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op donderdag 22 februari 2024, te 13.00 uur

door Brían Ó Fearraigh
geboren te Drogheda, County Louth

Promotiecommissie

Promotor:	prof. dr. ir. P.J. de Jong	Universiteit van Amsterdam
Copromotor:	dr. R. Bruijn	Universiteit van Amsterdam
Overige leden:	dr. E. Tzamariudaki	INP Demokritos
	prof. dr. C.W.J.P. Timmermans	Radboud Universiteit
	prof. dr. M.P. Decowski	Universiteit van Amsterdam
	dr. I.B. van Vulpen	Universiteit van Amsterdam
	dr. J.M. Sonneveld	Universiteit van Amsterdam
	prof. dr. M. de Jong	Universiteit Leiden

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Contents

Prelude

'An rud is annamh is iontach' is a saying in the Irish language, meaning that something that is rare is amazing, wonderful, beautiful. The word 'annamh' can also mean lonely.

The focus of this PhD thesis is on neutrino physics. Neutrinos are, in a sense, one of the rarest and loneliest particles that we know of.

Everything in the world that we can touch and interact with is made up of particles. We have all encountered the concept of atoms in school at some stage, these are the building blocks of all matter. Atoms themselves are made up of combinations of protons, neutrons and electrons, particles too small to be seen with the naked eye. The electron is a fundamental particle, not made up of anything else as far as we know. Neutrinos are fundamental particles even lighter than electrons. They come in three types. They are created in nuclear and particle decays, nuclear reactions, in all the stars in the sky, in the Sun, in the atmosphere, coming from the ground. Billions of neutrinos pass through you, me and everything else every second.

They are utterly, unimaginably tiny. For quite a while, neutrinos were believed to have no mass. Now, we know that neutrinos do indeed have mass, a property which allows them to 'oscillate' or change from one kind of neutrino to another in a regular rhythm. They only interact via a fundamental force of Nature termed the 'weak' interaction, and as the name suggests, the strength of this force is quite weak. The probability with which neutrinos interact with matter (like protons, neutrons, electrons) is therefore tiny. As a result, detecting neutrinos and measuring their properties is no easy task.

We do not need to give up hope yet though. Experiments are being carried out in vast detectors, to detect the results of neutrino interactions with matter and measure their properties. To achieve this, the larger the detection medium the better, which increases the probability of neutrino interactions. One such experiment is the KM3NeT (Cubic Kilometre Neutrino Telescope) experiment, which aims to contain and fill a huge volume of seawater in the Mediterranean Sea with photodetectors, to record the light from the products of neutrino interactions. KM3NeT is still under construction, consisting of two

separate detectors, which are currently between 12–16% of their envisaged, completed sizes. One detector, KM3NeT/ORCA, measures the properties of neutrino oscillations. By doing so, we can form a very accurate and complete description of how neutrinos oscillate between the three types. Furthermore, the masses of the three types of neutrino relative to each other are not known. KM3NeT/ORCA will carry out this measurement, and determine what we call the 'neutrino mass ordering'. KM3NeT/ARCA, the second detector, acts as a telescope. By estimating the direction and energy of detected neutrinos, KM3NeT/ARCA can pinpoint with extreme accuracy where in our Universe neutrinos are being created and coming to Earth from. These lonely particles offer us the chance to explore the Universe and add to our fundamental knowledge on how it works.

Thesis Outline & Contributions of the Author

Within the pages of this thesis, I will go into further detail on neutrinos, how they interact, and how we can detect them. I also describe how the KM3NeT experiment operates, and what I contributed to the science goals of the experiment with this research.

In **Chapter 1**, the physics behind cosmic rays (atomic nuclei which arrive to the Earth from outer space) and neutrinos is described. A brief history of the discovery of neutrinos is provided, followed by a description of their interactions, and the properties of neutrino oscillations and their masses. A summary is provided of the current measurements of the properties which describe neutrino oscillations.

In **Chapter 2**, the structure, technology, components and operations of the KM3NeT experiment is presented. The sources which induce signals in the detector are described, and the means by which the data signals are recorded, collected, and processed is outlined. Finally, the scientific research goals of the experiment are outlined, as well as the current status of the experiment.

In **Chapter 3**, an overview is given of the simulation software used within the KM3NeT experiment. Dedicated algorithms which reconstruct the direction and energy of detected neutrinos are described, as well as the use of photo-electron distributions as a function of the arrival time of photons, which are fundamental to the simulation and reconstruction software within KM3NeT. The simulation of neutrino interactions, cosmic ray interactions and atmospheric muons is also presented, with particular attention paid to the atmospheric muon simulation software, which forms the basis for the work carried out and described in Chapter 4. The description of the track reconstruction algorithm will form the basis for the work described in Chapter 5.

In **Chapter 4**, a tuning of the atmospheric muon flux simulation to KM3NeT data is presented. This work was entirely carried out by myself. I carried out a study into which parameters had the the largest impact on simulated physical observables in the KM3NeT/ORCA-6 detector, and define the statistical test used to compare these observable between data and simulation. I carried out the parameter scan and compared the data to the tuned simulation. A comparison is made between the results of the tuned simulation and for many KM3NeT/ORCA-6 data runs.

In **Chapter 5**, I developed a means of improving the direction estimate of track-like neutrino event signatures, particularly for the KM3NeT/ORCA detector. The motivation for this study is presented, followed by a description of using the length of the estimated track in its direction estimate, and how this is included in the track reconstruction algorithm. The direction and energy resolutions for track-like neutrino interactions in the complete, simulated KM3NeT/ORCA detector are presented, using this new procedure.

In **Chapter 6**, I introduce a novel **track+shower** reconstruction algorithm, which was developed and implemented by myself. The event interaction model and the likelihood function one can use to describe this event model is defined. I conduct a scan of the likelihood values for different model parameters in the simulated, complete KM3NeT/ORCA detector. A brief technical description of the algorithm is given, followed by the ability and performance of this new reconstruction technique in reconstructing track and shower-like events, and only shower-like events. The expected neutrino direction and energy resolutions are presented. In addition, an estimate of reconstructing the Bjorken-y parameter is presented, a parameter which can be directly reconstructed for the very first time by this track+shower reconstruction algorithm.

In **Chapter 7**, the novel track+shower algorithm is used to reconstruct neutrino interactions for data in the KM3NeT/ORCA-6 detector. In this chapter I explore how this new reconstruction technique can contribute to oscillation analyses using ORCA data. The distributions which result from the track+shower reconstruction are compared for different neutrino event types, with differences between the event types showing that the track+shower reconstruction algorithm can contribute to particle identification methods. A thorough and comprehensive study of the direction, energy and Bjorken-y resolutions for the various KM3NeT/ORCA-6 neutrino interactions is also carried in this chapter. I discuss how the reconstruction can be improved upon following these investigations. The work presented in this chapter is meant to form a basis for future studies and physics analyses using the track+shower reconstruction algorithm.

Relevance & Impact

The tuned atmospheric muon flux simulation offers a means of better describing and understanding the signal induced in the KM3NeT detectors by atmospheric muons. These atmospheric muons are studied in cosmic ray analyses by the KM3NeT experiment, and also form the main background signal in neutrino physics studies. The update to the track reconstruction software indicates a better reconstruction of the track direction for KM3NeT/ORCA, an essential part of studying neutrino oscillation parameters and measuring the neutrino mass ordering. The track+shower reconstruction technique offers a new means of reconstructing neutrino interactions, and is shown to improve the estimate of the neutrino energy in the KM3NeT/ORCA detector. This new algorithm also offers a means of directly reconstructing the Bjorken-y parameter. The work presented in this thesis, as carried out by the author, improves upon the reconstruction of neutrino interactions in the KM3NeT/ORCA detector. Improvements in the neutrino direction and energy resolutions are fundamental to fulfilling the KM3NeT experiment's potential and goal in measuring the neutrino mass ordering.

Chapter 1.

Cosmic Ray & Neutrino Physics

Two branches of physics in which there remain many unanswered questions are those of cosmic ray and neutrino physics. The origin of cosmic rays - atomic nuclei which hit the Earth from outer space - is still undetermined. Cosmic neutrinos, neutral subatomic particles likely created in the vicinity of the same sources as cosmic rays, also reach the Earth and point directly back to their origin, unaffected by magnetic fields they encounter en route. Both act as 'messengers' to study the astrophysical objects which produce them. Cosmic rays which interact with the Earth's atmosphere produce further neutrinos, as well as heavier subatomic particles such as muons. These atmospheric neutrinos offer the possibility of studying their fundamental properties. In this chapter we begin with a description of cosmic rays, proceed to their interactions in the atmosphere, and then follow the production of atmospheric neutrinos and muons. These particles will be detected and analysed using the KM3NeT experiment in the Mediterranean Sea, which will be described in later chapters.

This thesis describes research which covers both cosmic ray and neutrino physics, through a study of the models which describe atmospheric muon production and detection, and studies into improving the determination of fundamental neutrino properties.

1.1. Cosmic Ray Physics

1.1.1. Cosmic Rays in Space

The Earth's upper atmosphere is constantly being bombarded by particles. These particles, broadly termed *cosmic rays*, are protons and other atomic nuclei coming from both within and outside our Solar System. Some of these particles originate from the Sun, via

spontaneous emission in the form of giant plasma clouds called 'coronal mass ejections', or releases of energetic particles from the Sun's surface known as 'solar flares'. As a further distinction, the particles which come from the Sun are referred to as 'solar energetic particles' [1]. The term 'cosmic rays' is used to describe the particles which come from elsewhere in the Milky Way and beyond. However, the source of the high-energy cosmic rays observed thus far is currently one of the outstanding mysteries in astroparticle physics. Due to the enormous velocity and energy of some of these particles, their sources must be enormous and powerful astrophysical objects, none of which appear to be in the immediate vicinity of the Earth. The isotropy (arrival from all directions) of cosmic rays suggests an extra-galactic origin. No particle physics accelerator made by humankind can replicate the energies reached by these cosmic assailants. The sources of cosmic rays are termed 'cosmic accelerators', a broad term for the astrophysical bodies powerful enough to fling these particles outwards at velocities approaching the speed of light.

The majority of cosmic rays with energies between 1 GeV and 10^{11} GeV are protons, whilst cosmic rays of heavier nuclei are less frequent — cosmic rays are approximately 85% protons, 12% alpha particles (helium nuclei), 1% heavier nuclei (such as carbon, oxygen, nitrogen) and 2% electrons [2]. Their energies are highly dependent on the type of source and the propagation mechanism.

Figure 1.1 shows the all-particle cosmic ray spectrum in the range E (energy-per-nucleus) of 10^{13} – 10^{21} eV at the top of the atmosphere, as measured by various experiments. Here, the differential energy spectrum has been scaled by $E^{2.6}$, allowing for an exaggeration of the structure of the spectrum. The region between 10^{15} and 10^{16} eV is described as the 'knee'. Another steepening of the spectrum, known as the 'second knee', occurs at about $E \simeq 10^{17}$ eV. The change in the shape of the spectrum at $E \simeq 10^{18.5}$ eV is referred to as the 'ankle'.

These spectral features indicate different cosmic rays produced by different mechanisms. The rigidity R of a particle can be defined as

$$R = \frac{pc}{Ze}, \tag{1.1}$$

where p is the particle momentum, c is the speed of light, and Ze is the charge of the particle in multiples of the electron charge e. The rigidity relates the particle to influence by a magnetic field: the higher the rigidity/momentum of a particle, the less it is deflected by magnetic fields. Lower energy, lighter particles are produced more abundantly and closer-by (to the Earth) than heavier particles, which need very high

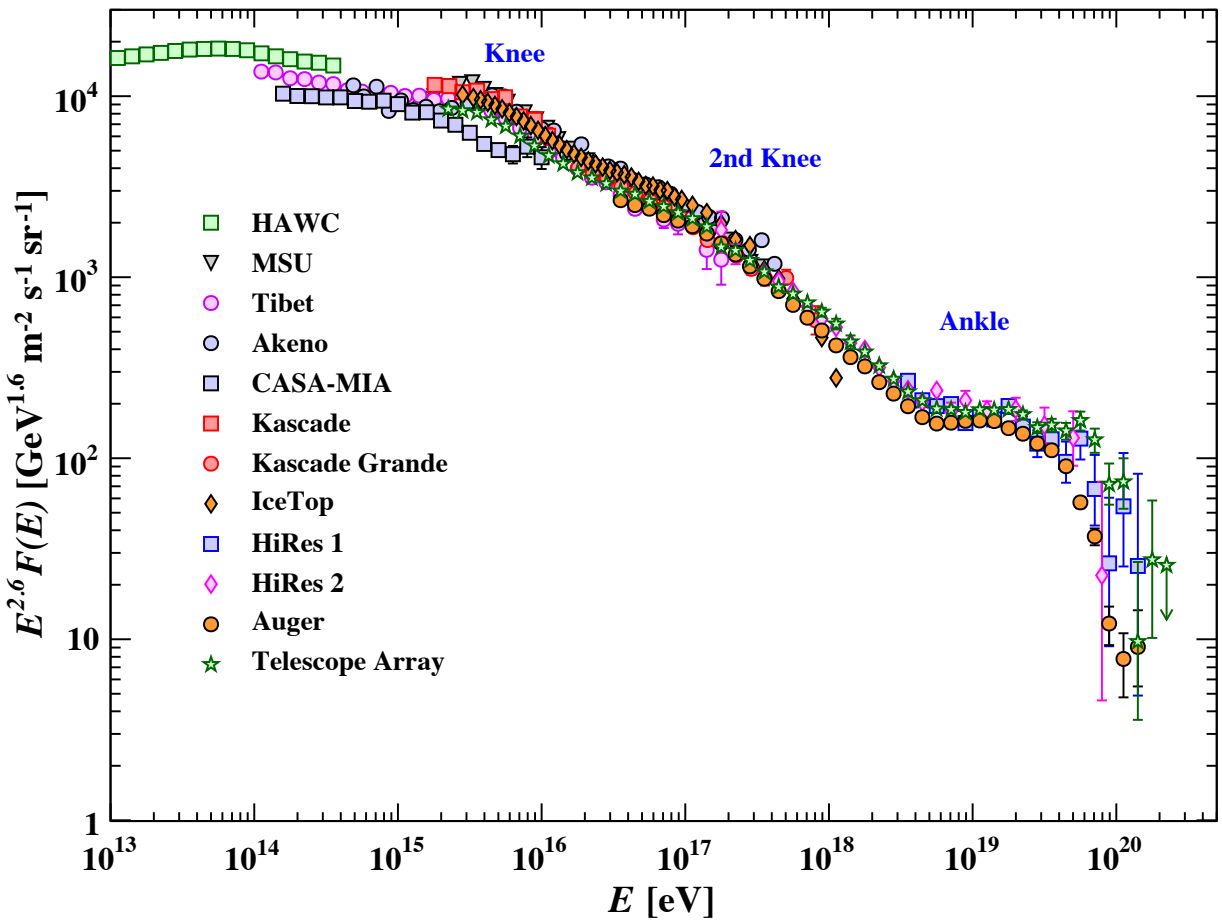


Figure 1.1.: The all-particle cosmic ray flux as a function of E (energy per nucleus) from various experiments. Taken from Workman et al. (2022) [3].

energies to be accelerated enough to reach the Earth and be detected. Thus, the features of this spectrum indicate different cosmic rays with rigidity-dependent acceleration mechanisms. As one follows the shape of the spectrum from the knee to the ankle, the composition changes, the cosmic rays become heavier, and the sources become further away.

Within the Solar System, solar energetic particles reach energies of $E \simeq 10^{10}$ eV. Beyond this, cosmic rays from $E \simeq 10^{10} - 10^{18}$ eV are expected to be of galactic origin — that is, coming from within the Milky Way [4]. The knee of Figure 1.1 can be interpreted as cosmic accelerators within the Milky Way reaching the maximum energy for the acceleration of protons. Both the knee and the second knee, i.e. a steepening of the spectrum, can be ascribed to a transition of the all-particle spectrum to heavier nuclei [5]. Beyond 10^{18} eV, where we see the ankle of the spectrum, the sources of these particles are considered to be extra-galactic. Further discussion of this spectrum can be found in Blasi (2013), van Dam (2021) [4, 6].

1.1.2. Cosmic Rays on Earth

The cosmic rays which collide with atoms in the upper atmosphere are referred to as *primary cosmic rays*, and upon collision, produce other particles, referred to as *secondary cosmic rays*. Both *electromagnetic* (electrons, photons) and *hadronic* (composite particles made up of quarks) particles are present in these interactions. The majority of the particles created in these interactions are mesons (comprising a quark and anti-quark) such as pions and kaons, which themselves decay into other particles. This 'cascade' of particle production, decay into other particles, and re-interaction with molecules in the atmosphere forms *air showers*. This phenomenon continues until the resulting particles lose all their energy. Air showers contain both hadronic and electromagnetic components, and can be quite extensive in their length and breadth.

Pions, the dominant mesons produced in air showers, follow a few decay modes. Neutral pions decay to photons, as

$$\pi^0 \to \gamma\,\gamma\,,$$

whilst charged pions follow the decay

$$\pi^+ \to \mu^+\,\nu\,,$$

and

$$\pi^- \to \mu^-\,\bar{\nu}\,,$$

where μ and ν denote the muon and the neutrino, respectively. Kaons follow similar decays to muons and neutrinos, mostly via

$$K^+ \to \mu^+\,\nu\,,$$

and

$$K^- \to \mu^-\,\bar{\nu}\,.$$

The branching fraction of these decays can be found in Workman et al. (2022) [3].

In this way, muons and neutrinos are produced aplenty in the atmosphere as a result of cosmic ray interactions. These muons, termed *atmospheric muons*, can travel to sea level and below and can be used to study the physics of cosmic ray interactions. These *atmospheric neutrinos* can be used to study the properties of neutrinos themselves. We will meet them again in later chapters.

Since the days of the first detection of cosmic rays, such as in 1912 by Victor Hess in a hot-air balloon flight [7], the detectors used to study cosmic rays, and their descendants, have advanced in technology and detection precision. In the 1930s, Pierre Auger pioneered the detection of air shower particles [8], measuring the arrival of particles at the same time, as recorded by equipment scores of metres apart — and thus concluding that the particles came from the same phenomenon. Nowadays, cosmic ray experiments can measure the particles from the electromagnetic and hadronic components of air showers (including atmospheric muons), and also detect the fluorescence light caused by air showers in the atmosphere across different energy ranges.

For example, the Pierre Auger Observatory [9] is one of the leading experiments to study and detect air showers from cosmic rays with energy above $\sim 10^{17}$ eV. The LHAASO [10] and Telescope Array [11] experiments measure cosmic rays with energies of 10^{11} – 10^{17} eV and above 10^{17} eV, respectively. Not all cosmic ray experiments are on Earth. For example, the AMS experiment [12] studies cosmic rays *before* their arrival on Earth; located on the International Space Station it plays a major role in the study of these particles and the cosmic ray spectral composition in the energy regime of 10^{9} – 10^{12} eV. Many outstanding questions remain, and AMS and other experiments aim to solve them.

Another means of studying the origin of cosmic rays is through electrically neutral particles such as gamma rays, which reach the Earth from extra-galactic sources, unperturbed by the magnetic fields encountered along the way. Gamma rays can originate from the same sources as cosmic rays. Experiments such as the FERMI-LAT satellite detector [13] and the ground-based gamma ray detectors such as HAWC [14], MAGIC [15] and HESS [16], cover the combined GeV – TeV energy regime for gamma ray detection.

Another piece of the same cosmic ray puzzle is the neutrino. As mentioned, cosmic neutrinos are produced by cosmic accelerators in the depths of outer space. The detection of these neutrinos plays a key role in the identification of their cosmic accelerators and the potential sources of cosmic rays. Such efforts to detect cosmic neutrinos and identify their astrophysical origins are carried out by neutrino experiments such as KM3NeT [17] (discussed in further detail in Chapter 2), IceCube [18], and others. In fact, the neutrino will form the main protagonist to this piece of work.

1.2. Neutrino Physics

Neutrinos hold a special place amongst the elementary particles described by the Standard Model of Particle Physics, a theory which gives a thorough description of three fundamental forces of Nature (the electromagnetic, weak and strong force, to be precise). Their ability to avoid detection and their status as the lightest elementary particle grants them this special rank in the Standard Model. The knowledge we have nowadays about neutrino physics and the sophisticated technologies needed to detect them are the result of arduous efforts of many previous physicists and experiments.

As now described in the Standard Model[1], neutrinos are light, spin-$\frac{1}{2}$ particles which interact with matter via the weak interaction. Their tiny interaction cross section determines their elusiveness for detection. We know today that neutrinos are realised in Nature as three distinct 'flavour' states: ν_e, ν_μ, ν_τ, pertaining to the lepton involved in their respective interactions. Forming our current picture of neutrino physics within the Standard Model is in itself a complex task.

1.2.1. A Brief History of the Neutrino

> "All great truths begin as blasphemies."
>
> George Bernard Shaw

In 1930, Wolfgang Pauli proposed an as-of-yet undetected particle, needed to explain the β-decay (a type of radioactive decay where an electron is emitted) energy spectrum. The electrons emitted as a result of β-decay showed a continuous energy spectrum, implying the presence of another, undetected particle alongside the electron. Pauli postulated a particle with spin-$\frac{1}{2}$ which has mass (differentiating it from photons) and is electrically neutral [19]. The sum of the energy of the electrons from β-decay and the new particle remain constant, explaining the continuous electron spectra. This new particle is what we now refer to as a neutrino.

In 1934, Fermi described the interaction governing β-decays [20] with a neutron converted to a proton, creating an electron and neutrino in the process, and using the name 'neutrino' for the first time (referring to it as a smaller counterpart of the electrically neutral 'neutron' discovered by Chadwick).

[1] To be more precise, as an addition to the Standard Model.

Although Bethe & Peierls speculated that the neutrino described by Fermi was unobservable due to its tiny cross section[2] [21], the experimental challenge was taken up in the 1950s by Raymond Davis Jr. and separately by Clyde L. Cowan and Frederick Reines.

The experiment ran by Raymond Davis Jr. sought to capture neutrinos with chlorine atoms, from the reaction

$$\nu\ {}^{37}\mathrm{Cl} \rightarrow {}^{37}\ \mathrm{Ar}\ e^{-}\ ,$$

with nuclear reactors providing the source of neutrinos.

In the end, Cowan and Reines were successful, observing

$$\bar{\nu}\ p \rightarrow e^{+}\ n\ ,$$

with a signal which implied a measured cross section of 6×10^{-44} cm^2 . They instead used liquid scintillator to measure the photons produced by the absorption of neutrons created by anti-neutrino interactions by Cadmium atoms. The distinction between detecting a neutrino and an anti-neutrino allowed for Cowan and Reines' successful detection and explains the lack of signal observed by Raymond Davis Jr. [22–24]. Raymond Davis Jr. is far from forgotten, however, as he is credited with observing neutrinos from the Sun with the Homestake experiment during the late 1960s.

In the same decade, an experiment performed by Goldhaber et al. (1958) [25] determined that neutrinos are left-handed particles (that the *helicity*, the projection of spin along the direction of momentum, is -1). It has been verified that the weak interaction only couples to left-handed fermions (such as neutrinos) and right-handed anti-fermions (such as anti-neutrinos). Neutrinos couple to the weak force in every frame of reference, obeying *chirality*, where a phenomenon is not the same as its mirror image.

Following speculation by Pontecorvo (1959) [26] (amongst others such as T.D. Lee and C.N. Yang) as to whether the neutrinos created in muonic interactions are the same as those neutrinos taking part in β-decay, the evidence for two separate types of neutrino, one muonic and one electronic, was presented by Danby et al. (1962) [27]. Following the discovery of the tau lepton [28] - indicating a corresponding tauonic neutrino - the tau neutrino was first evidenced in the year 2000 [29].

[2] Which was in itself contested by Bruno Pontecorvo, stating that a huge flux can overcome the cross section.

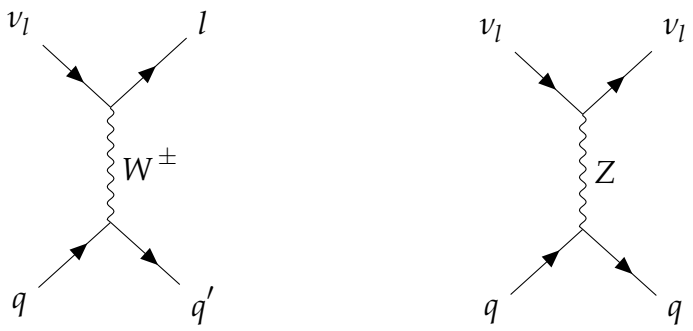


Figure 1.2.: Examples of neutrino interactions, in this case showing neutrino-quark interactions. In (a) we see the neutrino exchange a W boson, resulting in the creation of a lepton of the same flavour as the neutrino. In (b) we see the exchange of a Z boson. Since no charge is exchanged, the neutrino is not annihilated.

1.2.2. Neutrino Interactions

Weak Force mediators

The development of electroweak theory by Glashow, Weinberg, Salam, 't Hooft and Veldman [30–34] surmised the existence of the W and Z bosons, intermediaries of the weak force. Through the exchange of a charged W boson, a neutrino/anti-neutrino can interact with quarks in matter and result in the creation of a lepton. Since the W boson has charge, this interaction is termed a *charged-current* interaction. Through the exchange of a Z boson, a neutrino can also interact with matter and continue on its journey without annihilation. We refer to this as a *neutral-current* interaction, since the Z boson is without electric charge. These processes are illustrated in Figure 1.2.

Measurements carried out by LEP (the Large Electron-Positron Collider) included investigations into the width of the Z boson resonance [35]. Since this total width is composed of the hadronic contribution, electron contribution, muon contribution, tau contribution, and neutrino contribution, the measurement of the total width indicates the total number of neutrino species which interact via the Z boson. Please refer to Figure 1.3 for a depiction of this cross-section measurement as a function of energy. The number of neutrino species has since been further updated, and found to be [36]

$$N_\nu = 2.9963 \pm 0.0074\,.$$

This implies that only 3 neutrino species undergo the weak interaction, corresponding to the 3 flavours we have already encountered: the muon, electron and tau neutrino.

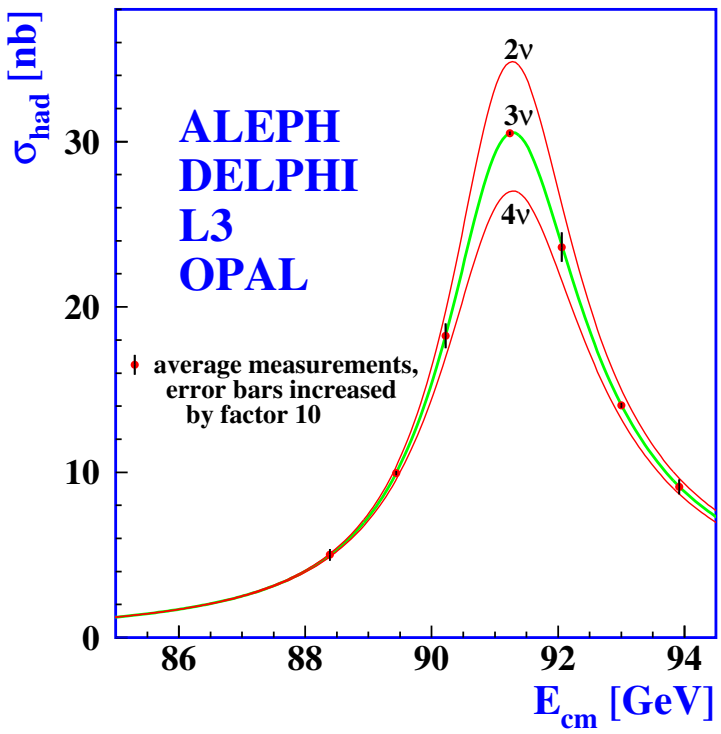


Figure 1.3.: Measurement of the hadron production cross section by the ALEPH, DELPHI, L3 and OPAL collaborations, around the Z boson mass. The predictions for the cross section given 2,3,4 (weakly interacting) neutrino species is shown. From S. Schael et al. (2006) [35].

Cross Sections - low energy neutrinos

The probability of interaction of neutrinos, i.e. the cross section, changes with the neutrino energy. Figure 1.4 illustrates this. Note that beyond $\sim$2 GeV the cross section becomes linear with the energy. Also, the cross section for anti-neutrinos is lower than that of neutrinos, and is not as precisely measured or with as many statistics.

The cross section can be grouped into different regimes. Here we focus on the lower energy neutrinos of less than 1 GeV to hundreds of GeV.

At energies below $\sim$2 GeV, neutrino-hadron interactions are dominated by *quasi-elastic scattering*. In these interactions, the neutrino scatters off the entire nucleon, as opposed to its constituent partons (quarks). For neutrinos and anti-neutrinos respectively, interactions are as follows (with n a neutron, p a proton, and l the resulting lepton):

$$\nu_l\, n \rightarrow l^-\, p\,,$$

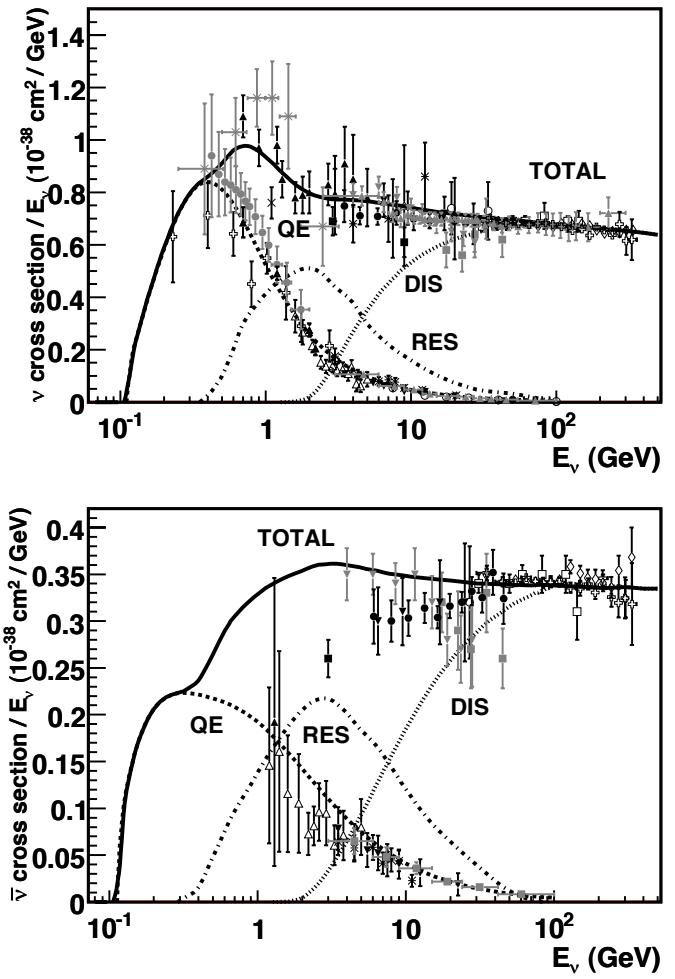


Figure 1.4.: The total neutrino and anti-neutrino per nucleon charged-current cross sections divided by neutrino energy, as a function of energy, for an isoscalar target. The different processes which contribute to the total cross section are shown. Top: cross section for neutrinos, bottom: cross section for anti-neutrinos. QE - quasi-elastic scattering, DIS - deep inelastic scattering, RES - resonance production. Taken from Formaggio, Zeller (2012) [37].

and

$$\bar{\nu_l}\, p \rightarrow l^+\, n\,.$$

For neutral-current interactions, the scattering is completely elastic, with the neutrino surviving the interaction.

At energies of a few GeV, *resonance production* describes when the neutrino can excite the nucleon to a higher energy state, with a baryonic resonance (a higher energy state of a baryon, a particle consisting of three quarks, e.g. N^*) decaying and usually with a pion in the final state. *Resonant pion production* is therefore a common mechanism and

dominates such interactions. To illustrate:

$$\nu_l\, N \rightarrow l^-\, N^*$$
$$N^* \rightarrow \pi N' \,,$$

where N,N' = n,p.

From Figure 1.4, we see that *deep inelastic scattering* soon dominates the cross section description, in the range of about 20 – 500 GeV. It is so-named since the scattering takes place with the constituent partons in the nucleon, rather than the proton or neutron itself. So, with enough energy the neutrino can probe the internal structure of the target nucleon. This can be described for charged-current interactions by

$$\nu_l\, N \rightarrow l^-\, X \,,$$
$$\bar{\nu}_l\, N \rightarrow l^+\, X \,,$$

and for neutral-current interactions by

$$\nu_l\, N \rightarrow \nu_l\, X \,,$$
$$\bar{\nu}_l\, N \rightarrow \bar{\nu}_l\, X \,,$$

where X denotes the nucleon following the interaction.

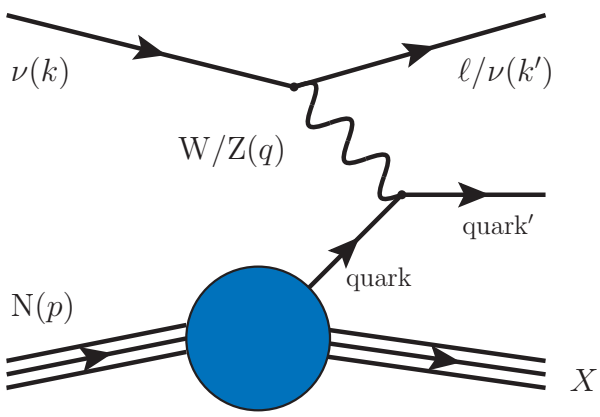


Figure 1.5.: A diagram of the deep inelastic scattering process. A weak force mediator (W/Z boson) is exchanged with the quark of a nucleon, resulting in an out-going lepton or neutrino and the formation of an hadronic shower. Taken from Garcia et al. (2020) [38].

Such a process is shown in Figure 1.5. The weak force mediator is exchanged with the partonic component of the nucleon itself. For a charged-current interaction, an out-going lepton is produced, whereas the neutrino carries on undetected in a neutral-current

interaction. In both instances, the resulting quark that is produced creates a hadronic particle shower. Here we can describe two kinematic properties of the interaction: the momentum transfer Q:

$$Q^2 = -q^2\,, \tag{1.2}$$

and the inelasticity parameter

$$y = \frac{E_\nu - E_l}{E_\nu}\,, \tag{1.3}$$

where q = k - k' (the difference between the incoming and out-going momentum), the neutrino energy E_ν and the energy of the out-going lepton E_l. The parameter y can be more easily thought of as the *fractional energy transfer from the neutrino to the hadronic shower*, and is sometimes referred to as the '**Bjorken-y**' parameter.

The cross section σ for neutrino/anti-neutrino and quark/anti-quark interactions (pertinent to deep inelastic scattering) can be described in terms of y [39]. The neutrino-quark and anti-neutrino-anti-quark differential cross section can be expressed as

$$\frac{d\sigma_{\nu q}}{dy} = \frac{d\sigma_{\bar{\nu}\bar{q}}}{dy} = \frac{G_F^2}{\pi}\hat{s}\,, \tag{1.4}$$

where q and $\bar{q}$ are used here to denote the quark and anti-quark, respectively, G_F is a constant which describes the weak interaction [20] and $\hat{s}$ is the centre-of-mass energy of the interaction. Similarly, for anti-neutrino-quark and neutrino-anti-quark interactions, the differential cross section can be expressed as

$$\frac{d\sigma_{\bar{\nu}q}}{dy} = \frac{d\sigma_{\nu\bar{q}}}{dy} = \frac{G_F^2}{\pi}(1-y)^2\,\hat{s}\,. \tag{1.5}$$

For these interactions, the differential cross section is modified by the term $(1-y)^2$.

Deep inelastic scattering cross sections, for neutrinos interacting with nucleons, are a combination of the cross sections for protons and neutrons. It can be shown that the the average neutrino differential cross section per nucleon N is [39]

$$\frac{d\sigma^{\nu N}}{dy} = \frac{G_F^2 m_N}{\pi}\,E_\nu\left[f_q + (1-y)^2 f_{\bar{q}}\right]. \tag{1.6}$$

The neutrino energy E_ν and nucleon mass m_N are used here, while f_q and $f_{\bar{q}}$ represent the fractions of the nucleon momentum carried by the quarks and anti-quarks, respectively. Similarly for anti-neutrino nucleon scattering, the differential cross section can be written as

$$\frac{d\sigma^{\bar{\nu}N}}{dy} = \frac{G_F^2 m_N}{\pi} E_{\bar{\nu}} \left[(1-y)^2 f_q + f_{\bar{q}}\right]. \tag{1.7}$$

In measurements of neutrino and anti-neutrino differential cross sections, for example those carried out by de Groot et al. (1979) [40], a large quark component in the nucleon and smaller anti-quark component in the nucleon is found. In terms of the Bjorken-y parameter, for neutrinos this results in an almost-flat cross section as a function of y, dominated by quark scattering and less so by anti-quark scattering (where the $(1-y)^2$ term comes into effect). For anti-neutrinos, a peaked distribution of the differential cross section as a function of y can be shown, now dominated by the quark contribution which is modified by the term $(1-y)^2$.

Cross Sections - high energy neutrinos

For neutrinos > 10^3 GeV (i.e. the TeV - PeV range), their interaction cross sections can be approximated as

$$\sigma_{\nu N}^{CC} = 5.53\text{x}10^{-36}\text{cm}^2 \left(\frac{E_\nu}{1\text{ GeV}}\right)^\alpha, \tag{1.8}$$

$$\sigma_{\nu N}^{NC} = 2.31\text{x}10^{-36}\text{cm}^2 \left(\frac{E_\nu}{1\text{ GeV}}\right)^\alpha, \tag{1.9}$$

where $\alpha \approx 0.363$, and where CC and NC indicate charged-current and neutral-current, respectively. This approximation is given in Formaggio, Zeller (2012) [37].

For high-energy neutrinos, the cross section for various interactions (with electrons or nucleons) is shown in Figure 1.6. A dominant feature is a resonant enhancement from the formation of an intermediate W boson in $\bar{\nu}_e e^-$ interactions, occurring at $E = M_W^2/2m_e$ = 6.3 PeV (in the rest frame of the electron, where M_W is the mass of the W boson and m_e is the electron mass). This feature was first predicted by Glashow in 1960 and so is often referred to as the *Glashow resonance* [41]. In 2021, the IceCube Collaboration measured a particle shower consistent with being produced at the Glashow resonance [42]. Note

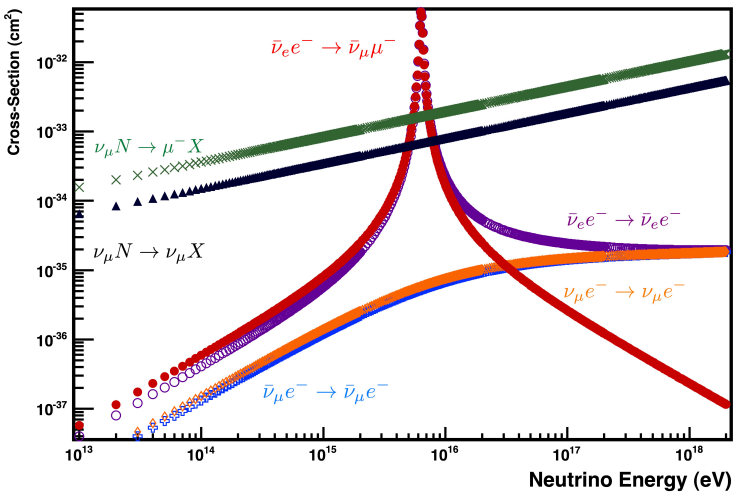


Figure 1.6.: The neutrino cross section for various processes across the high-energy regime. Shown are the electron and nucleon charged-current and nucleon neutral-current interactions, separately. Here N and X depict the nucleon before and following the interaction. Also visible is the leptonic W resonance. Taken from Formaggio, Zeller (2012) [37].

also that with increasing cross sections, the Earth becomes opaque to neutrinos in certain energy regimes.

The above description of neutrino cross sections is but an overview. For a detailed study, please consult Workman et al. (2022), Formaggio, Zeller (2012) [3, 37]. For a state-of-the-art calculation of neutrino cross sections in matter, including sub-dominant effects, please refer to Garcia et al. (2020) [38].

1.2.3. Neutrino Oscillations

Neutrinos undergo oscillations. This means that neutrinos inherently contain the probability to change flavour as they propagate through space and time. This simple yet powerful statement is the result of decades of research, which will not be detailed fully here but is summarised in Nobel Prize Outreach [43], with the discovery of neutrino oscillations being awarded a Nobel Prize in 2015. That neutrinos undergo flavour oscillations implies that neutrinos have mass, a fact that was not previously accounted for in the Standard Model of Particle Physics. Neutrino oscillation is a quantum mechanical phenomenon that can be described by mixing between mass and flavour eigenstates. The degree of mixing between these states is described by a mixing matrix.

A neutrino is created through the weak interaction as a flavour eigenstate $|\nu_\alpha\rangle$, where α = e, μ or τ. This flavour eigenstate is a linear combination (i.e. a superposition) of the three mass eigenstates $|\nu_1\rangle$, $|\nu_2\rangle$, $|\nu_3\rangle$, with respective masses m_1, m_2, m_3. Conversely, the neutrino can be described as a mass eigenstate which is a superposition of the flavour states. As the neutrino propagates through space-time, the phase of each mass state advances at a different rate to the others, due to the different neutrino masses. The advancing of the mass state phases with respect to one another is periodic and is a function of the distance travelled L and the neutrino energy E. This results in the observed phenomenon of neutrino oscillations: a neutrino of one flavour can be detected as another neutrino after travelling some distance, dependent on its energy.

These mass and flavour eigenstates can be related via

$$|\nu_\alpha\rangle = \sum_i U^*_{\alpha i} |\nu_i\rangle \,, \tag{1.10}$$

and

$$|\nu_i\rangle = \sum_\alpha U_{\alpha i} |\nu_\alpha\rangle \,, \tag{1.11}$$

where

- ν_α is a neutrino of definite flavour α = e, μ or τ
- ν_i is a neutrino of definite mass state i = 1, 2 or 3
- $U_{\alpha i}$ is the mixing matrix between the states. $U^*_{\alpha i}$ denotes its complex conjugate. For anti-neutrinos, the complex conjugate appears in Equation (1.11) and the conjugate in Equation (1.10).

We can relate the three neutrino flavour and mass states at a moment in time t by

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} . \tag{1.12}$$

Two-flavour approximation

An approximation to describe neutrinos oscillations considers oscillation between two flavours rather than three. This approximation can be pertinent, since neutrino oscillation

experiments usually focus on a neutrino of one particular flavour being detected as another flavour.

Consider a neutrino at some moment in time, which exists as a superposition of two flavour states, ν_α and ν_β. These flavour states can be related to two mass states ν_1 and ν_2 via a single mixing angle θ:

$$\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} \nu_i \\ \nu_j \end{pmatrix}. \tag{1.13}$$

It can be shown that the probability of this neutrino changing from flavour α to flavour β as it propagates *in vacuo* is

$$P(\nu_\alpha \to \nu_\beta) = |\langle \nu_\beta \,|\, \nu_\alpha(t) \rangle|^2 = \sin^2(2\theta)\, \sin^2\left(\frac{\Delta m^2_{ji}\, L}{4E}\right), \tag{1.14}$$

with L and E the distance travelled and energy of the neutrino, respectively. θ here is referred to as the *mixing angle* whilst Δm^2 is termed the *mass splitting*, the mass-squared difference between the two mass states. So, in the above example, $\Delta m^2_{ji} = m^2_j - m^2_i$. Neutrinos with no mass differences cannot undergo neutrino oscillation between different flavours. Note that the first term governs the amplitude of the oscillation and the mass splitting governs the wavelength.

Often, Equation (1.14) is written using the units commonly encountered in neutrino oscillation experiments:

$$P(\nu_\alpha \to \nu_\beta) = \sin^2(2\theta)\, \sin^2\left(1.27 \frac{\Delta m^2_{ji}[\mathrm{eV}^2]\, L[\mathrm{km}]}{E[\mathrm{GeV}]}\right). \tag{1.15}$$

Three-flavour description

The three-flavour neutrino oscillation probability is an extension of the two-flavour approximation. The probability that a neutrino, originally of flavour α, will be detected as a neutrino of flavour β is (in natural units, *in vacuo*):

$$P(\nu_\alpha \to \nu_\beta) = |\langle \nu_\beta \,|\, \nu_\alpha(L)\rangle|^2 = \left| \sum_i U^*_{\alpha i}\, U_{\beta i}\, e^{-i m_i^2 L/2E} \right|^2 , \tag{1.16}$$

which can also be written as

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 4\sum_{i>j} \mathrm{Re}\left(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}\right) \sin^2\left(\frac{\Delta m^2_{ji} L}{4E}\right) \\ +2\sum_{i>j} \mathrm{Im}\left(U^*_{\alpha i} U_{\beta i} U_{\alpha j} U^*_{\beta j}\right) \sin\left(\frac{\Delta m^2_{ji} L}{2E}\right). \end{aligned} \tag{1.17}$$

For anti-neutrinos, the imaginary part of the equation becomes negative. Here, $\delta_{\alpha\beta} = \langle \nu_\alpha \,|\, \nu_\beta \rangle$. $\delta_{\alpha\beta} = 1$ if $\alpha = \beta$ and $= 0$ if $\alpha \neq \beta$, thus for neutrinos of the same mass there would be no probability of flavour oscillation.

The mixing matrix U of Equation (1.11) is referred to as the PMNS (Pontecorvo-Maki-Nakagawa-Sakata) matrix. It describes the level of mixing between all the mass and flavour states. In full, this matrix can be defined as:

$$U_{\mathrm{PMNS}} = \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13} e^{-i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}, \tag{1.18}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$ and δ denotes a CP (Charge-Parity) violation phase (i.e., whether a physical interaction conserves charge and 'mirror' symmetry). The angle θ_{ij} parameterises the level of mixing between states i and j.

Thus we have three neutrino flavours ν_e, ν_μ, ν_τ, and three distinct mass states ν_1, ν_2, ν_3. We in turn have three mixing angles θ_{12}, θ_{23}, θ_{13} and three mass splittings Δm^2_{21}, Δm^2_{32}, Δm^2_{31}. Determining the values of the mixing angles and mass splittings with precision is one of the main research goals of neutrino oscillation experiments.

1.2.4. Neutrino Masses

Mass Ordering

We cannot directly measure the different neutrino masses from studying neutrino oscillations, only the respective mass splittings. The value of Δm^2_{21} has been determined and is of the order 10^{-5} eV^2, with Δm^2_{21} having a positive sign. Although the value of Δm^2_{23} ($\simeq \Delta m^2_{13}$) is measured to be of the order 10^{-3} eV^2, whether it is positive or negative is not known — i.e. it is unknown whether ν_3 is the lightest or the heaviest neutrino mass state.

This results in two possible orderings of the neutrino masses:

$$\nu_1 < \nu_2 < \nu_3 \,,$$

or

$$\nu_3 < \nu_1 < \nu_2 \,.$$

The former is referred to as *normal ordering* (NO), whilst the latter is referred to as *inverted ordering* (IO). Sometimes these cases are termed *normal hierarchy* (NH) or *inverted hierarchy* (IH). See Figure 1.7 for a diagrammatic representation of the neutrino mass ordering. The relative fractional amounts of the flavour state which comprise the mass states is also portrayed, i.e. the fraction gives an indication on the size of the mixing angle θ_{ij} between states. Note that such a diagram is representative of a neutrino in a vacuum.

Determining the neutrino mass ordering is of significant importance, for a few reasons. Firstly, an understanding of neutrinos and their masses is by itself an addition to fundamental knowledge. Neutrinos having mass is an *addition* to the Standard Model, and knowing their mass ordering ensures this addition is robust. The ordering of the masses is also significant in the understanding of flavour physics — quark masses and mixings, and other analogous properties [44]. It also has implications for cosmology and the evolution of the Universe, as neutrinos had an impact on the large-scale structure formation in the Universe. For a thorough review of neutrino cosmology, see Planck Collaboration (2021), Lesgourges and Pastor (2014) [45, 46]. Determining the neutrino mass ordering allows for further constraints on the CP-violating phase δ of Equation (1.18). For example, the neutrino experiment Hyper-K [47] aims to resolve the value of δ but relies on the neutrino mass ordering to be determined in order to do so.

The neutrino mass ordering can be determined experimentally, thanks largely to phenomena referred to as *matter effects*. When neutrinos pass through matter, the electron

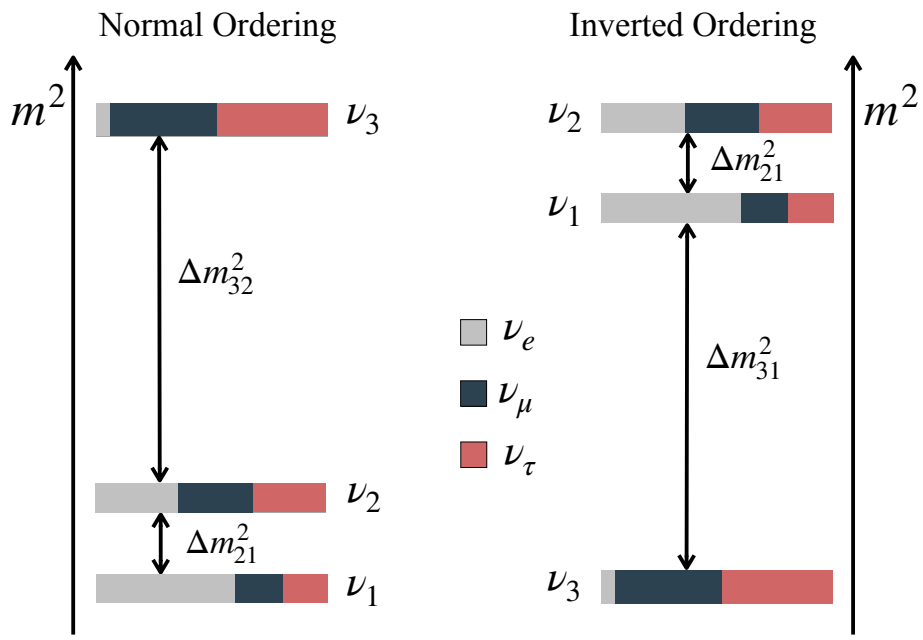


Figure 1.7.: The neutrino mass ordering (NMO for short) has two possible orderings, referred to as normal ordering (left) and inverted ordering (right). The sign of the Δm^2_{31} mass splitting is - at the time of writing - undetermined.

neutrino flavour can interact with *electrons* via charged-current interactions, resulting in an elastic forward scattering for this flavour alone. Figure 1.8 portrays these particular interactions.

In a vacuum, the Hamiltonian of the system is diagonal in the mass basis. In matter, the new, effective Hamiltonian can be diagonalised by adding an additional potential V to the system in a vacuum. V is given by

$$V = \pm\sqrt{2}G_F n_e\,, \tag{1.19}$$

with G_F and n_e the Fermi coupling constant and electron number density, respectively. This term V is positive for electron neutrinos and negative for electron anti-neutrinos. This effect introduces an asymmetry amongst the neutrino flavours travelling through matter, resulting in a distortion of the oscillation parameters.

For simplification, the two-flavour mixing scenario of Section 1.2.3 is used. We denote $\pm 2E_\nu\sqrt{2}G_F n_e$ by $\pm A$. In this case, the effective mixing angle θ^{eff} is given by

$$\tan 2\theta^{\text{eff}} = \frac{\tan 2\theta}{1 - \frac{\pm A}{\Delta m^2_{ji} cos 2\theta}} \tag{1.20}$$

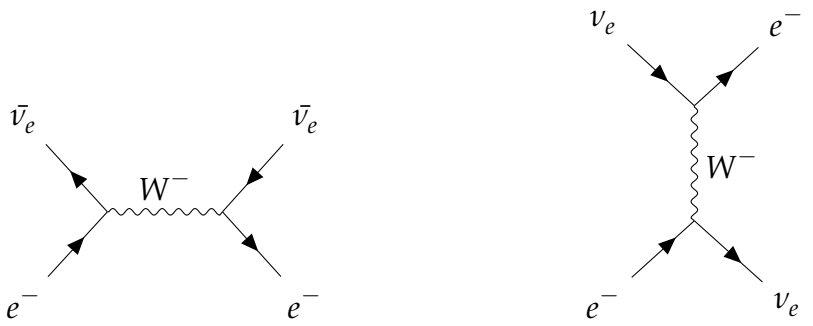


Figure 1.8.: Examples of neutrino coherent forward scattering due to the presence of electrons. For (a) anti-neutrinos and (b) neutrinos.

and the effective mass splitting by

$$\Delta^{\text{eff}} m^2 = \sqrt{(\Delta m^2_{ji} \cos 2\theta - \pm A)^2 + (\Delta m^2_{ji} \sin 2\theta)^2}\,. \tag{1.21}$$

For neutrinos traversing matter, these terms replace those in Equation (1.14). Note that in the limit of low electron number density, $n_e \to 0$, the oscillation parameters for neutrinos in a vacuum are recovered.

A resonance condition is met in Equation (1.20) for

$$\Delta m^2_{ji} \cos 2\theta = \pm A\,, \tag{1.22}$$

as $\theta^{\text{eff}} \to \pi/4$. Thus for certain values of Δm^2_{ji}, θ and E_ν (the neutrino energy), the probability of flavour transition of Equation (1.14) becomes maximal. This is known as *resonant enhancement*. For neutrinos traversing matter of near-constant density (relating to n_e), such as those created in the atmosphere and traversing regions within the Earth, a similar effect known as the *MSW effect*[3] occurs.

A related effect occurs for neutrinos passing through the Earth, termed *parametric enhancement*. For regions of abrupt changes in density, such as the boundary between the Earth's core and mantle, the neutrino flavour transition probabilities can again be enhanced.

Due to these combined *matter effects*, the neutrino mass ordering can be measured. In the case of normal ordering (see Figure 1.7) for neutrinos traversing matter, the probability of electron neutrinos changing flavour is enhanced, whereas it is suppressed for anti-neutrinos. In the case of inverted ordering, it is the flavour transition probability

[3] Named after Mikheyev, Smirnov and Wolfenstein.

of electron anti-neutrinos traversing matter which is enhanced, and suppressed for electron neutrinos. By measuring the number of neutrino events of a certain neutrino flavour, and comparing to the expected number of events for this flavour transition probability for both possible orderings, the neutrino mass ordering can be determined.

As mentioned, $A = \pm 2E_\nu\sqrt{2}G_F n_e$ is positive in the case of neutrinos, and negative in the case of anti-neutrinos. In the case of maximal mixing, $A = \Delta m^2_{ji}\cos 2\theta_{ij}$, à la Equation (1.22), and both A and $\Delta m^2_{ji}\cos 2\theta_{ij}$ must have the same sign. As a solution to the *solar neutrino problem* [43], which observed a deficit in the number of expected electron neutrinos from the Sun, neutrino oscillations along with matter effects were invoked. Since the Sun produces only electron neutrinos, and no electron anti-neutrinos, A is inferred to be positive (due to a deficit seen in the expected number of electron neutrinos, corresponding to maximal mixing).

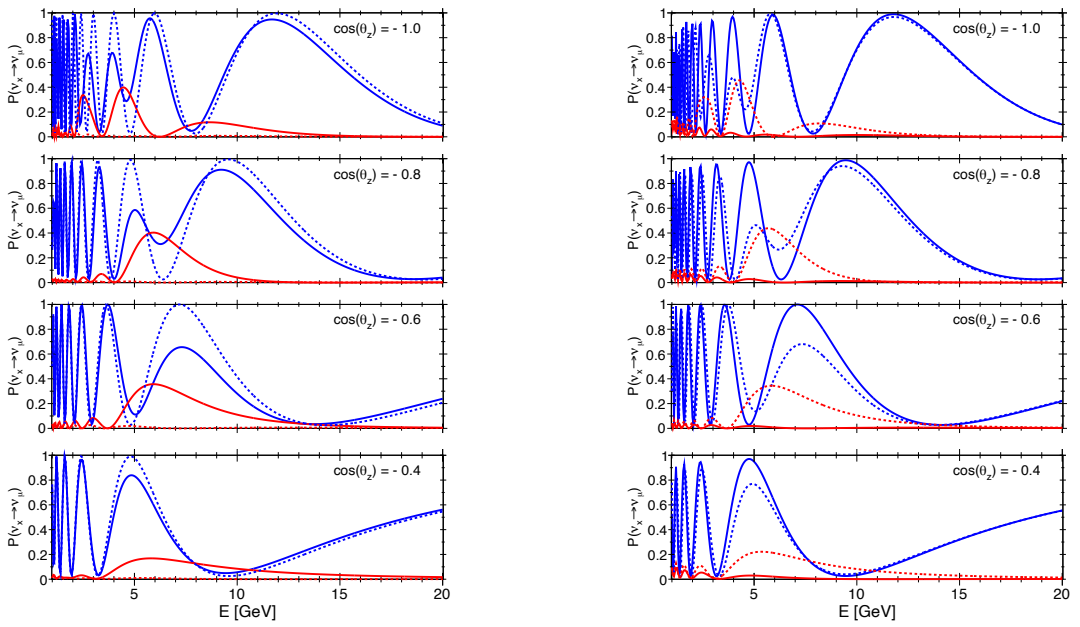
Now, in order for A to be positive, the terms $\Delta m^2_{ji}\cos 2\theta_{ij}$ should both be positive or negative. In the two-flavour approximation, the flavour state can be expressed as

$$\nu_e = \cos\theta_{12}\cdot\nu_1 + \sin\theta_{12}\cdot\nu_2\,.$$

If $\cos 2\theta_{12}$ is positive, then $\theta_{12} < 45^\circ$, leading to $\cos\theta_{12} > \sin\theta_{12}$. In this case, the flavour state ν_e is dominantly composed of the mass state ν_1, which is the lighter state because Δm^2_{21} is positive. If $\cos 2\theta_{12}$ is negative, then $\theta_{12} > 45^\circ$, leading to $\cos\theta_{12} < \sin\theta_{12}$. The flavour state ν_e is then dominantly composed of the mass state ν_2, which is the lighter state because Δm^2_{21} is negative. Thus, in either case, the mass state which couples most strongly to the ν_e flavour state is the lightest of the two mass states under consideration. With the convention that $\cos 2\theta_{12}$ is positive, the lightest mass state (which couples most strongly to the electron flavour) is denoted by ν_1, and $\Delta m^2_{21} > 0$.

For a more complete description of neutrino oscillation probabilities and matter effects, please refer to Ohlsson and Snellman (2001), Nosek (2017), Wolfenstein (1978), Choubey and Roy (2006), Hofestadt (2017) [48–52].

An example of the flavour transition probabilities for neutrinos and anti-neutrinos, according to both possible orderings, are shown in Figure 1.9 (a) and (b), respectively. The blue curves describe $P(\nu_\mu \to \nu_\mu)$ and the red curves describe $P(\nu_e \to \nu_\mu)$ transitions. These probabilities are shown as a function of the energy, for various zenith angle values. Here, the zenith angle θ_z refers to the direction of the incoming neutrino through the Earth, with $\cos(\theta_z) = -1$ meaning straight through from the opposite side ('upward-going neutrinos'). Different zenith angles imply different distances travelled through the Earth, and therefore different values of L in Equation (1.14). Due to the varying



(a) (b)

Figure 1.9.: An example of the flavour transition probability for $P(\nu_\mu \rightarrow \nu_\mu)$ (blue) and $P(\nu_e \rightarrow \nu_\mu)$ (red) as a function of the energy, for different zenith angles θ_z (distances travelled through the Earth). The solid lines depict the normal mass ordering and the dashed lines the inverted ordering. (a) shows neutrinos and (b) anti-neutrinos. This example is from a neutrino telescope studying atmospheric neutrino oscillations, taken from Adrián-Martínez et al. (2016) [17].

.

density profile of the Earth, the amount of matter traversed also changes, giving rise to matter enhancement effects and differences between the flavour transition probabilities for the two possible mass orderings. Note that the different orderings are quite similar, comparing neutrinos to anti-neutrinos, though a difference between the two is still introduced at certain energies and zenith angles (distances travelled) due to matter effects.

Another means of determining the neutrino mass ordering comes from experiments which attempt to measure neutrinoless double beta decays. Some elements decay by emitting two electrons, as opposed to one; such a process is termed *double beta decay*. If neutrinos are Majorana particles (the neutrino and corresponding anti-neutrino are the same), the neutrinos can annihilate, and the decay produces no neutrinos in the final state. The decay rate for neutrinoless double beta decay is proportional to the effective Majorana mass squared, $\langle m_{\beta\beta} \rangle^2$. Since the other terms relating the effective Majorana mass to the decay rate can be calculated, a value for $\langle m_{\beta\beta} \rangle^2$ can be found.

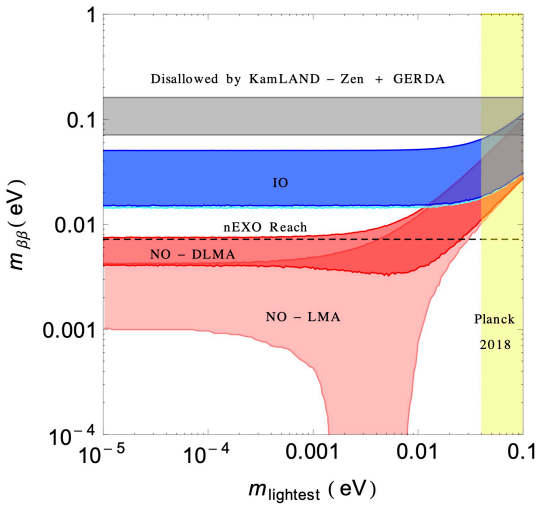


Figure 1.10.: The Majorana mass $m_{\beta\beta}$ as a function of the lightest possible neutrino mass, for both possible neutrino mass orderings. Alongside these regions are the constraints from current (grey region) and future (dashed line) neutrino-less double beta decay experiments. In addition, the region excluded by the constraint on the sum of the neutrino mass from Planck 2018 [45] is shown. From Vishnudath et al. (2019) [53].

Figure 1.10 shows $\langle m_{\beta\beta} \rangle$ as a function of the lightest possible neutrino mass, for both orderings [53]. The sensitivity of current and future experiments is shown. With enough sensitivity reach, a distinction can be made between the two orderings.

Additional mass measurements

A constraint on the *sum* of the masses of the neutrinos comes from cosmology. The sum of the neutrino masses is determined from measurements of the Cosmic Microwave Background (CMB) by the Planck Collaboration (2020) [45] to be

$$\sum m_\nu < 0.12\,\text{eV}\,,$$

under the assumption of the validity of the ΛCDM cosmological model.

Another means of measuring the (effective) neutrino mass comes in the kinematic study of weak interaction processes, such as the β-decay energy spectrum endpoint. By capturing the electron resulting from these decays, their energy spectra can be measured and the experimental result and the expected spectrum can be compared. The expected spectrum incorporates the square of the neutrino mass states m_i^2. In the quasi-degenerate energy regime of $m_i > 0.2$ eV, the neutrino mass measured by electron capture is the mass $m_\nu \approx m_i$ [54]. The KATRIN experiment [55] studies the β-decay endpoint of tritium (^{3}H)

decays, and as of 2022 has determined an upper limit of

$$m_{\nu_e} < 0.8\,\text{eV}\,,$$

with further sensitivity improvements expected for this measurement.

1.2.5. Current Neutrino Oscillation Measurements

A slew of neutrino oscillation experiments have attempted to measure the parameters associated with neutrino oscillations (outlined in Section 1.2.3), optimising the detectors based on the neutrino source energy range and baseline (distance from the source). Experiments such as Super-Kamiokande [56], SNO [57] and Borexino [58] detect(ed) solar neutrino oscillations, and so measure(d) θ_{12} and Δm^2_{21}. Experiments which measure neutrinos from nuclear reactors, such as Daya Bay [59] and Double Chooz [60], are sensitive to θ_{13} and Δm^2_{31}. Other experiments, sensitive to regions shown in Figure 1.9, measure atmospheric neutrinos. These include IceCube-DeepCore [61], Super-K and measure θ_{23} and Δm^2_{32}. In addition, detectors with a neutrino beam, such as T2K [62], NOvA [63] and MINOS [64], contribute to measurements of θ_{13}, θ_{23} and Δm^2_{32}[4].

A combined fit of the neutrino oscillation data, as of 2022, by NuFIT [65] is given in Figure 1.11. Here the best-fit point (bfp) and the 3σ significance intervals are given for the three-neutrino model oscillation parameters. These results are shown for both the normal and inverted mass ordering, with and without the inclusion of atmospheric neutrino data from Super-Kamiokande.

Note that here,

$$\Delta m^2_{3l} = \begin{cases} \Delta m^2_{31} > 0 & \text{for normal ordering} \\ \Delta m^2_{32} < 0 & \text{for inverted ordering} \end{cases}.$$

There are a few upcoming neutrino experiments, which aim to determine the neutrino mass ordering and measure θ_{23} with much greater precision. DUNE [66], for example, is a forthcoming neutrino beam experiment which can discriminate between both potential mass orderings. A successor to Super-Kamiokande, Hyper-K [47] has the ability to probe the value of the CP-violating phase δ, mentioned in Equation (1.18). Both experiments are in the early development stages.

[4] As can be seen, detectors are usually capable of focusing on one mass splitting only, depending on the detection method. This is why the two-flavour oscillation model remains relevant.

NuFIT 5.2 (2022)

		Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 2.3$)	
		bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range
without SK atmospheric data	$\sin^2\theta_{12}$	$0.303^{+0.012}_{-0.011}$	$0.270 \to 0.341$	$0.303^{+0.012}_{-0.011}$	$0.270 \to 0.341$
	$\theta_{12}/^\circ$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$
	$\sin^2\theta_{23}$	$0.572^{+0.018}_{-0.023}$	$0.406 \to 0.620$	$0.578^{+0.016}_{-0.021}$	$0.412 \to 0.623$
	$\theta_{23}/^\circ$	$49.1^{+1.0}_{-1.3}$	$39.6 \to 51.9$	$49.5^{+0.9}_{-1.2}$	$39.9 \to 52.1$
	$\sin^2\theta_{13}$	$0.02203^{+0.00056}_{-0.00059}$	$0.02029 \to 0.02391$	$0.02219^{+0.00060}_{-0.00057}$	$0.02047 \to 0.02396$
	$\theta_{13}/^\circ$	$8.54^{+0.11}_{-0.12}$	$8.19 \to 8.89$	$8.57^{+0.12}_{-0.11}$	$8.23 \to 8.90$
	$\delta_{\rm CP}/^\circ$	197^{+42}_{-25}	$108 \to 404$	286^{+27}_{-32}	$192 \to 360$
	$\frac{\Delta m^2_{21}}{10^{-5}\ {\rm eV}^2}$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$
	$\frac{\Delta m^2_{3\ell}}{10^{-3}\ {\rm eV}^2}$	$+2.511^{+0.028}_{-0.027}$	$+2.428 \to +2.597$	$-2.498^{+0.032}_{-0.025}$	$-2.581 \to -2.408$
		Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 6.4$)	
		bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range
with SK atmospheric data	$\sin^2\theta_{12}$	$0.303^{+0.012}_{-0.012}$	$0.270 \to 0.341$	$0.303^{+0.012}_{-0.011}$	$0.270 \to 0.341$
	$\theta_{12}/^\circ$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$	$33.41^{+0.75}_{-0.72}$	$31.31 \to 35.74$
	$\sin^2\theta_{23}$	$0.451^{+0.019}_{-0.016}$	$0.408 \to 0.603$	$0.569^{+0.016}_{-0.021}$	$0.412 \to 0.613$
	$\theta_{23}/^\circ$	$42.2^{+1.1}_{-0.9}$	$39.7 \to 51.0$	$49.0^{+1.0}_{-1.2}$	$39.9 \to 51.5$
	$\sin^2\theta_{13}$	$0.02225^{+0.00056}_{-0.00059}$	$0.02052 \to 0.02398$	$0.02223^{+0.00058}_{-0.00058}$	$0.02048 \to 0.02416$
	$\theta_{13}/^\circ$	$8.58^{+0.11}_{-0.11}$	$8.23 \to 8.91$	$8.57^{+0.11}_{-0.11}$	$8.23 \to 8.94$
	$\delta_{\rm CP}/^\circ$	232^{+36}_{-26}	$144 \to 350$	276^{+22}_{-29}	$194 \to 344$
	$\frac{\Delta m^2_{21}}{10^{-5}\ {\rm eV}^2}$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$	$7.41^{+0.21}_{-0.20}$	$6.82 \to 8.03$
	$\frac{\Delta m^2_{3\ell}}{10^{-3}\ {\rm eV}^2}$	$+2.507^{+0.026}_{-0.027}$	$+2.427 \to +2.590$	$-2.486^{+0.025}_{-0.028}$	$-2.570 \to -2.406$

Figure 1.11.: The NuFIT 5.2 (2022) combined fit of neutrino oscillation data. Shown are the best-fit parameter values under both neutrino mass orderings, with and without the use of Super-Kamiokande(SK) atmospheric neutrino data. For reference, please refer to `www.nu-fit.org` and Esteban et al. (2020) [65].

Another neutrino experiment, aforementioned, which will measure the neutrino mass ordering (amongst other physics goals), is KM3NeT. It will be described in comprehensive detail in the following chapter. The research carried out in this thesis is with data and simulations on behalf of the KM3NeT experiment.

Chapter 2.

The KM3NeT Experiment

The Cubic Kilometre Neutrino Telescope, or 'KM3NeT' for short, is a research infrastructure under construction on the Mediterranean Sea floor. KM3NeT foresees the completion of two neutrino detectors located at different sites in the Mediterranean Sea: KM3NeT/ARCA (**A**stroparticle **R**esearch with **C**osmics in the **A**byss) and KM3NeT/ORCA (**O**scillation **R**esearch with **C**osmics in the **A**byss). As suggested by the name, KM3NeT/ARCA will focus on neutrino astronomy, and KM3NeT/ORCA on the study of neutrino oscillations. The entire research infrastructure will comprise the same technology and design, with some changes in implementation depending on the physics goal.

In describing the KM3NeT experiment, we begin with the detection principle, and from there, describe the detectors, neutrino signals to be detected, backgrounds, and the data acquisition procedure. The research goals and current status of the experiment are then summarised.

2.1. Detection Principle

The KM3NeT infrastructure aims to instrument a large volume of water with light sensors which can detect the Cherenkov radiation from charged particles, in particular those resulting from neutrino interactions. For this purpose, KM3NeT employs the use of photomultiplier tubes - also widely known as PMTs - to detect the light signal of interest and amplify it into an electrical current. The principle behind Cherenkov radiation is first described.

2.1.1. Cherenkov Radiation

When a charged particle propagates through a dielectric medium, it can excite the molecules around it, which emit energy on returning to their ground state. This energy is emitted as spherical wavefronts of light which originate from the particle at each point in space and time. If the velocity of the particle v_p is greater than the speed of light in the medium, i.e. if

$$v_p > \frac{c}{n}, \tag{2.1}$$

where c is the speed of light in a vacuum and n is the refractive index of the medium, these spherical waveforms constructively interfere. As a result, a cone of light is emitted in the direction of travel of the particle [67]. This form of electromagnetic radiation is known as **Cherenkov radiation** [68]. Figure 2.1 shows an illustration of this effect.

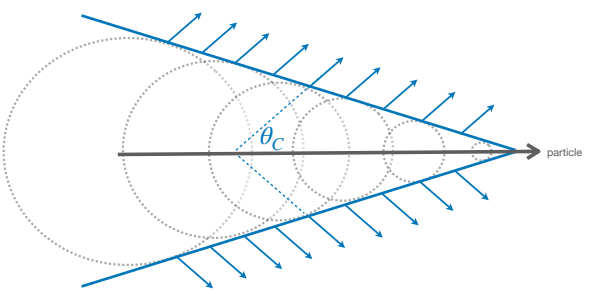


Figure 2.1.: As a particle moves through a dielectric medium, at a speed greater than the speed of light in the medium, Cherenkov light is released as overlapping spherical waveforms in the form a cone. The formation of the spherical waveforms is in fact continuous, here they are spaced out for illustrative purposes. The light is emitted at a characteristic angle θ_C.

This light cone is emitted at a characteristic angle θ_C. If β is the ratio between the speed of the particle and the speed of light, i.e.

$$\beta = \frac{v_p}{c}, \tag{2.2}$$

it can be shown that the angle of emission is given by

$$\cos\theta_C = \frac{1}{n\beta}, \tag{2.3}$$

with θ_C the Cherenkov emission angle.

Crudely, assuming the refractive index of seawater to be $\sim$ 1.33, and a particle moving approximately at the speed of light, we find $\theta_C \simeq 41^\circ$.

The number N of Cherenkov photons produced per unit path length x by a particle with charge qe (e is the electron charge), moving with speed $\beta = v/c$ through a medium (where v is the particle velocity and c is the speed of light in a vacuum), can be expressed as [3]

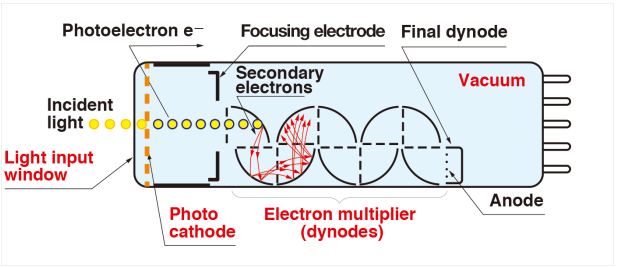
$$\frac{d^2N}{dxd\lambda} = \frac{2\pi\alpha q^2}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2}\right). \tag{2.4}$$

Here, α is the electromagnetic force coupling constant (= 1/137), λ is the wavelength of light and n is the refractive index of the medium. From Equation (2.4), the majority of photons are radiated in the ultraviolet region of the electromagnetic spectrum. Cherenkov radiation is visible as blue light in the visible region of the electromagnetic spectrum, and for example may be observed from underwater nuclear reactors.

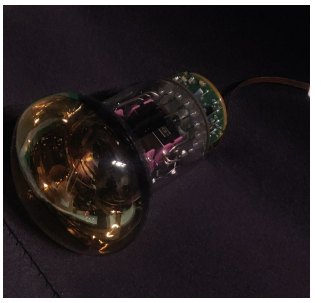
2.1.2. PMTs

PMTs are glass-encased vacuum tubes, used to detect photons in the ultraviolet, visible, and near-infrared frequency ranges of electromagnetic radiation. They multiply the electrons caused by incident photons into many more electrons, which are directed from one end of the tube to the other using a potential difference.

The basic operation of a PMT is as follows. A *photocathode* emits electrons via the photoelectric effect, following the absorption of incident electromagnetic radiation. A high voltage is applied across the photomultiplier tube, in order to create a large potential difference for accelerating electrons from one end to the other. A *focusing electrode* directs the accelerated electrons towards a series of electron multipliers known as *dynodes*. At the dynodes, further electrons are created through secondary emission, which is the incidence of particles of sufficient energy hitting a surface and resulting in the creation of secondary particles. A potential difference at each dynode ensures the movement of electrons towards the other side of the PMT. These electrons are focused upon an *anode*, which converts the electrons into a measurable electric current. Figure 2.2 shows a schematic of a PMT with its components, alongside an image of one used as part of the KM3NeT experiment.



(a)



(b)

Figure 2.2.: (a) Schematic of a PMT and its components, from `www.hamamatsu.com`. (b) A PMT used as part of the KM3NeT experiment.

The base of the PMT digitises the analog PMT signal pulse, as well as supplying the required high voltage. Two features of the resulting signal are recorded. The time at which the pulse crosses a predefined threshold pulse height is read out as the **arrival time** of the incident photon. The pulse duration - the time during which the signal pulse stays above this threshold - is also recorded. This metric for the number of photons which illuminate the PMT is referred to as the '**time-over-threshold**'. A threshold of 0.3 times the mean pulse height resulting from a single incident photon is used. This threshold allows for distinguishing the output signal from random noise due to the PMT electronics. A single datum containing the arrival time and time-over-threshold information recorded by a PMT is referred to as a '*hit*', corresponding to 6 Bytes of data (1 Byte for the PMT address, 1 Byte for the time-over-threshold, and 4 Bytes for the time)

The *quantum efficiency (QE)* is a unitless number which can be used to quantify the light sensitivity of a PMT, dependent on the wavelength of light illuminating the photocathode. It can be defined as the probability to convert a photon into an electron, or

$$\mathrm{QE} = \frac{N_{\mathrm{electron}}}{N_{\mathrm{photon}}}, \tag{2.5}$$

where N_{electron} is the number of electrons produced at the photocathode and N_{photon} is the number of absorbed photons. A specialised coating on the photocathode is used to greatly increase the quantum efficiency of a PMT, allowing it to peak at different wavelengths. Thus the quantum efficiency can be tuned depending on the incident electromagnetic radiation to be detected.

KM3NeT first employed the use of Hamamatsu R12199-02 type, three-inch PMTs. For the later batches of PMTs, the Hamamatsu R14374 three-inch model was adopted. Both models adopt a bialkali coating of the photocathode. For these PMTs, it takes about 30 ns for a photon impinging on the photocathode to result in electrons hitting the anode [69].

This time period is called the **'transit time'** of the PMT. This quantity is variable, with an associated spread on its mean value of 1.3 ns for the latest model of PMTs (measured from the full width at half maximum of the transit time distribution). This '**transit time spread (TTS)**' is the main uncertainty on the measured photon arrival time.

Another important PMT parameter is the '**gain**', describing the overall electron multiplication factor of a PMT (the average number of electrons released from the anode following one emitted photo-electron from the photocathode). For the KM3NeT experiment, the gain is 3×10^6 for the voltage range of 900 – 1300 V [17]. This value is chosen to maximise the lifetime of the PMTs (i.e. to minimise the effects of ageing) in accordance with the expected operational time of the detector [70], given the high counting rate the PMTs will experience in the deep sea.

2.2. Infrastructure

The KM3NeT infrastructure provides the means of instrumenting a large body of water with PMTs in order to detect the light from neutrino interactions. Due to the minuscule interaction cross sections of neutrinos and anti-neutrinos, described in Chapter 1, a large volume of interaction medium is needed in which neutrino interactions can be detected, in order to increase the interaction probability. Seawater is an abundant and natural medium fit for this purpose. The infrastructure and technology developed in order to instrument a volume of seawater with PMTs is described below.

2.2.1. Optical Modules

Pressure-resistant glass spheres of 0.44 m diameter are used to house 31 PMTs each. They also contain electronics required for supplying power and for the processing and readout of the PMT signals, and devices for monitoring information [70,71]. This entire assembly of the glass sphere, 31 PMTs and all contained electronics and devices is referred to as a '**digital optical module**'. The timing and positioning calibration devices within the optical module include: a '**nano-beacon**' - a small light emitting diode (LED) - which emits light in the violet-blue visible wavelength, a compass and tiltmeter for orientation calibration, and an acoustic piezo sensor used for position calibration. Calibration methods are described later in this chapter. A humidity and temperature sensor are also present in each optical module.

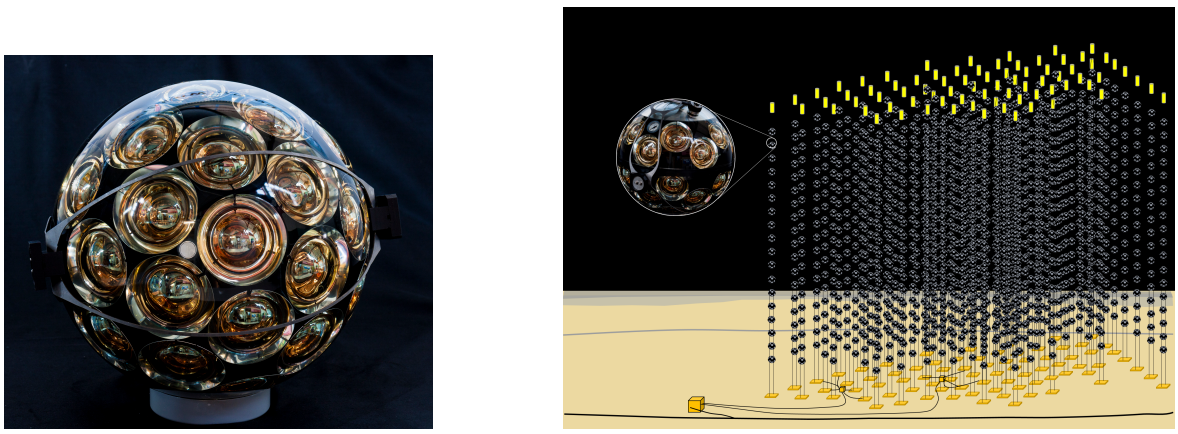
An electronics board known as the '**central logic board**' acts as the proverbial heart of the optical module electronics, controlling all of the instrumentation within and allowing for communication with the optical module when it is deployed in the sea [72]. Mounted on the central logic board is a Field-Programmable Gate Array (or FPGA for short), a type of integrated circuit which can be modified by the user. These FPGAs contain programmable logic circuits, used to receive data from the 31 PMTs and the piezo sensor, and to allow for data acquisition and communication with the optical module and its components. The arrival time and time-over-threshold sent by each PMT is recorded on the FPGA, with a base clock on each FPGA allowing for timestamping of the hits. A maximal threshold count rate of 20 kHz per PMT is implemented by the FPGA, with any recorded data that exceeds this threshold being discarded. This is referred to as the '**high-rate veto**' and can reduce the signal from the bioluminescence of marine life, as noted in Section 2.4.

A feed-through component for power and optical fibres for data transmission known as a '**penetrator**' is inserted at the top of the digital optical module. It is capped with a '**break-out box**', which isolates the electronic components within the optical module and connects to power cables and optical fibres on the outside. Also part of the optical module is a titanium collar, used for mounting to detection units, described in the following section.

An image of a digital optical module can be seen in Figure 2.3 (a), conveying the photon detection area captured by the PMTs. A complete overview of the optical module can be found in S. Aiello et al. (2022) [70].

2.2.2. Detection Units & Sea Floor Network

18 digital optical modules connected between two Dyneema® ropes form a structure referred to as a '**detection unit**'. These fibre ropes are 4 mm in diameter and have an extremely high tensile strength, allowing them to bear the load of the optical modules while remaining quite light. The optical modules are spaced apart, with a spacing that differs between KM3NeT/ARCA and KM3NeT/ORCA (described below). The optical modules are interspersed along these ropes like the rungs of a ladder, and attached via the titanium collar of each optical module. When deployed, these detection units are aligned vertically from the sea floor. They are held upright through the buoyancy of the optical modules, and additionally with a buoy at the top. This keeps the structure relatively rigid, and attempts to reduce the movement of the topmost optical module relative to the bottommost one due to sea currents.



(a) (b)

Figure 2.3.: (a) The bottom view of a digital optical module, showing the coverage achieved by the PMTs. The grey piezo sensor used for positioning calibration is visible left of centre of the image. The titanium collar for attachment to a detection unit (see text) is also visible. (b) An artistic impression of a KM3NeT detector building block, comprising 115 detection units, each accommodating 18 digital optical modules. A close-up of one optical module shows the titanium collar and external pressure gauge meter. Missing in this image is the break-out box (see text). Also present in this image is the sea floor network of optical fibres and electric power cables.

Along the ropes of each detection unit is a '**vertical electro-optical cable**', a plastic tube filled with oil to allow for withstanding water pressure in the depths of the Mediterranean Sea. This cable contains copper wires for power transmission to the optical modules, and optical fibres for data transmission to and from each optical module [17]. An anchor at the bottom of each detection unit contains a '**base container**', housing further optical components. Each detection unit is anchored to a junction box on the sea floor network. These house electronics and (sometimes) acoustic calibration devices. This network is fed power through a '**main electro-optical cable**' which connects to a shore station back on dry land. A network of separate acoustic emitters on the sea floor is under construction as part of the KM3NeT infrastructure, used to calibrate the positions of the optical modules in the water.

115 detection units (also colloquially termed 'lines') form a detector '**building block**'. An illustration of detection units forming a building block, attached to a sea floor network of power cables and optical fibres, is presented in Figure 2.3 (b).

2.2.3. Detector Overview

KM3NeT/ARCA will consist of two building blocks upon completion, while the KM3NeT/ORCA detector will be made up of one building block. These detectors will also differ in size. KM3NeT/ARCA will have an average spacing between optical modules of 36 m, and 90 m between detection units on the sea floor. It will enclose 2 x 0.48 km^3 of seawater, optimised for detecting neutrinos in the TeV–PeV range. The research infrastructure for KM3NeT/ARCA is located at at 36° 16′ N 16° 06′ E at a depth of 3.5 km, 100 km offshore of Portopalo di Capo Passero, Sicily.

KM3NeT/ORCA will exist as a denser configuration of the building block, with an average spacing between optical modules of 9 m along a detection unit, and 20 m between the detection units themselves. This geometry is optimised for the detection of neutrinos in the few GeV to 100 GeV range (c.f. Figure 1.9). The sea floor network for KM3NeT/ORCA is located at 42° 48′ N 06° 02′ E at a depth of 2.45 km, about 40 km offshore of Toulon, France. On completion, KM3NeT/ORCA will instrument about 6.7×10^6 m^3 of seawater, equivalent to 7.0 Mt.

The geometry of both detectors has been optimised according to their physics goals. The modular design of the detection units allows for the same technology to be used to construct two separate detectors.

2.3. Neutrino Event Description

The KM3NeT detectors aim to measure the Cherenkov light from particles, created by neutrinos interacting in the enclosed volume of seawater. Different neutrino interactions result in different three-dimensional profiles of light emission (and detection) in the detector, and so different neutrino events can have a distinctive event ‘signature’.

Following the discussion of neutrino interactions in Section 1.2.2, a neutrino or anti-neutrino ν_l of flavour l undergoing a charged-current interaction will produce a lepton of the same flavour. The three neutrino flavours of ν_μ, ν_e, ν_τ can thus result in a muon μ, electron e, or tau lepton τ, and likewise for anti-neutrinos. These leptons can produce Cherenkov radiation. For neutrinos which follow deep inelastic scattering interactions (Section 1.2.2), the quark of the nucleon with which the neutrino interacts is freed, and produces hadronic particles. These hadrons, with sufficient energy, can emit detectable Cherenkov radiation, as can the products of their interactions. Deep inelastic scattering is dominant for the energy ranges of neutrinos detected by KM3NeT/ORCA.

The neutrino interactions described above can be categorised into event signatures of different types, depending on their light-emission topology. Two of the most common type of neutrino event signatures follow the nomenclature in neutrino physics of '**track-like**' and '**shower-like**' events, or more colloquially, '**track**' and '**shower**' events. They will be described below. The use of this type of event classification is useful to separate quite distinct events. Neutrino interactions in fact create event signatures which are combinations of the two, and this type of event classification is merely a simplification of what actually occurs.The energy and direction of these event types can be determined - *reconstructed* - and used as a proxy to estimate the neutrino direction and energy. The reconstruction software used by the KM3NeT experiment is described in Chapter 3.

A portrayal of the event signatures discussed in this section is given in Figure 2.4. These neutrino event signatures are not to scale, nor are the detector components, they are for illustrative purposes.

2.3.1. Track Events

The muons and tau leptons created in (anti-)neutrino interactions travel through the KM3NeT detectors in a straight line, with minimal momentum transfer to electrons due to their large mass. The pattern of light deposited in the detectors follows the path of these particles, which emit Cherenkov radiation in the form of a Cherenkov cone. These straight-line events are referred to as track-like events. A pattern of detected hits which follow this topology can be used to identify such particles (and hence the possible flavour of the neutrino).

The length of muon or tau lepton tracks is dependent on their mass and energy. Muons at rest have a lifetime before they decay of about 2.2×10^{-6} s. When travelling near the speed of light they can travel large distances, depending on the energy, and so produce long track-like signatures in the detectors. Tau leptons at rest have a much shorter lifetime of 2.9×10^{-13} s and even when travelling near the speed of light create much shorter tracks in the KM3NeT detectors.

Muons lose energy when traversing seawater through various means. These include *ionisation* energy losses, the production of *delta rays* (ionised electrons which in turn cause further ionisation) and *photo-nuclear interactions* (inelastic collisions with atoms). According to Workman et al. (2022) [3], the average energy loss of muons with energy E

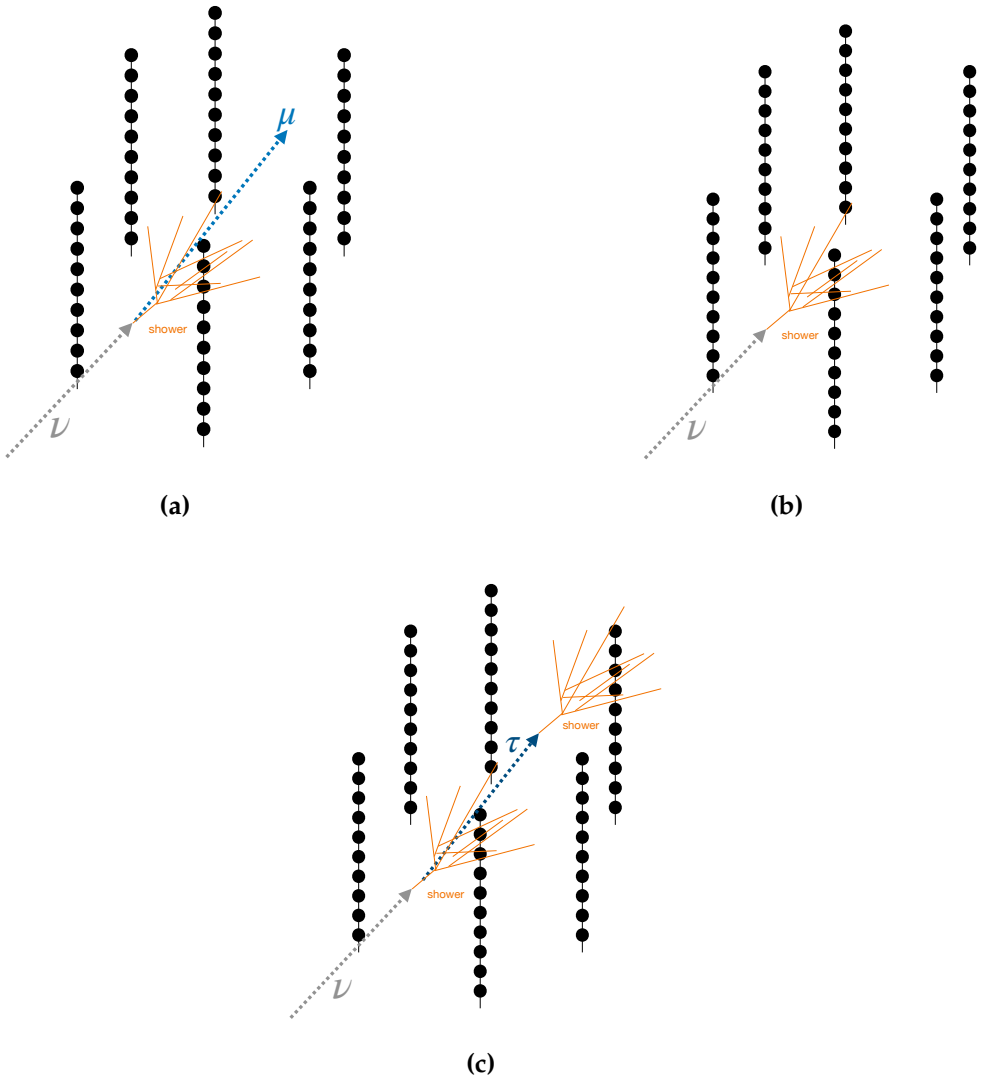


Figure 2.4.: Event signatures of neutrino interactions in KM3NeT. The detection units are depicted as black lines containing circular optical modules. Such diagrams are merely for purposes of illustration. In (a) we see a muon track μ, a typical event signature that is reconstructed in KM3NeT. The muon neutrino or anti-neutrino involved in this interaction can produce a hadronic shower at the interaction vertex, in addition to the out-going muon. In (b) is a typical particle shower interaction, also known as 'cascades'. The difference between hadronic or electromagnetic showers is described in the text. Both neutral-current and charged-current interactions result in these event types. Neutrinos or anti-neutrinos of all flavours produce such showers. Figure (c) depicts a tau neutrino or anti-neutrino interaction, resulting in a tau lepton track τ between two showers. Such an event is referred to as a 'double bang'. The tau lepton can instead decay into a muon track.

in seawater, per unit distance travelled x, can be approximated by

$$-\frac{dE}{dx} = \alpha + \beta \cdot E\,. \tag{2.6}$$

This gives an effective energy loss described by both ionisation and radiative energy loss processes. For lower energy neutrinos of 1–100 GeV (those detected by KM3NeT/ORCA), ionisation losses will dominate for the created muons, whereas radiative processes come into play at higher energies (e.g. for muons detected by KM3NeT/ARCA), with these muons producing additional particle showers along their tracks (through e^+e^- pair production) or emitting *bremsstrahlung* radiation. At lower energies, muons and tau leptons can be approximated as *minimum ionising particles*, where the particle energy loss rate is close to the minimum. Using Equation (2.6) and reference values of $\alpha = 0.277$ GeV m^{-1} and $\beta = 3.53 \times 10^{-8}$ m^{-1} [69,73], a muon of energy 1 GeV will travel about 4 m in seawater.

2.3.2. Shower Events

Electron neutrinos and anti-neutrinos can result in electrons and positrons created through charged-current interactions. Muons can decay to electrons, whilst tau leptons can decay to electrons, muons or hadronic particles. As mentioned, all neutrino interactions which undergo deep inelastic scattering processes (both charged-current and neutral-current interactions) can result in the creation of hadronic particles.

These hadronic particles decay into secondary particles of lower energy, which in turn lead to further particle decay and production, and so on and so forth until there is not enough energy in the system to create new particles. This avalanche of particle decays and interactions is termed a shower, akin to air showers produced by cosmic rays, as described in Section 1.1.2.

Each individual charged particle of the shower with sufficient energy may emit Cherenkov radiation, which can be detected. These shower-like events appear as bursts of light in a localised area of the detectors (compared to track-like events which traverse greater distances). For lower neutrino energies, as detected in KM3NeT/ORCA, the location of light emission is limited, due to the smaller volume in which the showers develop. For higher energies, as detected by KM3NeT/ARCA, the showers may be larger and more expansive. The Cherenkov light from these showers result in hits in the detectors. The timing and position information of these hits can be used to reconstruct the overall, average shower direction, energy, and vertex (the starting point of the interaction). The difference between electromagnetic and hadronic showers is described below.

Electromagnetic Showers

The electrons and positrons produced by electron neutrino or anti-neutrino interactions (or the decay of neutrino-induced muons or tau leptons) emit photons through bremsstrahlung, the electromagnetic radiation resulting from the deceleration of charged particles in the presence of the electric field of other charged particles. These photons in turn may produce further electron-positron pairs through *pair production*, and the cascade of particles continues. Such a cascade is referred to as an **electromagnetic shower**, since these particles primarily interact via the electromagnetic force.

The *point of shower maximum* is where the most light is radiated by the whole shower, this is where the average energy of the shower is insufficient in producing further particles, and the shower begins to diminish in size. The radiation length X_0 describes the characteristic amount of matter traversed for the processes of bremsstrahlung and pair production in a particle shower, measured in g cm^{-2} [3]. The length X of the shower where it reaches maximum size can be expressed as

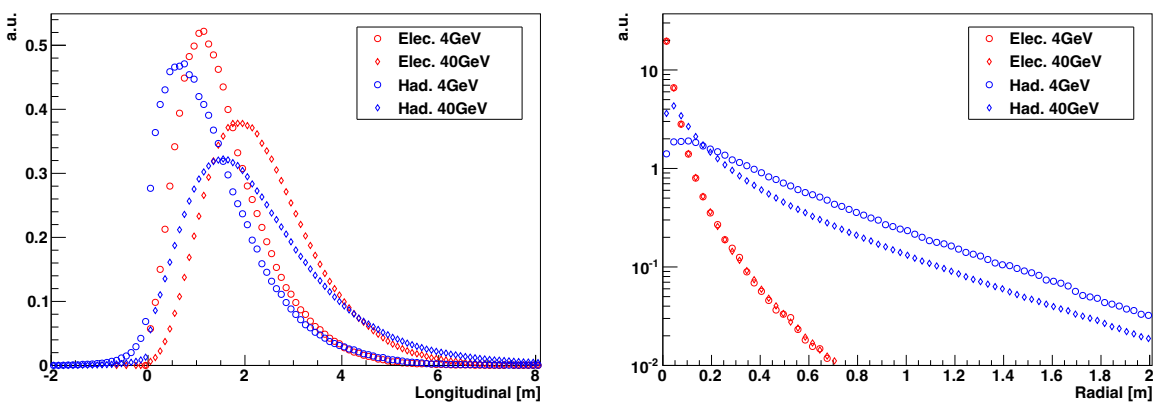
$$X = X_0 \frac{\ln E_0/E_c}{\ln 2}\,. \tag{2.7}$$

Here, E_0 is the energy of the primary particle and E_c is the critical energy, the energy at which bremsstrahlung and ionization energy losses become equal. In water, the values for X_0 and E_c are approximately 36 g cm^{-2} and 78.3 MeV (for electrons) [3].

Hadronic Showers

Hadronic showers are created at the interaction point of the neutrino, where the neutrino interacts with a nucleon and hadrons are then produced. These particles are composed of quarks and interact via the strong force. They in turn produce particle showers, which can consists of both hadronic particles or electromagnetic sub-showers; the electromagnetic sub-showers arise from electrons or photons (mainly coming from π^0 decays) with sufficient energy. Muons may also be produced in these decays (e.g. from charged pions); these muon tracks are usually shorter than those produced by neutrinos due to their lower energy. At lower energies in the GeV range, hadronic showers show more intrinsic fluctuations in the number of emitted Cherenkov photons and in the angular distribution of light in the detectors compared to electromagnetic showers [17]. In terms of event signatures though, hadronic showers are still recognisable as localised areas of light detection in the KM3NeT detectors.

The light emission profiles of hadronic and electromagnetic showers are depicted in Figure 2.5, in the radial and longitudinal direction (the longitudinal direction is along the shower axis, the radial direction is perpendicular to it). The longitudinal extension of the showers increases with the logarithm of the energy. The radial extension of the shower is dependent on how much energy is in the system, and the scattering angle between particles. At higher energies, relevant to KM3NeT/ARCA, the longitudinal elongation of the shower becomes quite important in determining its direction and energy (for electromagnetic and hadronic showers). In KM3NeT/ORCA the particle showers resemble less expansive, localised bursts of light.



(a) (b)

Figure 2.5.: (a) Longitudinal and (b) radial light emission profiles of electromagnetic (red) and hadronic (blue) showers in water, at energies of 4 and 40 GeV. The units of light emission are arbitrary. From Adrían-Martínez et al. (2016) [17].

The point of shower maximum is also used in describing hadronic showers. For these showers, only a fraction of the energy is converted to particles which emit Cherenkov radiation. Some neutral particles may be produced which yield no light. The detectable light can be referred to as the '**visible energy**' of the shower (which is normalised to the equivalent electromagnetic shower energy). A comprehensive description of modelling hadronic showers can be found in Matthews (2005) [74] — although it refers to the cosmic ray air showers of Section 1.1.2, it is relevant for hadronic showers caused by neutrino interactions.

The tau neutrino and anti-neutrino are particular cases in terms of event signatures. It has already been mentioned in Section 2.3.1 that tau leptons can resemble track-like events. However, as for all neutrino flavours, tau (anti-)neutrino interactions will likely result in the creation of a hadronic shower. In turn, tau leptons can decay hadronically *and* electromagnetically. Hence, the distinguishing interaction signature of a tau neutrino

or anti-neutrino is a hadronic shower, with a resulting tau lepton, which in turn results in another particle shower, be it hadronic or electromagnetic. Such an event can be referred to as a '**double bang**'. Another possible event signature is the resulting tau lepton showing a resulting muon track: a *muonic tau decay*. Such an event signature has the potential to be classified and reconstructed as a track-like event or a shower-like event.

2.3.3. A More Realistic Description of Neutrino Interactions: Track & Shower Events

The spatial/geometrical differences between track and shower events allow for them to be treated separately. However, at the neutrino energies of interest to the KM3NeT experiment, for track events a hadronic shower is always created in the interaction alongside a track signature (due to deep inelastic scattering, see Section 1.2.2). If this shower is in the vicinity of the detectors, it can result in a detected light signal more localised than that from a track. For example, muons produced by charged-current neutrino or anti-neutrino interactions will emerge alongside a hadronic shower. By considering both the track and shower components of this type of interaction and determining their respective directions and energies, a more precise determination of the direction and energy of the neutrino can be made. This in turn can lead to a more precise estimate of the neutrino mass ordering. In addition, this gives a handle on determining the fractional energy transfer, or inelasticity parameter, y (of Equation (1.3)). The same logic can be applied to tau neutrino interactions.

2.4. Backgrounds

A major consideration for any physics experiment is the background signal. The data arriving on shore from the detectors need to be processed to extract the neutrino events of interest, and also to reduce the sheer volume of data received. For KM3NeT, there are a few main background sources of light to consider, which can resemble the neutrino event signatures described above.

Potassium-40 decays

Potassium-40 (^{40}K) is a naturally occurring, radioactive isotope which undergoes two possible radioactive decays. These are

$$^{40}\mathrm{K} \rightarrow ^{40}\mathrm{Ca} + e^- + \bar{\nu_e}\,, \qquad (\beta - \text{decay, branching fraction: 89.3\%})$$

$$^{40}\mathrm{K} + e^- \rightarrow ^{40}\mathrm{Ar}^* + \nu_e \qquad (\text{electron capture, branching fraction: 10.7\%})$$
$$\hookrightarrow ^{40}\mathrm{Ar} + \gamma\,.$$

In the β-decay channel, the electron produced can have an energy up to 1.3 MeV, and so can produce Cherenkov light. In the electron capture channel, Compton scattering with the photon produces electrons which may also produce Cherenkov radiation [75]. ^{40}K decays can contribute to a significant fraction of the hits recorded by PMTs in the KM3NeT experiment. For example, ^{40}K decays can account for $\sim$4.7 kHz of the PMT single hit rate (per second) in the KM3NeT/ARCA environment [69]. The total background rate recorded per PMT in the KM3NeT experiment environment is 6-8 kHz [17], and so decays due to ^{40}K can account for $\sim$60-80% of the measured single hit rates.

The nature of these decays in the KM3NeT detectors is random. Due to the localised occurrence of these decays, they can be reduced from the data with simple filters (applied on shore, as explained later in this chapter), such as considering hits from more than one PMT on an optical module within a given time window.

Dark/Thermal Noise

There is a non-zero probability of thermal excitation of the PMT photocathodes, and the subsequent release of an electron, resulting in a signal from the PMT. The presence of an electric field also contributes to the possibility of electrons being accelerated inside the PMT environment. It is estimated that electronic noise in the PMT, along with natural radiation occurring in the glass of the PMT, can account for an additional single hit rate of $\sim$1.2 kHz [69, 76]. Thermal electrons can also be emitted from the anodes within the PMT, resulting in a signal from the PMT. Since this thermal noise pulse is lower in

amplitude than the thermal noise signal from the photocathode, the threshold signal pulse height of the PMT attempts to remove this signal.

Bioluminescence

Bioluminescence is the emission of light by living organisms. This phenomena occurs in various marine vertebrates and invertebrates, which dwell in the dark, deep sea. The light produced by these organisms can have a large impact on the photon sensors of KM3NeT, leading to light detection across one or more optical modules. Periods of strong sea currents correlate with high burst rates of bioluminescence in the detectors.

Atmospheric Muons

For both KM3NeT/ARCA and KM3NeT/ORCA, which aim to identify the muon tracks induced by neutrino interactions, a very significant background in the neutrino data analyses comes from atmospheric muons. These muons, introduced in Section 1.1.2, are created in cosmic ray air showers and travel to underwater depths and to the detectors, producing Cherenkov radiation as they do so. The light signals from these muons are identical to the muons produced from neutrino interactions.

Though the seawater surrounding and above the detectors provides a massive natural shielding from background particles, including atmospheric muons, some atmospheric muons still manage to reach the detectors. Due to energy loss processes, only high-energy muons can penetrate greater depths. A measurement of the atmospheric muon flux at underwater depths has been reported by the KM3NeT collaboration with data from the first deployed detection units [77]. The integrated atmospheric muon flux as a function of depth below sea level is shown in Figure 2.6.

An efficient method in removing atmospheric muons from the data in KM3NeT is simply by rejecting all muons with a downward-going direction. Since the chance of an atmospheric muon traversing the Earth is negligible, upward-going muons in the detector are certain to be those from neutrino interactions. However, a fraction of atmospheric muons will end up being mis-reconstructed as upward-going.

These atmospheric muons are far from simply useless background to KM3NeT. For instance, they can be used for calibration purposes, as mentioned in Section 2.5.2. In addition, these muons are invaluable to cosmic ray physics studies, as they provide a means of comparing simulations - assuming a particular cosmic ray flux and air shower

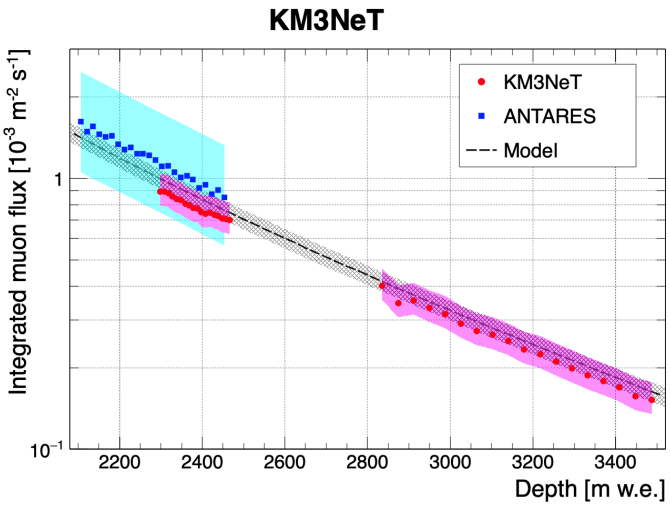


Figure 2.6.: The measured atmospheric muon flux as a function of seawater depth, from KM3NeT (red points) and ANTARES (blue points). The systematic errors are shown as the shaded pink region for KM3NeT, and as a blue shaded region for ANTARES. Also included is the Bugaev model of the atmospheric muon flux as a dashed black line, with uncertainties indicates by the grey shadowed area. From Ageron et al. (2020) [77].

simulation model - to data [78]. The higher energy muons detected by the KM3NeT detectors act as probes of the primary cosmic ray air shower interactions.

2.5. Data Acquisition

All hits recorded by the PMTs of the KM3NeT detectors, induced by neutrino signals or background, are sent to shore via a network of optical fibres and processed in real-time by a series of central processing units (CPUs). This is referred to as an '**all -data-to-shore**' principle. The collection and processing of this PMT signal forms the *data acquisition* stage of the KM3NeT experiment.

The optical fibre network allows for signals to be sent to control the optical modules and detection units. A bandwidth (the maximum amount of data transmitted over a given time) per optical module of 1 Gb/s is maintained. A relative timing offset of no more than 1 nanosecond between each pair of optical modules are ensured by the *White Rabbit* system which piggybacks on the network. This protocol ensures the time synchronisation of nodes in the optical Ethernet network to a master clock on shore [70]. The White Rabbit functionality extends to the optical modules through a 'White Rabbit Precision Timing Protocol' core on each optical module's FPGA. The clocks in the optical

modules are syntonized (operate on equal frequencies) to the master clock, but not necessarily synchronised (operating in phase).

The data from each optical module are sent to shore in data frames of a specified time period, usually of 100 ms. The data frames from all optical modules covering the same period are collected and sent to a node on the computer farm, and so different time periods are processed in parallel. In addition to this continuous physics data, summary data of the total number of hits per PMT are recorded with a sampling frequency of 10 Hz to inform on the actual optical conditions present in the detectors.

Acoustic data is also sent continuously from each digital optical module, in parallel with the optical data. This acoustic data, recorded using the piezo sensor on each optical module, can come from the acoustic emitters in the sea floor network or from other sources such as marine life. These piezo sensors take data with a sampling frequency of $\sim$195 kHz [70], which then transfer signal to the central logic board and then to shore via the optical fibre network. 'Monitoring data' is used for real-time online monitoring; this data shows for example the PMT rates. Additional information, termed 'slow control data', is also sent to shore. This slow control data contains information from the compass of the optical modules (e.g. yaw, pitch and roll information), and information such as the internal temperature of the optical modules.

2.5.1. Data Filters & Trigger Algorithms

All optical data and acoustic data are processed by the on-shore series of CPUs. This data flows through the following processes [79]:

$$\textit{DataQueue} \rightarrow \textit{DataFilter} \rightarrow \textit{DataWriter}\,.$$

Acoustic data for a specific optical module is sent by the *Acoustic DataQueue* to the *Acoustic DataFilter*, which performs a waveform selection depending on the desired signal waveform being analysed. The time of arrival of the signals are written to disk for positioning calibration purposes.

For optical data, the *Optical DataQueue* assembles the data from the same time period and sends them to the *Optical DataFilter* process. Here the data is grouped into data frames of 100 ms referred to as '**timeslices**'. The aforementioned summary data are coupled to the timeslices to form '**summary slices**'. In the *Optical DataFilter* process *trigger algorithms* (described below) are applied, to identify neutrino signal events from the background signal. If the trigger algorithms identify possible neutrino signals in the

timeslices, the first and last hit in the timeslice define the time window of a triggered event. The triggered event is written to disk by the *DataWriter* as a '**DAQEvent**'. Each triggered event is collected alongside all hits in the detector 10 microseconds before and after the first and last hits, respectively, to form a 'snapshot' of the event.

The total data rate for a detector building block of 115 lines will amount to about 25 Gb/s of data [17]. Assuming a background rate of 7 kHz per PMT (from potassium-40 decays and electronic noise), this amounts to $\sim$22 Gb/s of background data. This is a huge amount of data. In order to reduce the volume of data, and more importantly to select neutrino signals in KM3NeT, various levels of hit correlation are employed to filter the data.

Filter Levels

All hits, before any filtering or selection, are referred to as '**L0**' hits, i.e. the level-zero stage. Uncorrelated, random noise such as that from potassium-40 decays can be reduced by selecting correlated hits in the same optical module. These '**L1**' hits can be collected by selecting L0 hits which coincide between different PMTs on the same optical module within the same short time frame, usually within 10 ns[1]. Further level-two filters may be used, accounting for the PMT orientations within the optical module on L1 hits, resulting in '**L2**' hits. While neutrino events are more likely to induce L1 and L2 hits, these hits still contain background signal. For example, of the estimated rate of L1 hits per optical module of 1000 Hz, about 600 Hz is due to coincidences caused by potassium-40 decays [17]. The remaining 400 Hz comes from random coincidences which can be reduced by a factor of 2 in applying the L2 filter.

Trigger Algorithms

Neutrino-induced signal events in the KM3NeT detectors result in time-and-space-correlated hits. Dedicated trigger algorithms make use of coincidences between correlated hits to filter the data and extract neutrino signal events.

An important discriminator between background and signal events is in the maximal three dimensional distance condition for light (from e.g. a neutrino-induced muon) to travel through the detector. The difference in time between hits on PMT i and j should

[1] This value is used for physics purposes. For calibration work, longer time windows of up to 25 ns may be used.

be less than the time it takes for light to travel the distance between the PMTs. This can be written as:

$$|t_i - t_j| \leq |\vec{x}_i - \vec{x}_j| \cdot \frac{n}{c}\,, \tag{2.8}$$

where x indicates the position of the PMT, t represents the time of the hit, n is the refractive index in seawater and c is the speed of light.

This condition can be made more restrictive when assuming a direction of travel of a particle emitting Cherenkov radiation under an angle θ_C:

$$|(t_i - t_j)c - (z_i - z_j)| \leq \sqrt{\left((x_i - x_j)^2 + (y_i - y_j)^2\right)} \tan\theta_C\,, \tag{2.9}$$

with x, y, z denoting the position of the PMT. Finding a set of N hits that are pair-wise correlated according to this criteria is the basis of the trigger algorithms. These algorithms in turn employ cluster algorithms, to identify correlated hits — an example of addressing the *clique problem* [80].

There are three main trigger algorithms implemented for the filtering of data arriving on shore. Since neutrino events can be described by distinctive track and shower-like event signatures (as discussed previously), the trigger algorithms take advantage of this fact in discriminating neutrino events from background. These algorithms are as follows.

3DShower trigger. Here the maximal 3D distance is applied. A set of n causally connected L2 hits are considered, on an m number of optical modules within a sphere of radius r, corresponding to the maximal distance between PMTs.

3DMuon trigger. A scan over the full sky, and the directional filter condition of Equation (2.9) is implemented. A set of n causally connected L2 hits are considered within a cylinder of infinite length and radius R, on m optical modules. R can be set to a few times the absorption length of light (how far light can travel in seawater before it is absorbed). The directional condition assumes the direction of the muon, and this constrained cylinder of hits is scanned around a grid of the full sky. Further implementation of this geometry is used in the reconstruction software, outlined in later chapters.

A 'mixed' algorithm is also used: the **MXShower trigger**. As shown in Figure 1.4, at neutrino energies of $\sim$7–100 GeV, deep inelastic scattering is the dominant interaction mechanism. Many events have a signature which can resemble a track and a shower, due to the creation of a hadronic shower in all interactions. This trigger aims to identify

such events. In this algorithm, set of n causally related L0 hits are considered within a sphere of radius r, in addition to one L2 hit, on m number of optical modules.

The values of n, m, R are adjusted for the different trigger algorithms and KM3NeT detectors. Example values in KM3NeT/ORCA for the 3DMuon trigger are: $n = 5$, $m = 5$, $R = 120$ m, and $n = 5$, $m = 5$, $r = 250$ m for the 3DShower trigger.

2.5.2. Calibration Methods

As mentioned, the clock for each respective optical module is set to the time received from the master clock (at the shore station). Time delays exist in the optical fibre network, as it takes time for signals to travel to each optical module in a detection unit.

These time delays, and the transit time of the PMT, results in an overall **PMT time offset** for each PMT in the detector. Calibration efforts within KM3NeT attempt to correct for these time offsets (and others, described below), as well as in determining the physical positions of the optical modules when operating in the sea. Calibrating these time offsets in the detector with sub-nanosecond timing accuracy, and the positions with an accuracy of 20 cm, allows for the KM3NeT experiment to achieve its science goals. Each data run (a discrete period of data taking) is calibrated using the methods outlined in this section.

A pre-requisite for timing calibration procedures is that all PMTs in an optical module have the same required gain value, ensuring a homogeneous response for every PMT. A procedure referred to as '**high-voltage tuning**' is applied to PMTs in the laboratory (before integration into an optical module) and during operation of the detector under the sea, to set the PMT voltage which allows for an operational gain (see Section 2.1.2) of 3×10^6. The voltage across each PMT is tuned such that the mean value of the resulting time-over-threshold distributions corresponds to the required gain [70, 81].

Various methods are used to correct for the timing offsets. The dominant background sources of atmospheric muons and ^{40}K decays, both being well-understood signals, provide excellent means of studying PMT time offsets.

Inter-PMT calibration

It is important to determine the **relative PMT timing offset**, which is the difference between the PMT time offset and the digital optical module timing offset (defined below)

for the optical module in question. Potassium-40 decays are taken advantage of to determine the relative PMT time offset, from data taken *in situ* (i.e. data taken when the optical modules are operating in the sea). The coincidence rate for a pair of PMTs, as a function of the hit time difference between the PMTs, can be modelled as a Gaussian distribution, the mean of which depends on the difference in time offsets between that particular pair. A Gaussian distribution is used since the transit time spread of each PMT can be approximated by a Gaussian distribution. These offsets can be measured by fitting Gaussian functions to data from potassium-40 decays. This inter-PMT calibration method is outlined in Melis (2021) [69].

Inter-Digital Optical Module calibration

The purpose of inter-digital optical module calibration is to measure the time offsets of individual optical modules on a single detection unit.

One method of measuring these offsets is in the laboratory pre-deployment, with a laser pulse illuminating a number of PMTs per optical module. These light signals arrive on the PMT of each optical module with a fixed time delay. The time at which the laser is fired and the measured arrival time from the PMT provide these offsets so that the full detection unit is calibrated in time [82, 83].

The nano-beacon of each optical module can be 'flashed' for the purpose of measuring timing offsets of optical modules. These nano-beacon pulses have enough intensity to illuminate the optical modules on the same detection unit above the reference optical module [84]. These nano-beacons are flashed when the detection units are operating underwater.

Atmospheric muons also provide a means of time calibration for the detectors. One calibration method utilising atmospheric muons for calibration is referred to as '**L1 dt calibration**'. A likelihood function is used to describe the most probable time offsets given a set of L1 hits across multiple optical module pairs in the detector, for L1 hits from data and simulation [85]. This likelihood function is maximised to find the most probable values of the time offsets between optical modules along a detection unit, given a reference optical module. Maximum likelihood estimation is elaborated on in the following chapter, and is used for reconstruction purposes.

A second means of calibration with atmospheric muons can be carried out by calculating the **hit time residuals**, the difference between measured L0 hit times and the *expected arrival time* of light from atmospheric muons (discussed further in Chapter 3). In this

procedure, hit time residual distributions are estimated for every pair of optical modules, and their values are minimised for various detector configurations, which include optical module positions and timing offsets. A large set of hit time residuals close to zero for a given detector configuration indicates the most likely detector configuration (as time residuals close to zero imply the atmospheric muons have been reconstructed accurately, which occurs for a well-calibrated detector).

Acoustic Positioning Calibration

It is imperative to determine the real-time physical positions of the detection units and optical modules once they are operating in the sea. The positioning of the optical modules is required in the simulation and reconstruction of neutrino events, with the most accurate knowledge of the positions of the optical modules resulting in the most accurate determination of the neutrino direction and the true location of the resulting Cherenkov light detection. A positioning accuracy within $\sim$20 cm is aimed for in the KM3NeT infrastructure [17]. This goes hand-in-hand with the aim of a timing accuracy of 1 nanosecond; both the timing and positioning information of the hits need to be known with these error margins.

As noted in Section 2.2.2, a network of acoustic emitters is under construction as part of the KM3NeT infrastructure. This network includes autonomous beacons which reside on tripods, and other acoustic emitters which are placed on the junction boxes of the sea floor network. These all emit acoustic pings, a signal of known pulse shape, at a known frequency. A model which describes the arrival time of acoustic pings on the optical modules, dependent on the ping emission time and the location of the optical modules, is fitted to the data. In this way the positions of the optical modules are determined. These positions are subsequently used in the KM3NeT software packages, described in Chapter 3. Such position and orientation calibration is carried out every ten minutes.

2.6. Research Goals

As outlined in the KM3NeT Letter of Intent (2016) [17], the primary goal of KM3NeT/ARCA is the detection of high-energy astrophysical neutrinos of cosmic origin, and the identification of astrophysical objects, including cosmic ray accelerators, as a source of such neutrinos. In 2019, the IceCube neutrino experiment at the South Pole iden-

tified a neutrino which may have originated from a known blazar, TXS-0506+056[2] [86]. Though the astrophysical neutrino was found to be coincidental with the time and direction of a gamma ray flare from the blazar with a significance of 3.5σ, a 5σ discovery of a neutrino coincidental with the direction of an astrophysical source remains an open challenge.

In 2013, the IceCube Collaboration reported a measurement of a neutrino of extragalactic origin [87]. Figure 2.7 shows the discovery sensitivity of KM3NeT/ARCA to a diffuse neutrino flux corresponding to the signal measured by IceCube, as a function of observation time. The black and red lines indicate the track and shower-like signals in the detector, respectively. The blue line is the significance for the combined signal as a function of time. Clearly, KM3NeT/ARCA offers an extremely competitive discovery potential to the field of neutrino astronomy, claiming a 5σ discovery potential in little over half a year. Note that this figure may be understating the potential of KM3NeT/ARCA, as improvements have since been added to the KM3NeT software and analysis procedures.

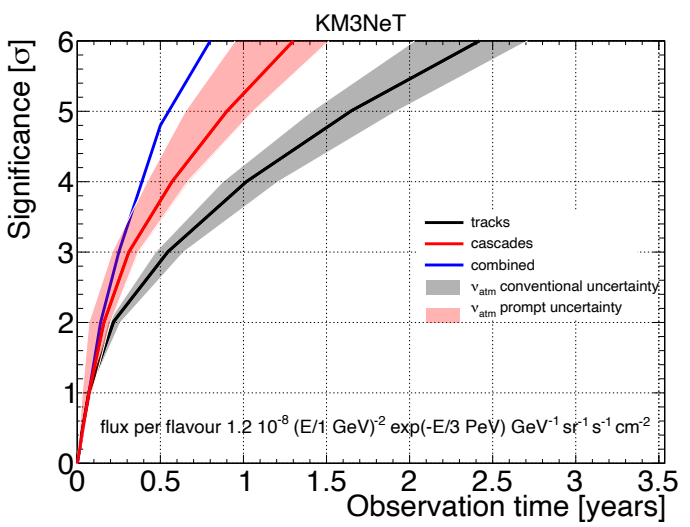


Figure 2.7.: The discovery significance of KM3NeT/ARCA to the diffuse neutrino flux, as reported by IceCube, as a function of observation time. In black is the track signal, in red the shower signal, and in blue the combined sensitivity from both signals. Also indicated is the error on both channels. From Adrían-Martínez et al. (2016) [17].

Located in the Northern Hemisphere, KM3NeT/ARCA will be strategically positioned to study neutrinos coming from the Galactic Centre of the Milky Way, a location of nearby potential neutrino sources. This detector will also contribute to the multi-messenger network that is being established across the globe: a collaboration of various experiments detecting different astrophysical messengers of photons, gamma-rays, neu-

[2] A blazar is a relativistic jet of material emitted from an active galactic nucleus (AGN): the compact region of a galactic centre with a high electromagnetic radiation output.

trinos, cosmic rays and gravitational waves, in order to carry out astroparticle physics and astronomy.

One means of characterising a neutrino detector is by the number of neutrino events KM3NeT/ARCA will detect, given a detector volume and assumed neutrino flux. This is referred to as the '**effective area**' of a detector, and can be defined as

$$A^{\text{eff}}(E_{\nu_l}) = \frac{R_{\text{triggered}}(E_{\nu_l})}{\Phi(E_{\nu_l})}, \tag{2.10}$$

where $R_{\text{triggered}}$ is the rate of neutrino events detected at the trigger level and Φ is the flux incident upon the Earth, both as a function of the neutrino energy and flavour E_{ν_l}. The expected effective area of KM3NeT/ARCA is shown in Figure 2.8 as a function of the neutrino energy, for the three neutrino flavours. The events shown here have passed through the trigger algorithms used to identify neutrino events. The Glashow resonance (see Section 1.2.2) is visible for the electron neutrino channel.

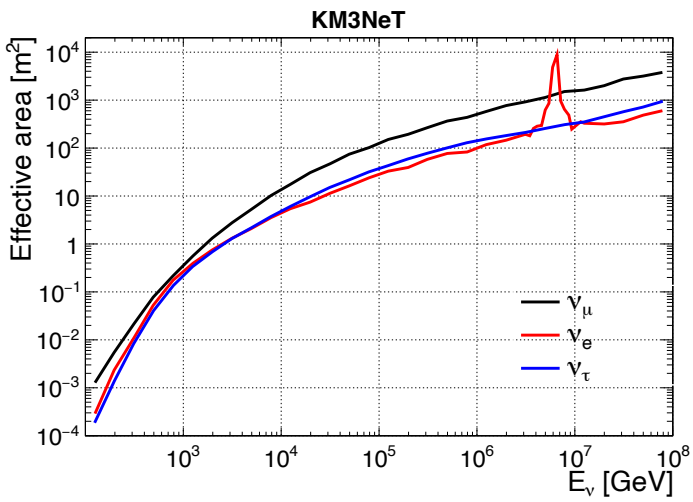


Figure 2.8.: The expected effective area of KM3NeT/ARCA for the three neutrino flavours, as a function of the neutrino energy, averaged over neutrinos and anti-neutrinos and including charged and neutral-current interactions. The Glashow resonance, described in Section 1.2.2, is visible for the electron neutrino channel. Taken from Adrían-Martínez et al. (2016) [17].

The primary focus of KM3NeT/ORCA is the determination of the neutrino mass ordering, described in Section 1.2.4. It has been estimated that after 3 years of data taking, KM3NeT/ORCA will achieve a 4.4σ sensitivity to the neutrino mass ordering if it follows normal ordering, and 2.3σ if the neutrino mass ordering follows the inverted ordering. This sensitivity is shown in Figure 2.9.

A precise measurement of the neutrino oscillation parameters θ_{23} and Δm^2_{32} is expected, narrowing the uncertainty range of these parameters given in Figure 1.11. θ_{23} is expected to be precisely measured to [+1.9°,-3.1°] for normal ordering and [+2°,-7°] for inverted ordering [88]. A measurement precision of 85×10^{-6} eV^2 or 75×10^{-6} eV^2 is expected for normal or inverted ordering, respectively, for Δm^2_{32}. Figure 2.10 indicates the precision of KM3NeT/ORCA in determining these values, compared to other neutrino oscillation experiments. These expected results were determined in 2021. Updates to the reconstruction and analysis methods have improved since, improving the results.

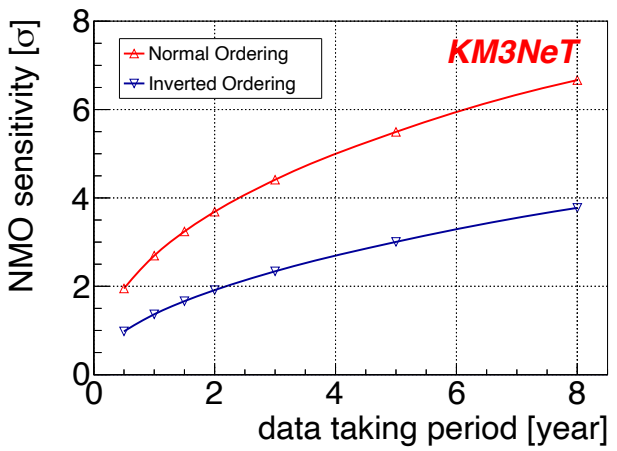


Figure 2.9.: The sensitivity of KM3NeT/ORCA to the neutrino mass ordering as a function of time for both potential orderings: inverted ordering (blue) and normal ordering (red). Taken from Aiello et al (2021) [88].

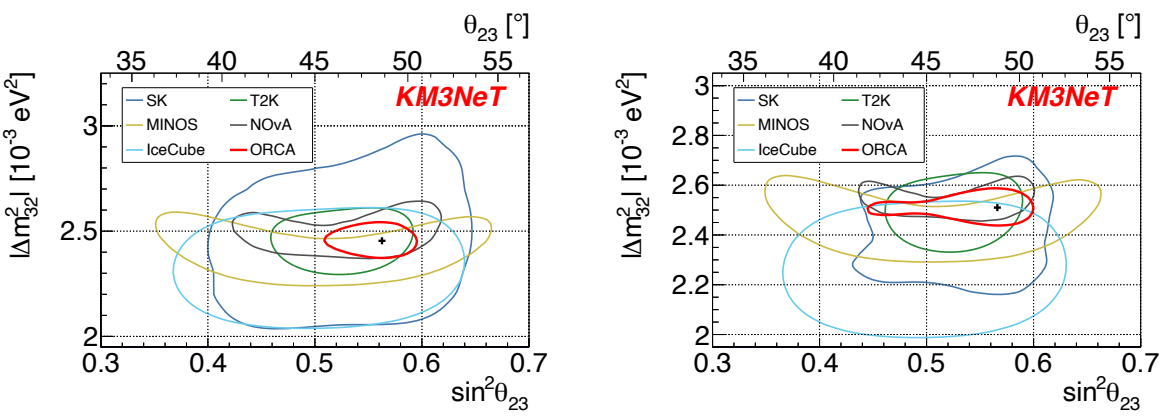


Figure 2.10.: Expected precision of KM3NeT/ORCA to the neutrino oscillation parameters of θ_{23} and Δm^2_{32}, for 3 years of data taking, for both potential mass orderings: (a) normal ordering (b) inverted ordering. Also shown are the expected precision contours for other experiments. Taken from Aiello et al. (2021) [88].

KM3NeT will also contribute to cosmic ray physics. Serving as an atmospheric muon detector, measurements of atmospheric muon properties such as the flux will contribute to the field. Since the detectors are at two different locations at different depths, KM3NeT/ARCA and KM3NeT/ORCA offer a unique measurement of atmospheric muons at different depths with the same experiment.

Considering the environment in which KM3NeT is located, collaboration with sea science institutes is also made possible with this experiment, e.g. in monitoring marine life.

2.7. Current Status

By the time this thesis has come to publication, some of the following information may be obsolete. KM3NeT/ARCA is currently operating with 28 detection units on the sea floor. KM3NeT/ORCA is taking data with 18 deployed detection units. These sub-detectors of KM3NeT are referred to as 'ARCA-x'/ 'ORCA-x', with x the number of detection units.

As an indication of the potential for KM3NeT, the simulated effective area is shown in Figure 2.11 for ANTARES and the KM3NeT/ARCA-6–21 detectors as a function of energy, for only the muon neutrino and anti-neutrino channel. KM3NeT/ARCA operated with 6 detection units in May 2021, before the deployment of additional detection units up to September 2023. ANTARES [89] is the predecessor to KM3NeT, an underwater neutrino detector distributed over 0.1 km^2, which operated from 2008 to 2022. It incorporates a different technology to KM3NeT, with a single 10-inch PMT forming the optical module, 3 of which are clustered together. With a limited number of detection units, all of the ARCA-few-line detectors (which comprise only 3–9% of the envisaged complete detector) show a better expected performance in the energy region of interest compared to ANTARES.

Another sign of great potential from KM3NeT is in the measurement of θ_{23} and Δm^2_{32} by the KM3NeT/ORCA-6 detector. ORCA-6 took data between January 2020 – November 2021 before the addition of more detection units. Figure 2.12 shows the precision in measuring these parameters from the ORCA-6 data [91]. This amounts to 510 days of data, and taking into account the instrumented volume of the detector, the data-taking time is presented as 433 kton-years of data. A comparison is also shown compared to a previous, preliminary analysis of 269 kton-years worth of data from the ORCA-6 detector [85].

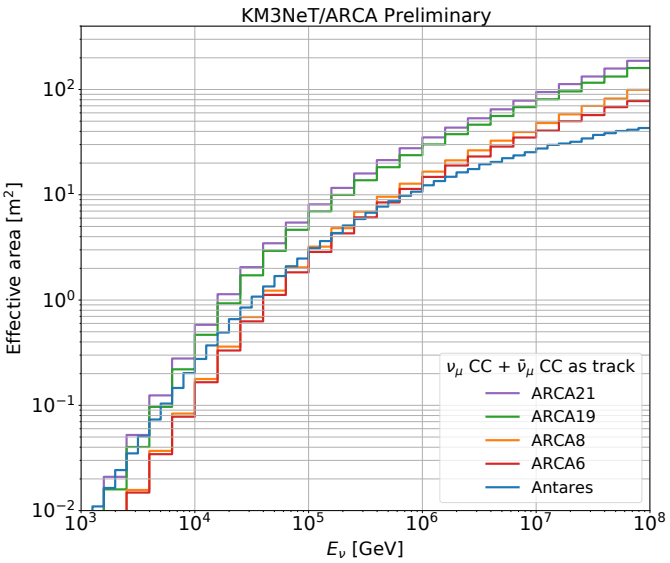


Figure 2.11.: The effective area of KM3NeT/ARCA-6–21 and ANTARES as a function of energy for the muon neutrino and anti-neutrino channel. Even with a few detection units, KM3NeT shows similar or better performance compared to ANTARES, its predecessor. From Muller et al. (2023) [90].

As KM3NeT/ORCA expands in size, its precision in measuring neutrino oscillation parameters of interest becomes greater. With 115 detection units, we recover the precision of Figure 2.10.

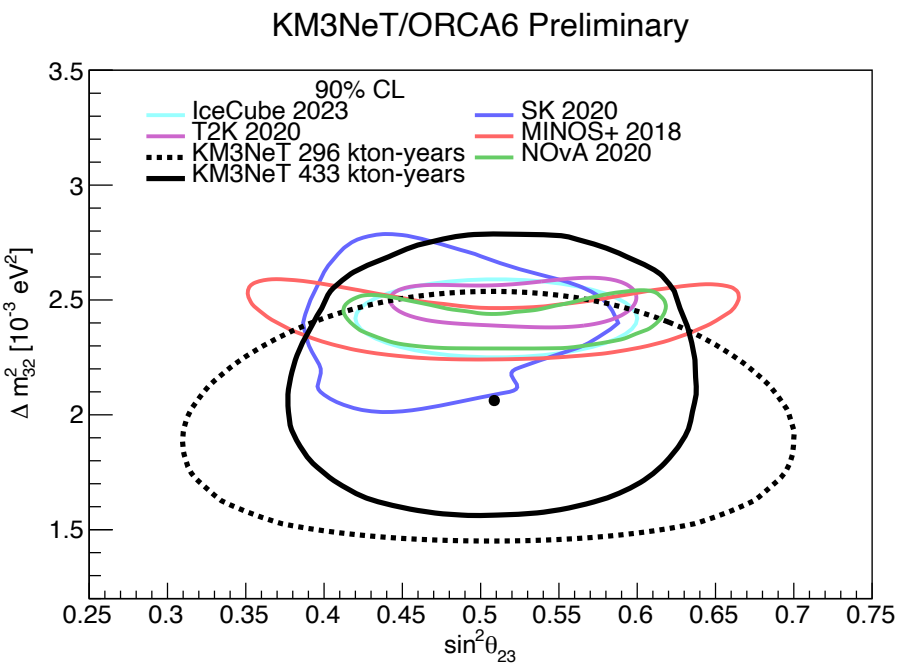


Figure 2.12.: The precision region to measuring θ_{23} and Δm^2_{32} from ORCA-6 data (solid line and best-fit point), in comparisons to other neutrino oscillation experiments [91]. The precision as determined from a previous analysis (dashed line) is also shown [85].

Chapter 3.

Physics Models & Reconstruction

In this chapter, the methods and physical models required for simulating events in the KM3NeT detectors are described. The simulations employed to do so incorporate our current understanding of the KM3NeT detectors and the physics of the measured signal and backgrounds. Confronting these simulations with the data allows for validating the physical models used, for understanding the detector response to the signal, and for testing software and improving measurements of the desired signal.

Reconstruction procedures are applied to both simulated data and data recorded by the detectors themselves, in order to estimate the direction and energy of track-like or shower-like events. These procedures are also described in this chapter.

3.1. Simulation Chain in KM3NeT

The sequence of software programs used to simulate and reconstruct KM3NeT data is here referred to as the 'simulation chain'. The various stages of the simulation chain will be described in this chapter. Particular attention is given to the parameterised description of the atmospheric muon flux, and to the reconstruction methods used, since they form the background knowledge for later chapters.

The simulation chain follows the order of the following (informally named) stages:

$$\text{Generation} \rightarrow \text{Light} \rightarrow \text{Trigger} \rightarrow \text{Reconstruction}\,.$$

In the generation stage, either neutrino interactions are simulated and their resulting particles are tracked through the KM3NeT detector in question, or the atmospheric

muons from cosmic ray air showers are simulated. These particles are then propagated through the detector and their Cherenkov light and detected number of photo-electrons by the PMTs is simulated in the light stage. In the trigger stage, the detector response is simulated, and the trigger algorithms are applied to filter simulated events of interest. The triggered events, composed of hits produced by the various particles, are then used in the reconstruction stage to determine the direction and energy of track-like or shower-like events.

The detector configuration is defined throughout the simulation chain using specific detector files, defining the number of PMTs and their position, orientation and timing offset. Information on the PMT quantum efficiencies, gain and timing offsets is defined in a separate PMT parameters file. This PMT information is usually set to nominal values or else updated following the calibration procedures outlined in Section 2.5.2. The positions of the optical modules can be simulated as static (e.g. if the positions are not known) with some default rotation, or are determined through acoustic position calibration (see Section 2.5.2); the correct positions and rotations can be deduced from the acoustic fit model parameters stored in separate files. This detector and PMT configuration information is used in the light, triggering and reconstruction software and is thus relevant to both simulations and recorded data.

All simulations are based on the concept of *Monte Carlo methods*. These methods are used in simulations, including the simulation chain stages defined above, to produce a set of n simulated data points, distributed according to some *probability density function*. These functions describe the probability that a random variable will fall within a particular range of values in the sample space. A random sequence of numbers is used to construct a sample of a population, 'pulled' from (i.e. generated according to) these probability density functions. Monte Carlo methods are described in detail in e.g. Metzger (2002) [92]. They are alluded to in this chapter, for example in the generation of neutrino events and the simulation of atmospheric muons from a parameterised flux.

In the simulation chain, some software programs take advantage of pre-calculated tables of the expected number of photo-electrons as a function of the photon arrival time on a PMT. These tables are widely applicable in the simulation chain, since they provide a means of simulating the number of detectable photo-electrons in the light stage, and in defining hit probabilities to be used in the reconstruction procedures outlined later in this chapter. In the following sections, the method used in calculating the expected number of photo-electrons for these tables is first described, followed by the simulation chain in detail.

3.2. Photo-Electron Distributions as a Function of the Photon Arrival Time

The KM3NeT collaboration makes use of distributions which describe the expected number of photo-electrons recorded by a PMT, as a result of Cherenkov light from a muon or a particle shower. These values are saved in the form of data look-up tables. Throughout this thesis, these tables shall be referenced. For the sake of brevity, these tables will be referred to as '**photo-electron distribution tables**'.

The number of photo-electrons recorded by a PMT at a particular time, resulting from various Cherenkov light emission processes, can be deduced through different methods. One means of doing this is to simulate (using Monte Carlo methods) the exact number of photo-electrons recorded by PMTs in the KM3NeT detectors for a given event type, such as a muon track or particle shower, for a specific set of parameters describing the event (such as direction, energy, and position relative to the PMT). This requires an extensive and repeated simulation of such events, in order to get a smooth, tabulated distribution. Another means of doing so is through analytical formulae which describe the expected photo-electron yield at a given time for muon and particle shower events. Using such formulae derives from first principles but requires some approximations to simplify things. This latter method is used to produce these tables of photo-electron distributions. These tables are produced within the `Jpp` software framework, a set of software tools, methods and applications in the `C++` programming language, used within the KM3NeT collaboration.

3.2.1. Coordinate System

The photo-electron distributions as a function of the arriving photon time are defined and calculated according to a specific coordinate system, namely one in which the muon or shower direction is pointing along the z-axis of the standard Cartesian coordinate system [93], as shown in Figure 3.1. The extrapolated muon or shower trajectory passes through the coordinate origin (0,0,0). The PMT is positioned in the $x-z$ plane, with an orientation defined by the angles $\theta_{\wp}$ and $\phi_{\wp}$. The angle $\phi_{\wp}$ is mirrored around the $x-z$ plane. The distance of closest approach R is defined as the minimal straight-line distance between the muon or shower and the PMT. For the shower however, it is more intuitive to use the distance from the Cherenkov light emission point to the PMT, denoted by D. The angle θ_0 represents the angle under which the light is emitted.

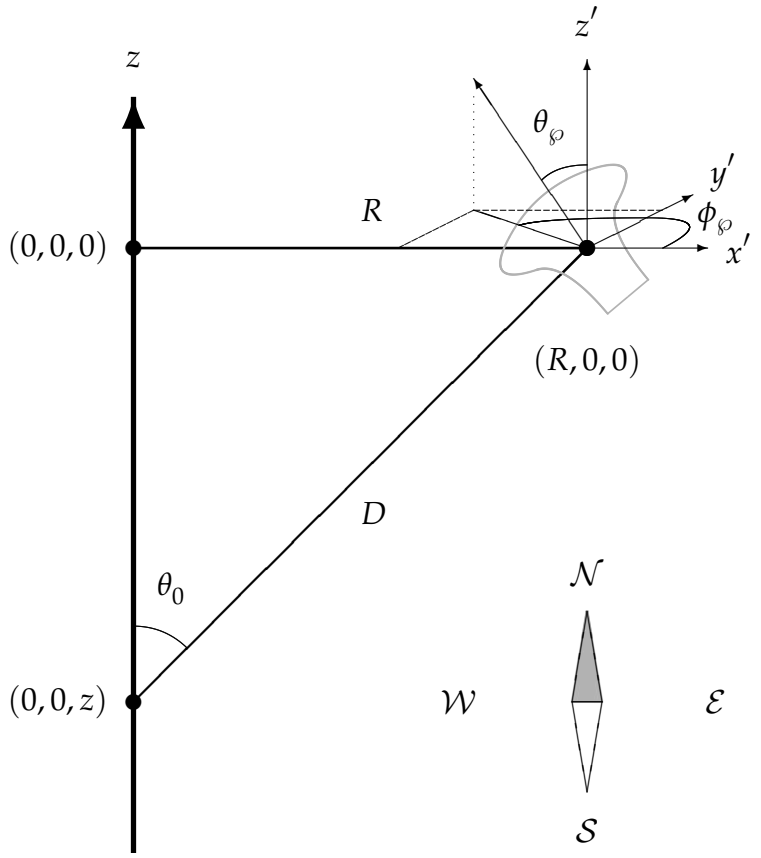


Figure 3.1.: The coordinate system in which the photo-electron distributions, according to muon track or shower events, are estimated. The muon or shower points in the z-direction, with the PMT described by an orientation according to the angles ($\theta_{\wp}$, $\phi_{\wp}$) and a position in the $x-z$ plane. The distance of closest approach between the PMT and muon or shower is represented by R, such that the PMT is located at position (R,0,0). The distance between the light emission point and the PMT is denoted by D. θ_0 is the emission angle of light. The compass directions are used to illustrate possible PMT orientations, described in Figure 3.2. Figure taken from M. de Jong (2012) [93].

3.2.2. Calculation of the Number of Detectable Photons

The number of detectable photons are calculated for direct and single-scattered Cherenkov radiation, assuming a muon or a particle shower. The ingredients entering this calculation are summarised in this section. Accounting for single-scattered radiation is an approximation, though the scattering length in water is large enough to justify this [93]. The calculation of multiple-times-scattered radiation is even more complex.

Assuming a light cone from Cherenkov radiation, with the characteristic Cherenkov emission angle $\theta_0 = \theta_C$ (as per Equation (2.3)), the number of detectable photons per unit area (where x is the unit track length) and per unit wavelength at a distance R can

be expressed as [93]

$$\Phi_0(R,\lambda) = \frac{d^2N}{dxd\lambda}\frac{1}{2\pi R\sin\theta_C}, \tag{3.1}$$

following from Equation (2.4).

For particle showers, hadronic and electromagnetic showers are treated in the same way. The number of detectable photons N per unit shower energy E and per unit wavelength λ can be expressed as

$$\frac{d^2N}{dEd\lambda} = \frac{dx}{dE}\frac{d^2N}{dxd\lambda}, \tag{3.2}$$

where x is the equivalent track length per unit shower energy. The angular distribution of light emission from an electromagnetic particle shower can be parameterised as [93]

$$\frac{dP_*}{d\cos\theta_0 d\phi_0} = c\cdot e^{b|\cos\theta_0-\cos\theta-\cos\theta_C|^a}, \tag{3.3}$$

where θ_0 is the emission angle of light — considering the averaged emission angle of light from particles in the shower, this is not necessarily the Cherenkov angle θ_C. Here, ϕ_0 is the angle of the PMT with respect to the shower, and a, b and c are determined constants. The parameter c is defined such that P_* (the light contribution) is normalised to unity for the full solid angle.

The number of detectable photons per unit wavelength, per unit energy and per unit solid angle can thus be expressed as a function of the angle of emission by

$$\Phi_1(\cos\theta_0,\lambda) = \frac{d^2N}{dEd\lambda}\frac{dP_*}{d\cos\theta_0 d\phi_0}. \tag{3.4}$$

The ionisation and radiative energy loss of the muon is given by Equation (2.6). In addition, the energy loss due to secondary electrons (those generated as a product of ionisation) referred to as delta rays can be found in Zeyla et al. (2020) [3]. These equations are used to modify the calculation of the number of detectable photo-electrons in $\frac{d^2N}{dEd\lambda}$, incorporating muon energy loss processes.

The coordinate system defined above is used to define the derivative of the arrival time of light on the PMT as a function of the wavelength, i.e. $\delta t/\delta\lambda$.

3.2.3. Numerical Computation & Examples

For direct Cherenkov radiation from a muon, the distribution of photo-electrons as a function of the incident photon arrival time is given by

$$\frac{d\mathcal{P}}{dt} = \Phi_0(R,\lambda)\, A \left(\frac{\delta t}{\delta \lambda}\right)^{-1} \varepsilon\, QE(\lambda)\, e^{-d/\lambda_{abs}}\, e^{-d/\lambda_s}\,. \tag{3.5}$$

Here, A is the photocathode area of the PMT, ε is the angular acceptance (the sensitivity as a function of incident angle) of the PMT, d is the distance travelled by the radiation, QE is the quantum efficiency (see Equation (2.5)), and λ_s and λ_{abs} are scattering and absorption length parameters, respectively. The quantum efficiency and angular acceptance are included to account for the light detection efficiency. This equation is shown as an example of such a calculation. The same distributions for shower light and for muonic energy loss processes are calculated in a similar manner, for both direct and single-scattered radiation.

These functions are evaluated numerically, involving multi-dimensional integrals. Their values for a set of muon or shower and PMT parameters in the parameter space are stored in the form of data tables. The use of multi-dimensional interpolations provides access to the values, their derivative and partial integrals [94]. The PMT response to detecting the light can be accounted for by convoluting the arrival times according to a Gaussian distribution with a width equal to a chosen PMT transit time spread.

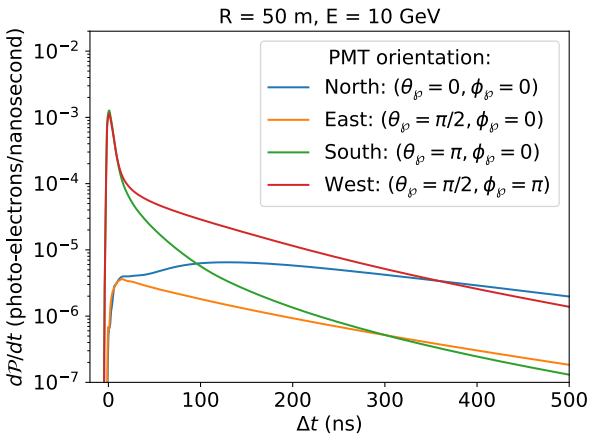
While these distributions describe the number of photo-electrons per nanosecond, the total number of photo-electrons up until a given time can also be calculated and stored in data tables. These cumulative distributions are useful when calculating the total number of hits to expect on a PMT given a particular source of Cherenkov radiation, and are used in the software used to simulate the light stage of the simulation chain.

Two examples of the number of photo-electrons per nanosecond ($d\mathcal{P}/dt$) as a function of the time difference between the arrival time t and the expected arrival time of light t_0 on a PMT are illustrated in Figure 3.2.

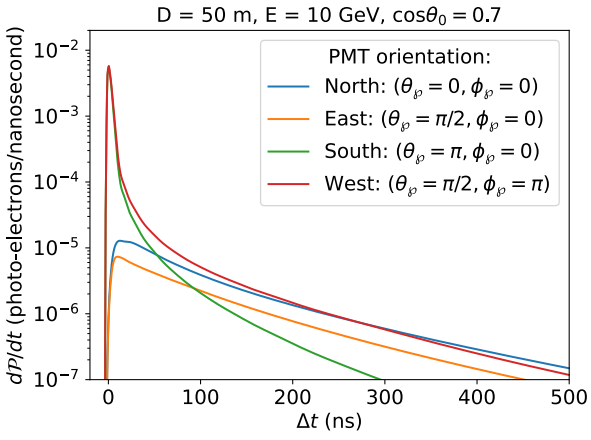
This quantity

$$\Delta t \equiv t - t_0\,, \tag{3.6}$$

is referred to as the **time residual**. The value of the time residual is zero when the observed arrival time coincides exactly with the expected arrival time. For the muon, the



(a)



(b)

Figure 3.2.: The number of photo-electrons per nanosecond as a function of the time difference between the actual and expected arrival time of light, assuming a Cherenkov cone hypothesis, for (a) muon light and (b) light from an electromagnetic shower. The energy used is 10 GeV, with a distance from the emission point to the PMT D of 50 m for the shower. A cosine of the emission angle $\cos\theta_0 = 0.7$ is assumed for the shower. Reproduced from M. de Jong (2012) [93].

expected arrival time of light on a PMT due to Cherenkov radiation is defined as

$$t_0^{\text{muon}} \equiv \frac{R \tan\theta_C}{c}, \tag{3.7}$$

where c is the speed of light *in vacuo* and θ_C is the characteristic Cherenkov angle. The value of θ_C used (for direct Cherenkov radiation) is $\sim 42.565^\circ$, calculated as an average over a range of wavelengths. Here t_0 represents the shortest path for Cherenkov light to reach the PMT from its emission point along the muon trajectory. Appendix A contains a

derivation of t_0^{muon}. For the case of an electromagnetic shower,

$$t_0^{\text{shower}} \equiv \frac{D \cdot n}{c}, \tag{3.8}$$

where n is the refractive index. It is also derived in Appendix A.

The number of photo-electrons as a function of the time residual for a muon and particle shower are shown in Figure 3.2 (a) and (b), respectively. An energy E = 10 GeV is used for both cases, with R = 50 m assumed for the muon, while a cosine of the emission angle $\cos\theta_0 = 0.7$ is assumed for the shower with a distance between the shower and the PMT of D= 50 m. Other parameter values, such as that of the angular acceptance, are the same as in M. de Jong (2012) [93]. In Figure 3.2 (a) the total number of expected photo-electrons as a function of the time residual, as a result of direct and scattered light from the muon, direct and scattered light from electromagnetic showers along the muon track, and direct and scattered light due to delta rays is shown. In the case of Figure 3.2 (b), the total number of expected photo-electrons as a function of the time residual, for direct and scattered light from an electromagnetic shower, is shown. Four different PMT orientations, determined by the values of $(\theta_{\wp}, \phi_{\wp})$ are used, corresponding to the orientation displayed in Figure 3.1.

For a PMT orientated 'West' or 'South' (looking at the event), direct and scattered light is detected, with a sharp peak at $\Delta t \sim 0$ ns. In the case of the PMT oriented 'East' or 'North' (looking away from the event), scattered light alone is detected. These orientations and the resulting light yield are determined by the coverage of the track segment by the PMT. Note that the electromagnetic shower results in more light arriving at the PMT than from a muon in this example.

3.3. Event Generation

The term 'event' is used in this section to refer to the primary particles (such as neutrinos or cosmic rays) which are generated with a unique timestamp, and the associated particles which result from their interactions.

3.3.1. The Can

The '**can**' is an imaginary volume around the instrumented volume of the detector (be it KM3NeT/ARCA or KM3NeT/ORCA). The can serves as the volume in which Cherenkov

radiation from particle interactions may reach the detector. Taking the instrumented volume of the detector as a cylinder, this volume is then enlarged by a number of times the absorption length of light in seawater to obtain the can. All of the events leading to Cherenkov radiation in the can are stored for the subsequent particle and light propagation stage of the simulation chain.

The **interaction volume** is the volume in which a neutrino interaction can produce detectable particles, and is either larger than or the same as the can volume, for the cases of track-like and shower-like events, respectively. In the case of charged-current electron neutrino interactions and all neutral-current interactions (so, shower-like events), the interaction volume is the same as the can, where light may reach the detector. In this case the target media of the interaction volume are the most abundant elements found in seawater. For track-like events, these may be detected even if the neutrino interaction occurs outside of the can. The interaction volume is defined using the maximum range of a muon in rock and water. Since tau neutrino decays may result in muons, the interaction volume for charged-current tau neutrino interactions are also treated using the maximum muon range. The interaction volume for track-like events is a cylinder made of a layer of rock and a layer of seawater. For atmospheric muons, on the other hand, these are simulated starting from the point where they enter the can.

3.3.2. Event Weights

The particles one wishes to generate, be they neutrinos or primary cosmic rays, are usually generated according to an energy spectrum from which the energy of the particles is sampled. The energy spectrum used in particle generation may not be the same as the physical energy spectrum. For example, for high-energy particles one expects fewer particles as the energy is increased. Using a physical spectrum would result in few particles at high energy and result in wasting computational resources on the lower energy ones. A 'non-physical' flux is often used, where the physical spectrum can be re-obtained through **event weights**. Important physical effects, such as the neutrino interaction cross section and oscillation of neutrinos after passage through the Earth, are accounted for in this 'non-physical' flux.

The event weight is calculated from

$$W_{\text{event}} \equiv \frac{1}{f(E,\theta,\phi)^{\text{generation}}} \cdot f(E,\theta,\phi)^{\text{physical}};$$
$$W_{\text{generation}} \equiv \frac{1}{f(E,\theta,\phi)^{\text{generation}}},$$

where W_{event} is the event weight, $W_{\text{generation}}$ is the generation weight and $f(E,\theta,\phi)$ represents the spectrum, which depends on the energy E and the zenith and azimuth directions θ and ϕ, respectively. In this formulation, the spectrum according to which events are generated takes into account normalisation terms, such as the number of events generated and the direction and energy ranges considered. Calculating the event weight allows for simulating high-statistics samples of events across a wide energy range, and allows for changing the physical flux without re-generating events.

3.3.3. Neutrino Interaction Simulation

In simulating neutrino events, the neutrino is generated at a vertex position within the interaction volume, where it produces other particles. The neutrino energy is sampled from an energy spectrum of the form $dN/dE \propto E^{-\gamma}$. The energy spectrum can be for atmospheric neutrinos, or cosmic neutrinos from astrophysical sources such as diffuse, point-like or extended sources. The neutrino direction is assumed to be isotropic. The scattering angle of the neutrino and its momentum transfer to target nucleons are sampled from the relevant cross section models. All three neutrino flavours are simulated, for both charged and neutral-current interactions. In practise it is sufficient to simulate the neutral-current interactions for one flavour only and weight the events correctly, since neutral-current interactions will lead to the same final state particles.

The neutrino vertices may be simulated on a surface outside of the interaction volume. This is done for example by the gSeaGen [95] software program, a GENIE [96]-based code for simulating neutrino interactions for the purposes of neutrino experiments such as KM3NeT. The generated neutrinos then interact within the interaction volume. GENIE is developed to simulate neutrino interactions for neutrino oscillations studies, particularly in the few-GeV energy regime. Within gSeaGen, differential cross sections for radiative processes and nuclear interactions are computed. As of 2021, gSeaGen has been extended to be applicable to simulating neutrino interactions from energy ranges of MeV to PeV [97], and can simulate neutrinos coming from diffuse, point-like or extended sources. gSeaGen is the standard software package used for simulating events in the

KM3NeT/ORCA detector. All neutrino simulations mentioned in later chapters of this thesis make use of gSeaGen.

GENIE relies on parameterised models for the different scattering processes and energy regimes to describe neutrino-nucleon scattering. Another program, KM3BUU [98], acts as a wrapper around the GiBUU neutrino simulation tool [99], which splits the neutrino interactions into electro-weak scatterings and the subsequent propagation of generated particles through the nucleus. GiBUU provides an alternative means of simulating neutrino-nucleon interactions. Relevant for simulations of KM3NeT/ARCA events, Genhen [100] generates neutrino interaction vertices anywhere inside the interaction volume. It is a FORTRAN-based neutrino simulation code which was used by the ANTARES collaboration and also for KM3NeT/ARCA neutrino event simulations, before the high-energy extension of gSeaGen.

These programs can use external propagator software to propagate muons from charged-current muon or tau lepton neutrino interactions to the can, if the interactions occur outside of the can but the muons can still travel to it. For muon propagation, MUSIC [101] or PROPOSAL [102] can be used. Tau leptons are propagated under the assumption they are minimally ionising particles, and their decays are simulated with the TAUOLA [103] software.

3.4. Cosmic Ray & Atmospheric Muon Simulation

For the simulation of atmospheric muons, two main approaches are used. One may arduously generate primary cosmic ray interactions and track the resulting particles from the interactions. Such an approach requires models to describe the hadronic, electromagnetic, and low-energy interactions of the particles forming air showers (see Section 1.1.2). One such program to simulate cosmic rays and their resulting air showers in detail is CORSIKA [104]. Particles ranging from protons to heavier nuclei such as iron are treated as primary cosmic rays. The particles produced in their interactions are tracked through the atmosphere, including atmospheric muons. These atmospheric muons are simulated to sea level, and then propagated using external software to the can. CORSIKA encapsulates our physical understanding of cosmic ray air showers in the choice of hadronic and electromagnetic interaction models it uses. It provides a means of studying the cosmic ray or atmospheric muon flux and the accuracy of the hadronic or electromagnetic interaction models used. Such an approach requires a lot of computational resources.

Alternatively, one may parameterise the atmospheric muon flux at the depths of the KM3NeT detectors. Such an approach is a far quicker means of simulating atmospheric muons, by simulating events using Monte Carlo methods from parametric equations which describe the atmospheric muon flux at a certain depth. A disadvantage of this method is the parameterised model is tied to a specific flux and to specific interaction models. One such program to simulate atmospheric muons in this way is MUPAGE [105, 106]. The following sub-section is dedicated to MUPAGE, since it forms the basis of Chapter 4.

3.4.1. Parametric Muon Flux Simulation with MUPAGE

MUPAGE is a fast atmospheric muon bundle generator (a group of muons 'bunched' together forms a bundle), which generates muons on the surface of the can according to a specific parameterisation. It may be used for underwater or in-ice neutrino detectors. Parametric equations within MUPAGE describe distributions of the atmospheric muons, as a function of depth, in terms of the **flux** (number of particles arriving per unit area per solid angle per second), **lateral spread** (radial distance from the bundle axis), **single muon energy spectrum** and **multiple muon energy spectrum** (the energy for muon bundles is treated separately from that of single muons, which dominate the flux). MUPAGE is valid from 1.5 — 5.0 km of vertical depth in water, and for a zenith angle from 0° (completely down-going atmospheric muons) to 85° (where 90° implies muons coming directly from the horizon). It incorporates the **multiplicity** of the bundles, i.e. the number of muons in each bundle. The parameters which form the parametric equations of MUPAGE are tuned to a cosmic ray air shower simulation with resulting atmospheric muons, produced with the HEMAS [107] (Hadronic, Electromagnetic and Muonic components in Air Showers) simulation software, which is similar to CORSIKA.

One can vary the values of the parameters within MUPAGE in order to acquire a better description of the data from the KM3NeT experiment. Such a procedure is described in the following chapter. There are multiple reasons for which MUPAGE does not completely describe the KM3NeT data. The HEMAS air shower simulation that MUPAGE is tuned to originates from 1990. The hadronic and electromagnetic interaction models used for cosmic ray air shower simulation have advanced since. Comparisons of the muon multiplicity simulated by HEMAS and CORSIKA have been shown to agree at a level of 10% [106]. In turn, HEMAS and the parametric distributions of MUPAGE are compared and cross-checked to data from the MACRO experiment [108]. The MACRO collaboration used a transition radiation detector to measure the energy of single and

multiple muons, reaching the detector at different depths below the surface in rock. The parameterisations within MUPAGE have been shown to agree to the MACRO data at a level of 5% [106]. The data set from the MACRO experiment is from 2002, and the default MUPAGE parameterisations will not necessarily describe atmospheric muons measured by KM3NeT, measured at different depths and in a different medium.

The 'tuning' of MUPAGE parameters to KM3NeT data is described in Chapter 4. The relevant parameters for the tuning procedure are highlighted below. For a detailed description of the 40(!) free parameters within MUPAGE, please consult Becherini et al. (2006) [106].

Flux Parameterisation

The parametric formula to describe the flux of muon bundles with multiplicity m, at a vertical depth h (in kilometres water equivalent = 10^6 kg m^{-2}), and given zenith angle θ (in radians) is given by

$$\phi(m;h,\theta) = \frac{K(h,\theta)}{m^{\nu(h,\theta)}}\,. \tag{3.9}$$

ϕ is given in units of m^{-2} s^{-1} sr^{-1}.

For single muons,

$$\phi(m=1;h,\theta) = K(h,\theta) = K_0(h)\cos\theta \cdot e^{K_1(h)\cdot\sec\theta}\,, \tag{3.10}$$

with

$$K_0(h) = K_{0a}\cdot h^{K_{0b}}\,, \tag{3.11}$$

$$K_1(h) = K_{1a}\cdot h + K_{1b}\,. \tag{3.12}$$

The fraction of the multiple muon flux with respect to the single flux is given by

$$\nu(h,\theta) = \nu_0(h)\cdot e^{\nu_1(h)\cdot\sec\theta}\,, \tag{3.13}$$

with

$$\nu_0(h) = \nu_{0a} \cdot h^2 + \nu_{0b} \cdot h + \nu_{0c}\,, \tag{3.14}$$

$$\nu_1(h) = \nu_{1a} \cdot e^{\nu_{1b} \cdot h}\,. \tag{3.15}$$

The other parametric equations in MUPAGE follow the same form as the ones presented here. The parameters K_{0a}, K_{1a} and ν_{1b} are used to tune the parameterised atmospheric muon flux, as described in Chapter 4.

Lateral Spread Parameterisation

In cosmic ray air shower interactions, secondary particles are produced in clusters, with most of the energy concentrated in the very forward region, coinciding with the shower axis. As a consequence, in muon bundles, the most energetic muons are expected to arrive closer to the axis of the shower, following the energy distribution of the parent particles. Thus, one can parameterise the energy of muons in a bundle using their radial distance to the shower axis R (the lateral spread).

A variable M which relates to the multiplicity m is used. It is defined as

$$\begin{aligned} M &= m, \quad \text{if } m \leq 3 \\ M &= 4, \quad \text{if } m \geq 4 \end{aligned}\,. \tag{3.16}$$

The average radial distance $\langle R \rangle$ (in metres) is parameterised within MUPAGE as a function of h (depth), M, and θ (zenith angle), as

$$\langle R \rangle = \rho(h, \theta, M) = \rho_0(M) \cdot h^{\rho_1} \cdot F(\theta)\,. \tag{3.17}$$

$\langle R \rangle$ decreases when h increases, as only the most energetic, penetrating muons reach greater depths underwater or in ice, which are closer to the cosmic ray air shower axis. For a given depth, $\langle R \rangle$ decreases with increasing multiplicity, again relating to the more energetic muons originating from high-energy air shower particles, focused in the forward direction. In addition, $\langle R \rangle$ does not depend on the zenith angle θ below $\sim 50°$, then it decreases with increasing θ [106].

The parameter ρ_{0b} is used for the tuning procedure, and is defined within the equation

$$\rho_0(M) = \rho_{0a} \cdot M + \rho_{0b} \,. \tag{3.18}$$

Energy Spectrum - single muons

The energy spectrum of single muons is treated separately from that of multiple muon bundles. The energy loss processes of single muons along the shower axis with energy E_μ, per depth X, are described using Equation (2.6), with the parameters α for ionization energy losses and β for radiative energy loss. The parameter β is used for the MUPAGE tuning procedure.

The expected energy distribution for the number of single muons N of energy E_μ, at a depth X, is written as [106]

$$\frac{dN}{d(\log_{10} E_\mu)} = G \cdot E_\mu e^{\beta X(1-\gamma)} [E_\mu + \epsilon(1 - e^{-\beta X})]^{-\gamma} \,, \tag{3.19}$$

where γ is the spectral index of the primary, parent particles and $\epsilon = \alpha/\beta$. The normalization factor G is defined as

$$G(\gamma, \beta, \epsilon) = 2.30 \cdot (\gamma - 1) \cdot \epsilon^{\gamma-1} \cdot e^{(\gamma-1)\cdot\beta\cdot X} \cdot (1 - e^{-\beta\cdot X})^{(\gamma-1)} \,. \tag{3.20}$$

Here ϵ and the muon energy are in units of TeV, and γ is dimensionless.

Within MUPAGE, γ is parameterised as

$$\gamma = \gamma(h) = \gamma_0 \cdot \ln(h) + \gamma_1 \,, \tag{3.21}$$

with γ_1 used for the tuning procedure also.

Energy Spectrum - multiple muons

For the case of multiple muons, an equation similar to the single muon case, Equation (3.19), is used. In this case however, γ^*, β^* and ϵ^* are used to distinguish from the single muon case, and depend on the multiplicity M and the radial distance to the

shower axis R, as well as on the vertical depth h and the zenith angle θ [106]. In brief:

$$\frac{dN}{d(\log_{10} E_\mu)} = G \cdot E_\mu e^{\beta^* X(1-\gamma^*)} [E_\mu + \epsilon^*(1 - e^{-\beta^* X})]^{-\gamma^*} . \tag{3.22}$$

The parameters b_{1b} and d_{0b}, relating to γ^* and ϵ^* respectively, are also tuned in Chapter 4.

Parameter Values

The **default** values of the parameters relevant for the tuning procedure of Chapter 4 are presented in Table 3.1, in the syntax used by MUPAGE. For the full list of MUPAGE parameter values, please refer to Table C.1.

Parameter	Value
K0a	0.0072
K1a	-0.581
K1b	0.034
ni1a	0.03
ni1b	0.470
BETA	0.420
g1	3.961
RHO0b	28.26
B1b	6.124
D0b	3.955

Table 3.1.: Default values of the MUPAGE parameters, used to tune to KM3NeT data. The syntax of the parameter names within the program is used here. From Carminati, Margiotta and Spurio (2008) [105].

3.5. Cherenkov Light Emission & Detection

Following event generation and particle production, the Cherenkov light emission by these particles, its propagation through seawater, and its detection by the PMTs is simulated, resulting in the detected number of photo-electrons per PMT.

There are two main methods used to carry out this function. The cumulative number of photo-electrons recorded by a PMT at a given time can be estimated from the data

tables described in Section 3.2, for either track or shower event parameters with respect to the PMT. Thus for these events types, and including direct and single-scattered light, the number of photo-electrons recorded by each PMT is evaluated. This is carried out by a program called `JSirene`, an application within the `Jpp` framework.

The advantage of this method to evaluate the number of photo-electrons is that the number of photo-electrons can be readily estimated from pre-existing tables. A more detailed - but slower - procedure tracking the individual particles and photons and evaluating the number of photo-electrons induced by each one can be carried out instead. This offers a more complete description of each individual photon and the resulting number of photo-electrons on each PMT, in particular at the few-GeV energies relevant for KM3NeT/ORCA where there are fewer photons and multiple interaction processes at play. One such program to do this is KM3SIM [109], a program which propagates all particles separately - including photons - using the GEANT4 [110] particle simulation package. In both of these methods, the detector response is included in the form of the PMT quantum efficiencies.

3.6. Triggering and Detector Simulation

As outlined in Section 2.5.1, trigger algorithms are used to extract physics events from data, which is dominated by background signals. Tracks, particle showers or mixed track and shower event signatures (mentioned in Section 2.3) are identified. A program called `JTriggerEfficiency` within the `Jpp` framework performs this task of applying the trigger algorithms and filtering the simulated data (into L0, L1, L2 hits), in the same way that is done for actual data. The events that pass the trigger conditions are stored as triggered events, again in the same way that is done for the data. Hits due to ^{40}K decays are also added to the simulation. In this simulation stage the detector conditions following from calibration procedures can be included, in order to properly simulate the timing and position of the hits in the detector. The detector response is incorporated in this stage of the chain as the PMT rates and PMT quantum efficiency values.

Often, simulating a particular data-taking run is required, when one wishes to simulate the exact detector response for a given period. This includes the specific detector conditions such as the PMT quantum efficiencies, the levels of high-rate veto, and the measured background rates. The specific trigger algorithms implemented during a run are also used in this case. This is referred to as 'run-by-run' mode, where the trigger algorithm parameters, background rates and statuses of the PMTs are read from the data

and used to simulate the same data set in the given run. Such a procedure is used in Chapter 4 to simulate the atmospheric muons detected for a specific detector run.

3.7. Reconstruction

In order for KM3NeT/ARCA to achieve its goals of identifying and discovering the sources of astrophysical neutrinos (and cosmic rays), the direction and energy of cosmic neutrinos needs to be determined, to associate neutrinos with a likely source. For KM3NeT/ORCA, the direction and energy needs to be determined for many events to provide sensitivity to oscillation parameters and the neutrino mass ordering (see Chapter 1). Thus, it is imperative to reconstruct properties of the neutrino such as its direction and energy.

Since we do not observe the neutrino directly, our best handle on the neutrino direction and energy is in reconstructing the direction and energy of the products of its interaction, which usually take the form of the event signatures described in Section 2.3. We reconstruct the direction and energy of tracks, electromagnetic showers and/or hadronic showers associated with neutrino interactions to estimate these same properties for the neutrino. As described in Chapter 2, the KM3NeT experiment provides the arrival time of photon-induced hits in the detector. These hits have associated positions, according to the positions of the optical modules within the detectors. Thus we know the time, position and amount of charge associated with hits, and from this information need to determine the event signature type, and its direction and energy (as a proxy for the neutrino).

To reconstruct an event requires fitting a model (in fact, model parameters) to data. A model describing a muon track or a particle shower can be compared to data to estimate how well the model describes (or 'fits') the data. Model parameters can include the direction, energy, vertex position and time of the event to be reconstructed. Means of parameter estimation (i.e. model fitting techniques) are outlined in this section, followed by a description of the methods used to reconstruct track-like or particle shower-like events in the KM3NeT detectors.

3.7.1. Parameter Estimation

Parameter estimation is a procedure used to obtain an estimate on the value of the model parameter θ, with an associated error z in the form $\theta \pm z$. Parameter estimation is described in detail in e.g. Metzger (2002) [92]. Methods of parameter estimation are used in reconstructing the model parameters of track-like and shower-like events.

Maximum Likelihood Estimate

The *likelihood function* $\mathcal{L}(x;\theta)$ describes the probability to obtain the outcome x, given the parameter (or parameter set) θ. For a set of n independent measurements x (with the individual measurement x_i, i=0,1...n) sampled from a probability density function $f(x;\theta)$, the likelihood function is given by

$$\mathcal{L}(x;\theta) \equiv \prod_{i=0}^{n} f(x_i;\theta)\,. \tag{3.23}$$

The value of θ which best describes the model f will be that which results in the maximum value of the likelihood. This best estimate of θ implies an unbiased one, with minimum variance [92].

The likelihood function can result in extremely small or extremely large values. To make life computationally easier, the minimum value of the negative log-likelihood ℓ can be found instead (which also results in a sum rather than a product of values). It is defined as

$$\ell \equiv -\ln \mathcal{L}(x;\theta) \equiv -\sum_{i=1}^{n} \ln f(x_i;\theta)\,. \tag{3.24}$$

A common issue in maximum likelihood estimation for non-linear systems is the *starting value problem* - where one begins the fitting procedure in order to determine the *global* maximum of the likelihood in the parameter space, where the best estimate of the model parameters lie. Depending on the shape of the negative log-likelihood in the parameter space, local minima may exist which should be avoided.

Least Squares Fitting

Consider N linearly-related data points x_i, y_i, with uncertainty on y_i given by σ_i, and a functional relationship of the form $y = f(x, a, b)$ where a and b are parameters to be determined. The values of a and b can be found by minimising the following chi-squared function:

$$\chi^2 \equiv \sum_{i=1}^{N} \frac{[y_i - f(x_i, a, b)]^2}{\sigma_i^2}. \tag{3.25}$$

This procedure of minimising the difference between data points and their expected value is called 'least squares' fitting. When the model is nonlinear, the same equation can be used. In that case, if the deviations $y_i - f(x_i, a, b)$ are normally distributed, the best estimate from this method is the same as from maximum likelihood estimation [92]. If the deviations are not normally distributed, the method of least squares fitting is less desirable, and a minimisation of the log-likelihood is preferred in giving a best estimate of the model parameters.

M-Estimators

Least squares fitting can be prone to influence by outliers, i.e. data points several standard deviations from the expected mean. Functions referred to as M-estimators can be defined which negate the influence of outliers. These are a broad category of functions which aim to minimise a function of the form

$$M \equiv \sum_{i=1}^{N} \rho(x_i, \theta). \tag{3.26}$$

These functions have the form $\rho(z)$ with a minimum at $z = 0$, and flatten at higher values of z [69].

3.7.2. Hit Probabilities

The maximum likelihood principle maximises the product of probability density functions which describe expected measurements, given model parameters. The probability of obtaining hits (the measurements) in the KM3NeT detectors can be defined. These can in turn be used to define likelihood functions to be maximised. The tables of the number of photo-electrons as a function of the arrival time of light for muon or particle

shower events, described in Section 3.2, readily provide values for these functions, for specific values of track or shower model parameters. In this section some definitions of hit probabilities are provided.

The Poisson probability for the number of occurrences k of a discrete random variable X , with parameter λ the expected value of X, is given by

$$P_{\text{Poisson}}(k|\lambda) \equiv \frac{\lambda^k \cdot e^{-\lambda}}{k!} . \tag{3.27}$$

The probability of detecting the number of photo-electrons $n_{p.e.}$ on a PMT, given an expected number of photo-electrons $\hat{n}$, can be found following Poissonian statistics. The probability to detect at least one photo-electron on PMT i is

$$P_i(n_{p.e.} \geq 1; \hat{n}_i) = 1 - P_i(n_{p.e.} = 0; \hat{n}) = 1 - \frac{\hat{n_i}^0 \cdot e^{-\hat{n}_i}}{0!} = 1 - e^{-\hat{n}_i} . \tag{3.28}$$

Defining $n(t)$ as the number of photo-electrons detected by PMT i at a given time t, the probability of observing a hit in the time window $[t_1, t_2]$ is

$$P_i^{\text{hit}} = 1 - e^{-\int_{t_1}^{t_2} n(t) \cdot dt} . \tag{3.29}$$

This can referred to as the '**hit PMT**' probability.

The '**non-hit PMT**' probability is the chance of observing no hits on the PMT in the time interval $[t_1, t_2]$, given by

$$P_i^{\text{non-hit}} = e^{-\int_{t_1}^{t_2} n(t) \cdot dt} . \tag{3.30}$$

Another useful quantity is the probability density function for obtaining the first hit on a PMT in some time interval between t_1 and t_2, given by the '**first hit probability density**' $P_{\text{first hit}}$:

$$P_i^{\text{first hit}} = n(t) \cdot \frac{e^{-N(t)}}{1 - e^{-N_{\text{total}}}} . \tag{3.31}$$

$N(t)(= \int_{t_1}^{t} n(t) \cdot dt)$ is the cumulative number of photo-electrons up to that given point in time, and N_{total} is the total number of photo-electrons in the considered time window. The term $1 - e^{-N_{\text{total}}}$ normalises the above expression to unity.

The values of n, $N(t)$ and N_{total} can be obtained from the photo-electron distribution tables described in Section 3.2. The hit probabilities defined in this section can be used to define likelihood functions for detecting hits (or no hits) on a PMT. Although not explicitly stated, the probability definitions above reflect the probability of obtaining hits or no hits on a PMT for the specific track or shower model parameters under consideration.

The time residuals used to obtain values from the photo-electron distribution tables are considered in the reconstruction software within specified **time windows** around the value of the expected arrival time t_0 of $[t_0 + t_{min}, t_0 - t_{max}]$. t_0 follows the definition for that of a track or shower, given in Section 3.2. The values of t_{min}, t_{max} are usually chosen in the range of approximately [-100, 500] ns respectively, reflecting the distributions shown in Figure 3.2.

3.7.3. Muon Track Reconstruction

The parameters to be reconstructed for the hypothesis of a muon track in the KM3NeT detectors are the direction of the muon, its energy, and the time and position at which the first detected light is emitted (the 'start position'). The start position is often not the same thing as the interaction vertex, since this can lie outside the detector. The 'stop position' of the track is in turn the last point at which detected light is emitted. The distance between the start and stop point of the track defines the track length. For a minimum ionising particle, as per Equation (2.6), the track length is proportional to the energy and is a useful estimate for it.

Fitting muon parameters is no exception to the starting value problem in maximum likelihood estimation. How do we begin to fit all of these parameters? As a solution, an initial fit (a '**prefit**') is carried out to determine the direction of the muon, where a scan over starting point directions is made, resulting in a set of prefit candidates which are used for more accurate direction fit procedures. Doing this reduces the complexity of the problem. In turn, it is easier to reconstruct the start position and the energy of the muon track separately (rather than direction, energy and start and end points all at once).

Reconstruction Chain Ordering

Multiple algorithms are used in sequence to reconstruct the aforementioned parameters. This sequence of algorithms is referred to as a '**reconstruction chain**'. The algorithm names and their functions are summarised in Table 3.2.

Function	Algorithm name
Prefit direction estimate	`JMuonPrefit`
Intermediate direction fit	`JMuonSimplex`
Full direction fit	`JMuonGandalf`
Track length estimate	`JMuonStart`
Energy estimate	`JMuonEnergy`

Table 3.2.: Stages of the track reconstruction chain and the corresponding program name.

These algorithms and their usage are subject to development. At the time of writing, two different sequences of these reconstruction algorithms are used for track reconstruction in KM3NeT/ARCA and KM3NeT/ORCA. The different orderings are a result of investigations into the speed and performance of these algorithms when applied to the different detectors.

For KM3NeT/ARCA, the track reconstruction chain is as follows:

`JMuonPrefit` → `JMuonSimplex` → `JMuonGandalf` → `JMuonStart` → `JMuonEnergy`.

For KM3NeT/ORCA, the track reconstruction chain is as follows:

`JMuonPrefit` → `JMuonSimplex` → `JMuonStart` → `JMuonGandalf` → `JMuonStart` → `JMuonEnergy`.

For KM3NeT/ORCA the track length is estimated twice, since it is a useful constraint in determining the direction of the muon track. This track length inclusion in the direction fit is part of the work for this thesis and is described in Chapter 5. After the direction is determined, the track length is estimated again for further accuracy. The fitting procedures in this reconstruction chain are described in further detail below. Throughout the chain, the coordinate system used is the same as that in which the photo-electron distribution tables are evaluated, with the direction of the track along the z-direction.

Direction Prefit

An initial fit is made over more than a thousand directions, isotropically distributed over the sky, with a chosen grid angle between directions. The position and time of the muon track are estimated for each individually assumed direction using an iterative procedure to minimise a χ^2 function, as a function of the time difference between the arrival time and the expected arrival time, assuming a Cherenkov cone hypothesis. At the minimum of this function, the model parameters best describe the data.

For each direction, the coordinate system of Figure 3.1 is used to define the expected arrival time of photons on the PMT. This arrival time t_j is

$$t_j \equiv t_0 + \frac{z_j}{c} - \left(\frac{R \cdot \tan\theta_C}{c}\right), \tag{3.32}$$

as derived in Equation (A.12), with z_j the distance to the $z = 0$ plane, c the speed of light in a vacuum, and θ_C the Cherenkov angle. The distance of closest approach between the muon trajectory and the PMT is given by R, defined as

$$R_j \equiv \sqrt{(x_j - x_0)^2 + (y_j - y_0)^2}, \tag{3.33}$$

where (x_0, y_0) is the position of the muon travelling in the z-direction when it crosses the $z = 0$ plane.

The time difference between all pairs of consecutive hits i, j can be described using the relation

$$\begin{aligned} t_j'^2 - t_i'^2 - 2(t_j' - t_i')t_0' = x_j^2 - x_i^2 - 2(x_j - x_i)x_0 \\ + y_j^2 - y_i^2 - 2(y_j - y_i)y_0, \end{aligned} \tag{3.34}$$

as derived in Appendix B.

This equation can be solved using a least squares minimisation. A cluster of causally related hits is selected from the data in order to limit effects induced by background signal. The possible outliers are removed if their contribution to the total χ^2 is larger than 3 standard deviations. In this particular algorithm the best fit is determined by a **quality parameter**, according to which the fits are sorted, from best to worst. This quality parameter Q is defined as

$$Q \equiv \mathrm{NDF} - \frac{1}{4}\frac{\chi^2}{\mathrm{NDF}}, \tag{3.35}$$

where NDF is the number of degrees of freedom in the fit.

Dividing by the number of degrees of freedom is carried out in order to be able to compare the fit directions with small associated hit statistics and fit directions with a large number of hits. The division by 4 is chosen to provide sufficient separation between the qualities of the individual fits [111].

Intermediate Direction Fit

The program `JSimplex` performs an intermediate fit of the muon trajectory to data. All hits within a specified distance R and time window are used to compare the arrival time with the expected arrival time due to Cherenkov light from a muon track. In the case of multiple hits from the same PMT, the earliest hit is used. An M-estimator as a function of the time residuals is minimised. This step has been found to improve the speed in the following direction fit for events in the KM3NeT/ORCA detector.

Advanced Direction Fit

Here, the final value of the position and direction of the track is estimated. Taking as input a chosen number of fits from the previous algorithm, a scan is performed around the corresponding track hypothesis and the maximum likelihood method is used to evaluate the track direction and position. The photo-electron distribution tables for muon light are used in this fit procedure.

The expected arrival time of Cherenkov light from a muon, and the resulting time residual $\Delta t = t_{hit} - t^{\text{expected}}$, is estimated for each hit. All the first hits in the detector are considered within a cylinder of radius R, and within a time window of [-50,450] ns. The track length, if already determined, is used to define the length of the cylinder. If not, the track and cylinder in which hits are considered is treated as infinitely long.

The probability that a hit is only caused by background signal is estimated using Equation (3.31) for an estimated time residual Δt, using a constant background rate r (which can be given as input to the program, or read from the input data). The likelihood that a hit was caused by background only is

$$\mathcal{L}_{\text{b}} = -\ln\left[P^{\text{background}}_{\text{first hit}}(\Delta t, r)\right]. \tag{3.36}$$

Similarly, the probability that a hit resulted from a muon-like signal is found using Equation (3.31), where the probability values are obtained from the muon photo-electron distribution tables. The likelihood of obtaining a first hit on a PMT, including background, given the track parameters $(E, R, \theta, \phi, \Delta t)$ is

$$\mathcal{L}_{\text{s+b}} = -\ln\left[P^{\text{signal}}_{\text{first hit}}(E, R, \theta, \phi, \Delta t) + P^{\text{background}}_{\text{first hit}}(\Delta t, r)\right], \tag{3.37}$$

where 's+b' stands for signal plus background. The function to be minimised takes the form

$$\lambda \equiv \sum_{\text{first hits}} \ln\left(\frac{\mathcal{L}_{\text{b}}}{\mathcal{L}_{\text{s+b}}}\right). \tag{3.38}$$

This is the same as the maximum likelihood estimation procedure, with an offset in the likelihood value induced by the background likelihood.

The track parameters which result in the minimum value of λ are the most compatible with the signal hypothesis and least compatible with the background-only hypothesis. λ is minimised using the Levenberg-Marquardt method [112, 113]. Following this procedure, the direction, time and position of the track is estimated. The quality parameter $q = -\lambda_{\text{minimum}}$ is used as an indicator of the quality of the fit. The larger the value, the better the fit. This parameter is used in Chapter 4 of this thesis.

Track Length Estimate

The hits within a cylinder of pre-defined radius and within a defined time window are back-projected onto the track under the Cherenkov angle. The hits that exceed a pre-defined signal-over-background threshold along the track are used to identify the start and end point of the track. The vertex position of the track is then moved to the determined start point.

Track Energy Estimate

An iterative search within a pre-defined energy range is carried out. A likelihood function is defined using the hit PMT and non-hit PMT probabilities of Equation (3.29) and Equation (3.30), respectively. The photo-electron distribution tables for muon light are used to provide estimates of the probabilities for the track parameters. The energy

resulting in the maximum of the likelihood within the sampled energy range is taken as an estimae of the muon energy at the track starting point.

3.7.4. KM3NeT/ORCA Shower Reconstruction

A separate series of algorithms are used to reconstruct the energy and direction of particle showers in KM3NeT/ORCA. These make use of the electromagnetic shower photo-electron distribution tables of Section 3.2. Particle showers at the energy range detected by KM3NeT/ORCA tend to be quite localised, and reconstructing their energy and direction can be difficult.

Again, the starting value problem makes this fit procedure difficult. An attempt is first made to estimate the vertex position. The point of shower maximum (defined in Section 2.3) is the brightest point of the shower. It is approximated as the shower position. A hit selection of L2 and L0 hits which satisfy the criteria of Equation (2.8) is carried out, to minimise background. For hits i and j of position and time (x_i, y_i, z_i, t_i) and (x_j, y_j, z_j, t_j), respectively, the vertex (x_0, y_0, z_0, t_0) is estimated from

$$(t'_j - t'_0)^2 = (x_j - x_0)^2 + (y_j - y_0)^2 + (z_j - z_0)^2 \,. \tag{3.39}$$

Here, $t'_j = nct_j$ and $t'_0 = nct_0$. An estimate of the vertex position and time is again carried out using an M-estimator of the time residuals, similar to the intermediate direction fit in the track direction estimation.

As a next step, photo-electron distribution tables pertaining to the isotropic emission of light by a shower - without a direction and assuming an expanding bright point for the shower signature - are used to determine the shower vertex position and time of the event. These tables are pre-defined, in addition to the ones described in Section 3.2. To estimate the vertex position and time, a likelihood function similar to Equation (3.38) is minimised, using only first hit probabilities. For this case of a particle shower, the parameters D (distance from the vertex to the PMT), $\cos\delta$ (the emission angle of the light), Δt (the time residual between the actual and expected arrival time of the light) and E (shower energy) are used. The signal hypothesis likelihood is added to the background hypothesis likelihood:

$$\mathcal{L}_{\text{s+b}} = -\ln\left[P^{\text{signal}}_{\text{first hit}}(D, \cos\delta, \Delta t, E) + P^{\text{background}}_{\text{first hit}}(\Delta t, r)\right] . \tag{3.40}$$

In the next steps of this shower reconstruction chain, the direction and energy are determined. The isotropic photo-electron distribution tables are used for a prefit of the energy. Using the hit PMT and non-hit PMT information (Equation (3.29) and Equation (3.30), respectively) and the isotropic photo-electron distribution tables, the hit PMT and no hit PMT probability are used to determine a likelihood function to be minimised, for a pre-defined energy range.

A prefit of the shower direction is then carried out, using the (directional) electromagnetic shower photo-electron distribution tables of direct and scattered light described in Section 3.2, with all previous estimates of the energy and vertex as starting points. In this step the hit and non-hit PMT probabilities are evaluated. This fit is then carried out again over all parameters to determine the energy and direction of the particle shower using the electromagnetic shower photo-electron distribution tables.

Fitting shower parameters at lower energies is clearly a complex procedure. A more thorough overview of the `Jpp` shower reconstruction algorithms can be found in Domi (2019) [114].

3.7.5. KM3NeT/ARCA Shower Reconstruction

A separate reconstruction algorithm is used to determine the direction and energy of particle showers produced by high-energy neutrino interactions, namely those detected primarily by KM3NeT/ARCA. This algorithm, named `aashowerfit`, uses the hit and non-hit PMT information and assumes an expanding sphere for the shower topology. A description of this algorithm can be found in Melis et al. (2018) [111]. It has recently been updated and expanded to use the first hit information to improve its performance, and also for the reconstruction of tau neutrino event signatures [115].

3.7.6. Quantifying the Performance of Reconstruction Procedures

It is useful to quantify how well parameters of interest have been reconstructed by reconstruction procedures. One way of doing this is recording the errors associated with the fitted parameters, with these errors usually provided by the minimisation algorithms. Another useful metric for how well a quantity has been determined is the **resolution**. The angular resolution and energy resolution are briefly defined in this section, as they will be referred to frequently in later chapters.

Angular Resolution

A valuable metric for quantifying the performance of the direction reconstruction is the **angular resolution**, conveying the accuracy in which the direction of the reconstructed event can be determined. The **angular deviation** can be defined as the angle between the true direction and the reconstructed direction. For two vectors $\vec{A}$ and $\vec{B}$, the angle $\Delta\theta$ between the two vectors can be found using the dot product between the two vectors:

$$\Delta\theta \equiv \arccos\left(\frac{\vec{A}\cdot\vec{B}}{|\vec{A}||\vec{B}|}\right),$$

where $|\vec{A}|$ and $|\vec{B}|$ denote the magnitude of both vectors, respectively.

To illustrate the angular resolution, the angular deviation $\Delta\theta$ can be calculated for every reconstructed event, resulting in a two-dimensional distribution of $\Delta\theta$ as a function of the event energy. Taking quantiles of this distribution, such as the median, can illustrate the angular resolution as a function of energy. Such plots of angular resolution can be found in Chapters 5 and 6. This quantity is expected to improve (become smaller) as the energy of the event in question increases, as more energetic events produce more hits in the KM3NeT detectors, and with more information the ability to determine the direction should improve.

Energy Resolution

There are various means of illustrating the energy resolution resulting from reconstruction efforts. A useful metric for signifying how well the energy has been reconstructed is

$$\log_{10}\left(\frac{E_{\text{reco}}}{E_{\text{true}}}\right),$$

the logarithm to the base 10 of the reconstructed energy divided by the true energy. The closer the value is to zero, the closer the true energy is to the reconstructed energy. This offset $\log_{10} E_{\text{reco}} - \log_{10} E_{\text{true}}$ can be referred to as the **bias**.

The quantity $\log_{10}\left(E_{\text{reco}}/E_{\text{true}}\right)$ can be shown as a function of the energy. One expects the reconstructed energy to become closer to the true energy as the energy increases, as there are more hits deposited in the detector, providing more information for the reconstruction procedure. The width of $\log_{10}\left(E_{\text{reco}}/E_{\text{true}}\right)$ as a function of energy indicates the **energy resolution**, with a more narrow spread around the median value

indicating a smaller error margin. The standard deviation σ can be used to quantify the width. Such illustrations of the energy bias/resolution are shown in Chapters 5 to 7.

Chapter 4.

Tuning the Parametric Atmospheric Muon Flux Simulation to KM3NeT Data

In this chapter, efforts to tune the parametric equations describing the atmospheric muon flux within the MUPAGE program to KM3NeT data are presented. MUPAGE is used to generate atmospheric muons in the KM3NeT detectors by describing the generated flux through parametric equations. An overview of the atmospheric muon simulation with MUPAGE can be found in Section 3.4.1. The atmospheric muons generated by MUPAGE according to its internal parameterisation are propagated through the simulation chain described in Chapter 3, from which we extract simulated *observables* in the detector. These physical observables can be for example the atmospheric muon zenith angle distribution or the reconstructed muon energy distribution. The observables obtained from the simulation are used to tune the MUPAGE parameterisation to KM3NeT data.

This tuning method is proven to be successful, resulting in a better agreement between the distributions of observables in the KM3NeT/ORCA-6 data and those in the simulation, for the considered data-taking period. These tuned parameters are then used to simulate a larger data set over a longer time period, and are again shown to improve the description of the KM3NeT/ORCA-6 data. An outlook is given at the end of this chapter, including possible improvements to the tuning procedure.

4.1. Motivation

To illustrate the need for tuning the MUPAGE parameters to KM3NeT data, Figure 4.1 shows the level of agreement between the un-tuned MUPAGE simulation and data from the KM3NeT/ORCA-6 (sub-)detector for select observables. The **nominal** values of

all MUPAGE parameters are used here, i.e. those listed in Table C.1. The observables shown are the reconstructed cosine of the zenith angle distribution (the angle in which the muons impinge on the detector, from completely down-going ($\cos(\theta_{\text{zenith}}) = -1$) to completely up-going ($\cos(\theta_{\text{zenith}}) = 1$)), the reconstructed energy (the energy of the muons, as evaluated by the track reconstruction software), the number of overlays (the number of simultaneous triggers caused by an event), the number of triggered hits, and the track reconstruction quality parameter. The quality parameter is the value of the likelihood from the `JGandalf` muon direction reconstruction procedure (see Section 3.7.3).

For this tuning of MUPAGE parameters to data, it is beneficial to limit the ranges of the distributions such that one is not tuning on background noise or mis-reconstructed events. Therefore, the range of the reconstructed zenith angle is limited to $\cos(\theta_{\text{zenith}}) = [-1, 0.5]$. The muons reconstructed with an upward-going direction are mis-reconstructed. In addition, the reconstructed energy range is limited to $[5, 10^6]$ GeV. Take note of the 'double bump' feature in the reconstructed energy distribution. This is thought to come from the small detector size. With a reduced distance in the horizontal direction, there is a limit to the maximum reconstructed track length of the muon. Thus, muons with an energy reconstructed proportionally to the track length (as a minimum ionising particle) have a maximum bound. The left bump is thought to come from such muons, creating an artificial two bump feature. This feature is seen in other analyses of KM3NeT/ORCA-6 data. With regards to the number of overlays, the pattern shown here is thought to be a result of many events inducing simultaneous triggers that are multiples of 3. With 3 trigger algorithms (see Section 2.5.1), events appear more likely to be triggered upon by multiples of all three algorithms.

These observables are shown for a particular data-taking run in Figure 4.1, as an example to motivate the need for tuning of the MUPAGE parameters to data in order to improve the agreement between the data and simulation. For example, the simulation and data for $\cos(\theta_{\text{zenith}}) > -0.6$ are not in very good agreement, and the same holds for atmospheric muons with reconstructed energy $> 10^3$ GeV. There is also a clear difference between the simulation and data for the reconstructed quality parameter. For the atmospheric muons which are completely down-going, the data value differs from the simulation by $\sim 8.5\%$.

The parameters within MUPAGE are ripe for tuning, having never been tuned to KM3NeT data before. And so, in this chapter, the approach to tuning the parameters within MUPAGE to improve the description of observables in KM3NeT data is described. The strategy in choosing parameters to tune is outlined in the following section. These parameters are selected and tuned on KM3NeT/ORCA data. The choice to tune on

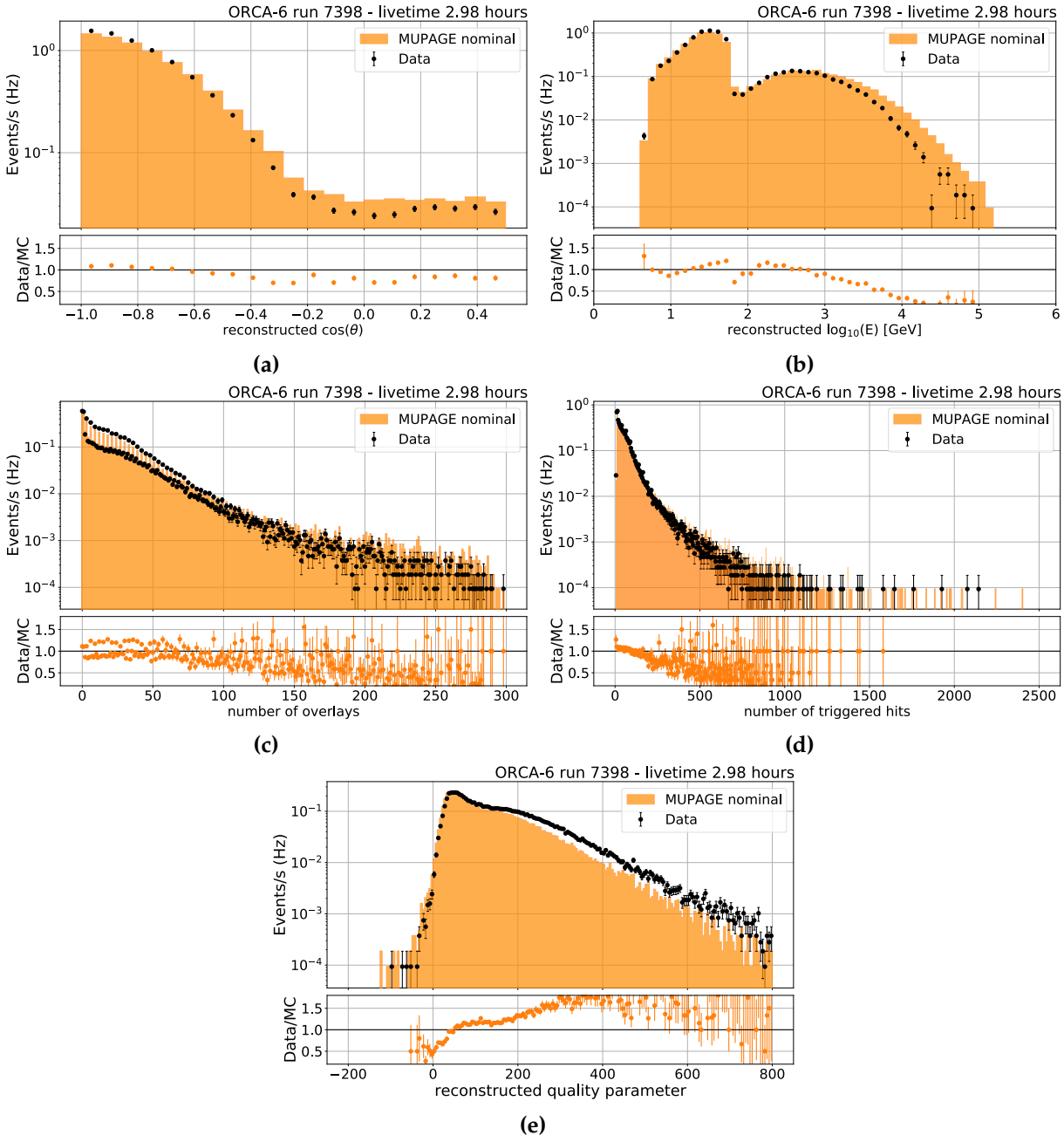


Figure 4.1.: The distributions of the (a) cosine of the zenith angle, (b) reconstructed energy, (c) number of overlays, (d) number of triggered hits and (e) quality parameter (i.e. the reconstructed likelihood), comparing the data from a select ORCA-6 data run and the simulation, for the nominal MUPAGE parameter values. The ratio of the data to the MC is also shown. These distributions are the focus of the MUPAGE tuning exercise of this chapter.

KM3NeT/ORCA-6 data was made due to the availability of high quality, well-calibrated data at the time. The level of agreement between the data and MUPAGE simulation before and after tuning is quantified using a chosen metric.

4.2. Methodology

From the distributions shown in Figure 4.1 above, it is apparent that both the shape and scaling can be adjusted to bring about a better agreement between the data and simulation. Of the 40 free parameters within MUPAGE (see Table C.1), a select few are identified which have a sizeable impact on the shape or scale of the distribution of some observables. This is achieved by simulating the observables for the set of default MUPAGE parameter values, and then altering the values of one parameter at a time.

The simulation chain of

$$\texttt{MUPAGE} \rightarrow \texttt{JSirene} \rightarrow \texttt{JTriggerEfficiencyRunByRun} \rightarrow \text{Muon Track Reconstruction Chain}$$

is chosen to simulate reconstructed observables in KM3NeT data. All of these simulation programs are described in Chapter 3.

4.2.1. Identifying Impactful MUPAGE Parameters

This procedure of varying MUPAGE parameters compared to the nominal value is carried out for simulated KM3NeT/ORCA-4 data, which was available at the time, in order to identify the parameters with the most impact on the observables. The ORCA-4 detector operated from July 2019 to January 2020, before the addition of two more detection units to make ORCA-6. 10,000 atmospheric muons are simulated for each MUPAGE parameter at its nominal (un-tuned) value. The same simulation is carried out using 3 times the default value and 1/3 of the default value (to ensure a wide range of values is covered). The number of muons being 10,000 is chosen for a statistically significant effect, to ensure that changes in the resulting distributions are not due to statistical fluctuations but due to the new parameter values. Altering the parameter values affects the processing time, since the generated flux is altered and thus the lifetime (the time taken to generate the required number of particles) of the atmospheric muon simulation changes. For the cases where the processing time was enormously affected, 2 times and 0.5 times the nominal parameter value are considered instead, to save processing time.

The parameters which are found to have the most noticeable impact on distributions of observables are K_{0a}, K_{1a}, ν_{1b}, ρ_{0b}, β, b_{1b} and d_{0b}. The effect of varying the parameter ν_{1b} on the simulated observables is shown in Figure 4.2 as an example. The results for the other parameters are presented in Appendix D. Varying K_{0a} simply scales the

distributions with no effect on their shapes. The parameter ν_{1b} affects the shape of the reconstructed zenith angle, reconstructed energy and number of overlays in different ways. The range of zenith angles in these figures is from 0° – 180°, with *in this case* 180° meaning completely down-going atmospheric muons. The effect on the multiplicity, the number of muons in a muon bundle, is also shown. The number of overlays acts as a proxy for the multiplicity, which is available in the simulation but not (directly) measurable in the data. The number of overlays relating to the multiplicity can be understood by individual clusters of causally-related hits in the detector being caused by separate muons in the bundle (depending on the lateral spread of muons in the bundle).

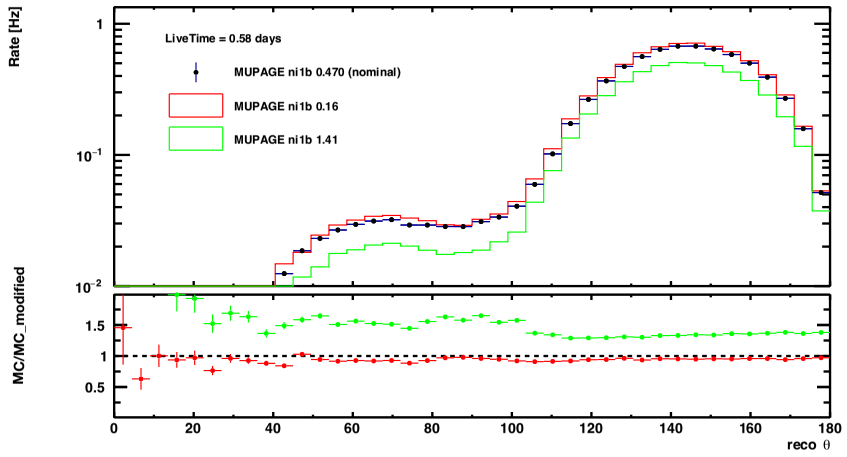
The distributions of the reconstructed zenith angle, reconstructed energy, number of overlays and multiplicity are shown for the nominal parameter value in MUPAGE, as well as for the samples in which the parameter value has been varied. These distributions are presented as the rate of events (number of events per second) as a function of a range of parameter values. Underneath each distribution is the ratio of the default value distribution to the altered distribution. The statistical error bars on the nominal value points are included.

The parameters K_{0a}, K_{1a}, and ν_{1b} pertain to the parameters used to describe the atmospheric muon flux, the parameter ρ_{0b} relates to the lateral spread of muons in a bundle, while β is used to parameterise the single muon energy spectrum. The parameters b_{1b} and d_{0b} parameterise the multiple muon energy spectrum. Please refer to Section 3.4.1 for a reminder of the parametric equations in MUPAGE. As can be seen, these MUPAGE parameters affect all of the observables in non-trivial ways. There are degeneracies in how these parameters affect the same observable, for example between ρ_{0b} and d_{0b}. This already makes the exercise of tuning parameters to KM3NeT data non-trivial.

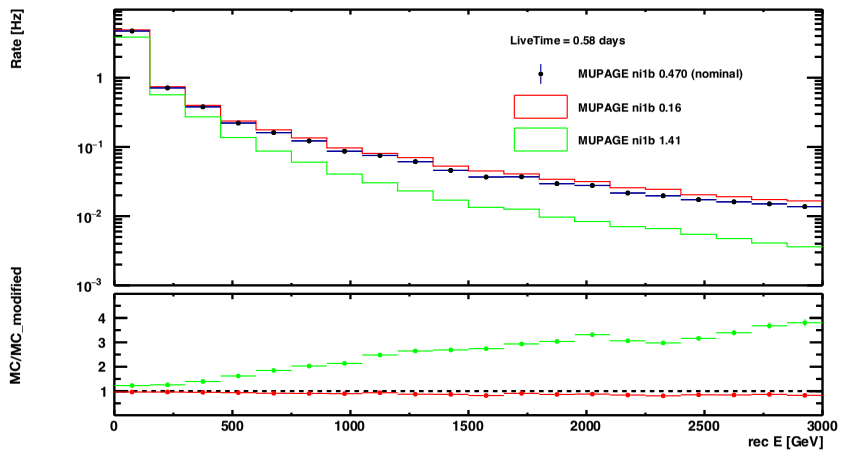
4.2.2. The Significance Test

For the process of varying MUPAGE parameters, propagating the events through the simulation chain to the reconstruction stage, and then comparing the simulated observables to sea data, a metric is needed to quantify the level of agreement between the data and the simulation.

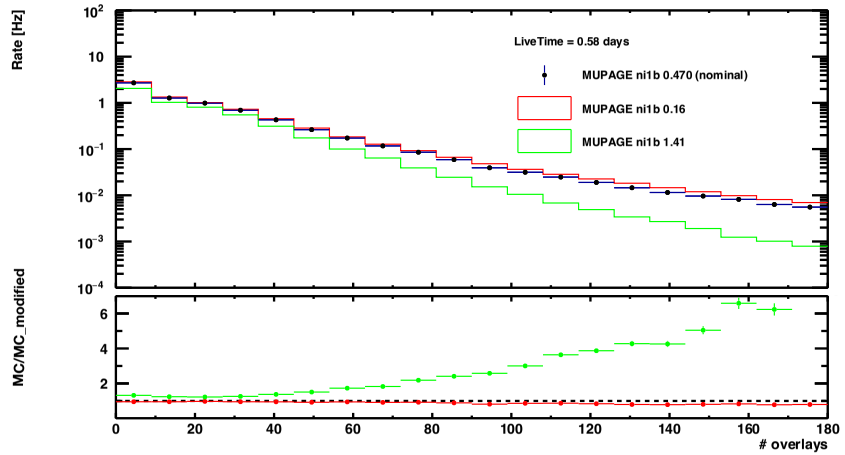
A metric referred to as the '**significance**' is defined. The significance S for two histograms (graphical representations of distributions) H_a and H_b can be used to quantify



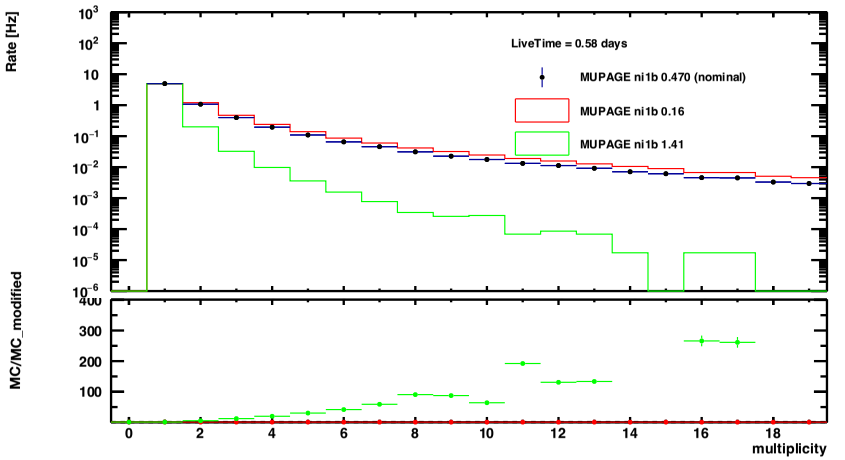
(a) ν_{1b} - reconstructed zenith angle



(b) ν_{1b} - reconstructed energy



(c) ν_{1b} - overlays



(d) ν_{1b} - multiplicity

Figure 4.2.: The effect of varying ν_{1b} on the distributions of the (a) reconstructed zenith angle, (b) reconstructed energy, (c) number of overlays, and (d) multiplicity. Increasing this parameter value has a greater effect on the distributions than decreasing it. A large difference is induced for higher values of the reconstructed energy and the number of overlays. As shown in (d), increasing the value of ν_{1b} reduces the multiplicity significantly.

the level of agreement between them. S may be defined as

$$S = \frac{1}{N} \cdot \sum_{i=0}^{N} \frac{|a_i - K \cdot b_i|}{\sqrt{\sigma^2(a_i) + K^2 \cdot \sigma^2(b_i)}}, \tag{4.1}$$

where N is the number of bins in each histogram. A bin refers to the interval chosen for a histogram into which the range of values are divided; the number of entries falling into that interval is then displayed. S is calculated for each bin i in the two histograms. K represents the normalisation factor or scaling factor between the histograms, it is the integral (total number of entries) in histogram H_a divided by the integral of histogram H_b. The standard deviation is denoted by σ.

Since the number of entries for an individual bin follows a Poissonian distribution, the denominator in the above equation can be expressed as

$$\sqrt{\sigma^2(a_i) + K^2 \cdot \sigma^2(b_i)} = \sqrt{a_i + K^2 \cdot b_i}\,.$$

Equation (4.1) comes from the definition of the significance test described by Li & Ma (1983) [116], used to estimate the statistical significance of observations in gamma-ray astronomy. It can be seen from Equation (4.1) that if $a_i = b_i$, i.e. the two distributions being compared are the same, then the significance $S = 0$. This metric offers a quantitative value on the difference between two distributions, and can be *minimised* such that the distributions become more and more equal. Note that the inclusion of K in the definition of the significance implies the *shape* is compared between the histograms, and not the scaling factor. And so, the observables between data and simulation for different values of select MUPAGE parameters can be compared through the significance test, determining the best agreement in the shape of the distributions, and then can be re-scaled through the parameter K_{0a}.

4.3. MUPAGE Tuning to KM3NeT Data

4.3.1. Data Run Selection

Following the addition of two more detection units to the KM3NeT/ORCA detector, these efforts to tune MUPAGE to data switched focus from ORCA-4 to ORCA-6. A data run from ORCA-6 is chosen to tune MUPAGE parameters on. This run should

be well-understood, well-calibrated, and should experience minimal bouts of high-rate veto.

One such run which fits these requirement is run **7398**, taken in February 2020. This run had a sufficient run duration (> 10,000 seconds, for sufficiently long data taking). In addition, this run was used as a reference for calibration procedures on ORCA-6 data, due to the detector experiencing a low sea current during the run. This data-taking run is used to compare the observables resulting from the nominal MUPAGE values to the data, shown in Figure 4.1. This same run is used for the tuning procedure described below.

4.3.2. Iterative MUPAGE Parameter Scan

The MUPAGE parameters K_{1a}, ν_{1b}, ρ_{0b}, β, b_{1b} and d_{0b} are varied one at a time, scanning a range of 0.6 – 2.1 times their default value, and used to obtain distributions of the select observables. As will be shown, this range is sufficient for varying the parameter values in order to find a new optimum value, according to the significance test. The individual significance for each observable is calculated between the simulation and the data, and the total **average significance** is then used in order to quantify the best agreement between data and simulation for all the observables considered. Since the minimum significance (i.e. best agreement between data and simulation) induced by one parameter value for one observable will not necessarily result in the minimum significance for another observable, focusing on the average significance value for all observables simplifies the matter. It gives one metric for one set of MUPAGE values.

When an optimum value of a MUPAGE parameter is found (i.e. a MUPAGE value which results in the minimum average significance when comparing distributions between simulation and data), this value is chosen as the new default value for that parameter, and the scan over parameters is carried out again. This process is continued until no lower value of the significance can be found — the minimum in the significance landscape has been reached. This method of tuning MUPAGE parameters to KM3NeT data is reported in Ó Fearraigh, B (2021) [117].

An example is shown for the first iteration of this parameter scan. The average significance as a function of the parameter value, for all 6 MUPAGE parameters, is presented in Figure 4.3. For some of the parameters, a lower value of the average significance is found. This value is used as the new 'default' MUPAGE parameter value and the scan is carried out again. As can be seen, the range of parameters scanned

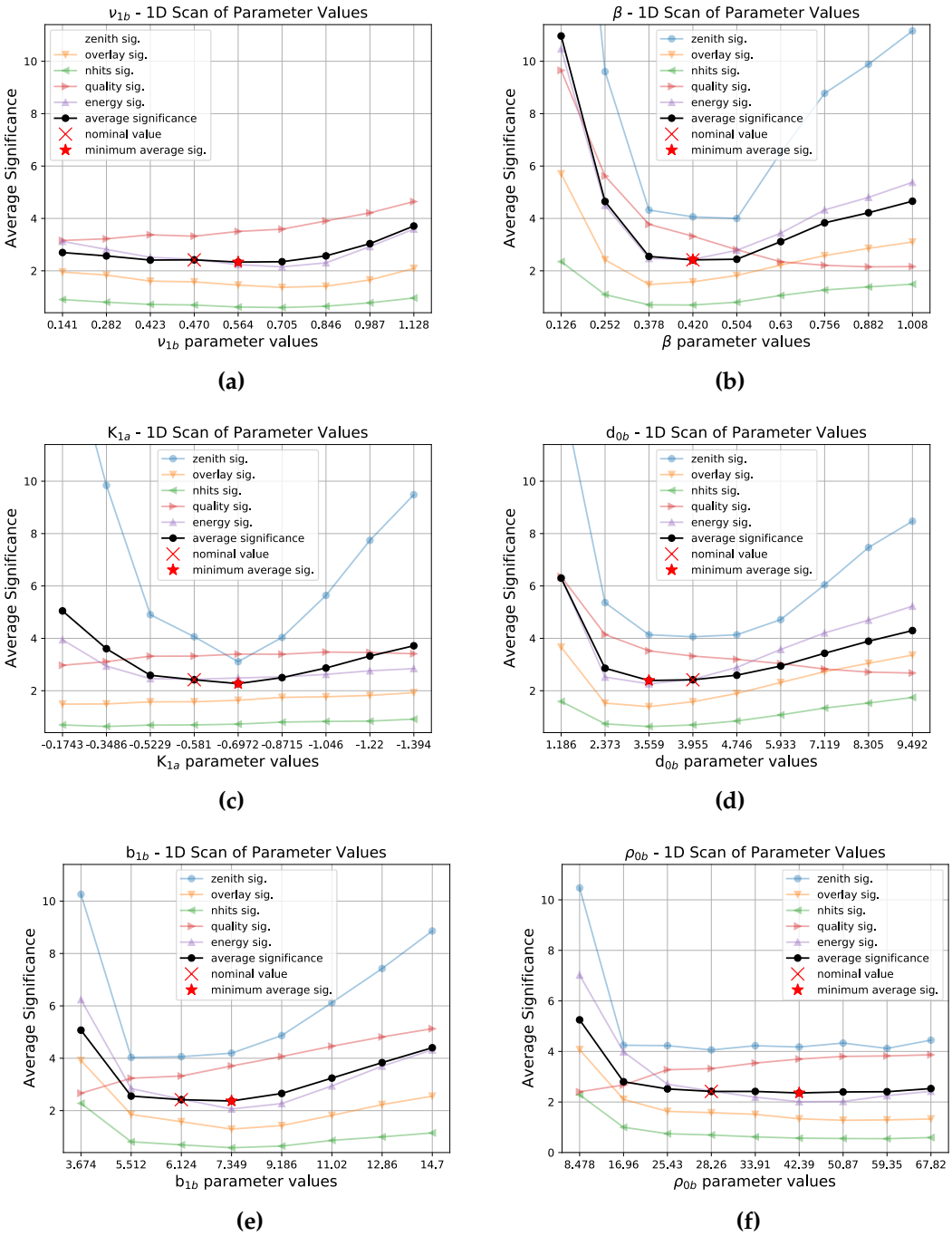


Figure 4.3.: The average significance as a function of the parameter values, for the parameters (a) ν_{1b}, (b) β, (c) K_{1a}, (d) d_{0b}, (e) b_{1b} and (f) ρ_{0b}, including the significance curves for each observable. The nominal MUPAGE value and the minimum significance value are indicated in each figure. The new minima found for each parameter scan are used as the new nominal values for another iteration of the parameter scan.

appears to be sufficient for finding a new minimum within this range, albeit for the parameter ρ_{0b} the average significance shows a rather flat behaviour for larger parameter values in the range. For the same scan, the contribution from each individual significance

to the average, for each observable is also shown in Figure 4.3. The significance of the zenith angle and energy distribution are the most influenced by the changes in the parameter values. The significance of the reconstruction quality shows a behaviour that does not follow that of the other observables. This observable appears to be the least sensitive one when varying MUPAGE parameter values. The parameters which least affect the shape of the number of hits or the number of overlays are those pertaining to the flux, i.e. K_{1a} and ν_{1b}.

Parameter	Nominal Value	Tuned Value
K_{0a}	0.0072	0.009896
K_{1a}	-0.581	-0.6972
ν_{1b}	0.470	0.564
β	0.420	0.630
ρ_{0b}	28.26	42.39
d_{0b}	3.955	3.203

Table 4.1.: MUPAGE parameter values before and after the tuning procedure.

This iterative scan is continued, until a new minimum is found for as many parameters as possible, and their values are updated. The scan is continued until a new minimum cannot be found for the remaining parameters. The new, tuned values for the MUPAGE parameters are shown in Table 4.1.

The scaling of the distributions can be addressed through changing the parameter K_{0a}. The amount of time needed to generate the number of muons which can describe the data run, i.e. the generated lifetime, is used to alter the scaling. The un-tuned MUPAGE parameter values resulted in a generated lifetime of 10,582 seconds when simulating this ORCA-6 run. The parameter values shown in Table 4.1 result in a generated lifetime of 14,544 s. With the un-tuned value of $K_{0a} = 0.0072$, the new value can be found via

$$K_{0a}^{\text{tuned}} = 0.0072 \times (14,544/10,582) \simeq 0.009896\,.$$

The distributions of the observables are presented in Figure 4.4 for the tuned MUPAGE parameters, and are compared to the KM3NeT/ORCA-6 data run and the un-tuned MUPAGE parameterisation. The tuned MUPAGE parameterisation shows a better description of the data compared to the un-tuned MUPAGE parameterisation, most noticeably for the reconstructed zenith angle and reconstructed energy distributions. For example, the most down-going bin for the reconstructed zenith angle distribution shows

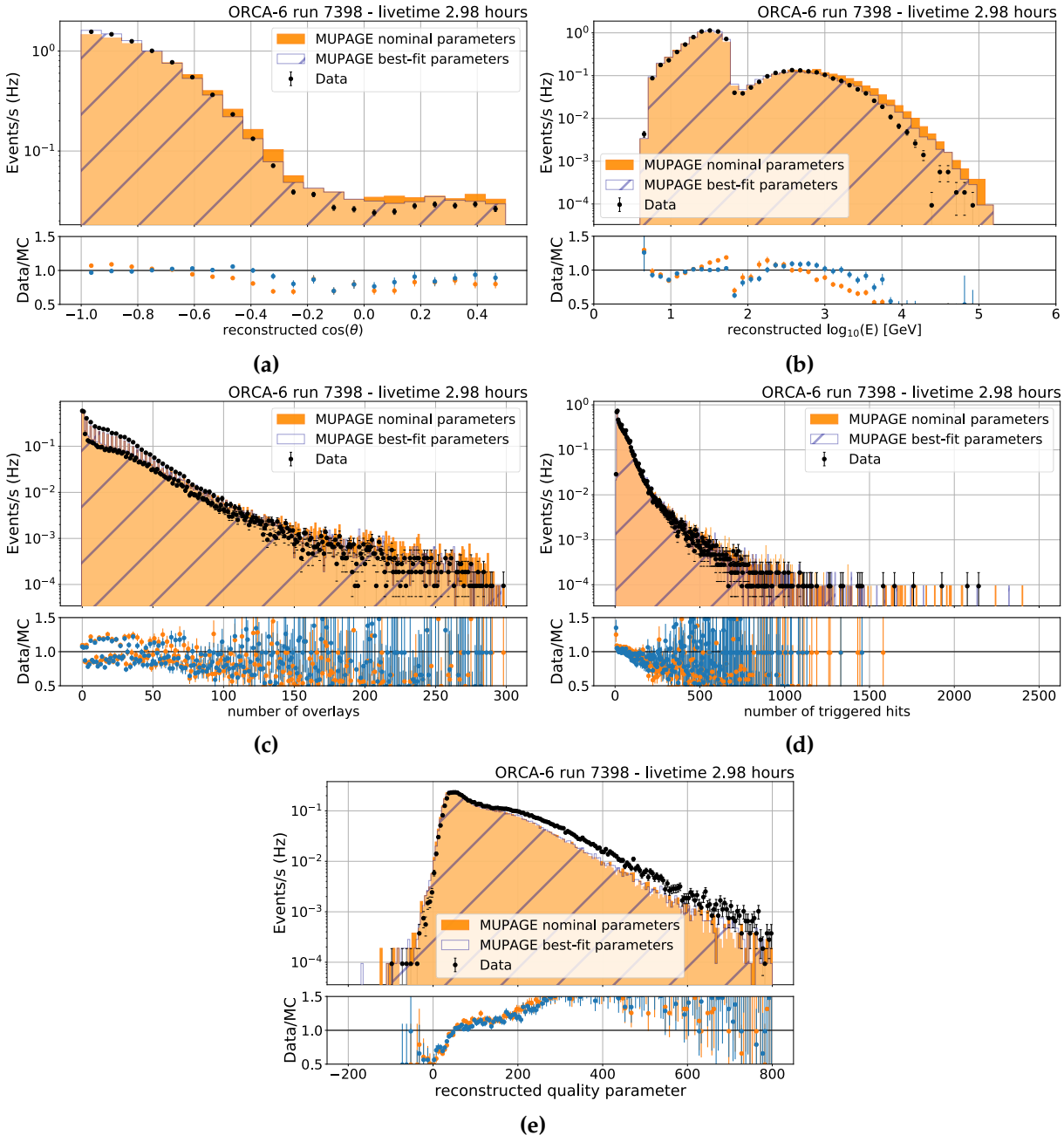


Figure 4.4.: The distributions of the (a) cosine of the zenith angle, (b) reconstructed energy, (c) number of overlays, (d) number of triggered hits, and (e) quality parameter (i.e. the reconstructed likelihood), comparing the data from a select ORCA-6 data run and the simulation, for both the nominal MUPAGE parameter values and the new values obtained through the parameter scan described in the text. The ratio of the data to the MC is also shown. The new, tuned values result in an improved description of the data, in particular for the zenith angle and reconstructed energy distributions.

a difference between the data value and simulation value of 2%. There is also a much improved agreement between the data and simulation for almost the entire range of the

reconstructed zenith angle distribution. A better agreement between the simulation and data has also been induced for the high-energy muons.

This improvement comes into effect for the ranges of parameter values which are important for physics analyses in KM3NeT/ORCA, for example the down-going atmospheric muons from $\cos(\theta_{\text{zenith}})$ = [-1, 0.0] and reconstructed energy from 10 to 100 GeV. These regions of the parameter space form the most significant background in identifying neutrino events which are pertinent to determining the neutrino oscillation parameters and the neutrino mass ordering.

4.3.3. Two-Dimensional Scan

An attempt is made to further improve the description of the KM3NeT data by MUPAGE by varying more than one parameter value at once, i.e. by carrying out a two-dimensional scan of the parameter space. A scan over more than one parameter is attempted to reduce the remaining deviations seen between the observables from data and simulation.

An additional parameter to be tuned is introduced here: γ_1. This parameter affects the slope of the energy spectrum in the MUPAGE parameterisation (see Section 3.4.1). It was not used in the iterative tuning procedure in order to limit the number of varied parameters. Three parameters are chosen for this scan which relate to the energy spectrum: γ_1, β and b_{1b}. For each value of one parameter, a range of values of the other parameter is used. All of the other MUPAGE parameters are not changed, their values are set to their nominal values. The tuned values of Table 4.1 are here used as the nominal values for those parameters.

Again, for each set of varied parameter values, the simulation chain of atmospheric muon generation, light propagation and detection, (run-by-run) triggering and muon track reconstruction is ran. The same ORCA-6 run as before is used. The average significance is calculated to estimate the best agreement between simulation and data.

The parameter space of the average significance for these 2-D scans is shown in Figure 4.5. A grid of points is shown, with each point representing a unique set of the two parameters being varied, each with an associated value of the average significance. The average significance in between each points is filled via the interpolation method used to create these plots. The position of the nominal value is also indicated in the parameter space. For the case of Figure 4.5 (a), the average significance for different values of (γ_1,β) is calculated. No value of the average significance is found with a value lower than that for the nominal parameters for the parameter space values indicated by

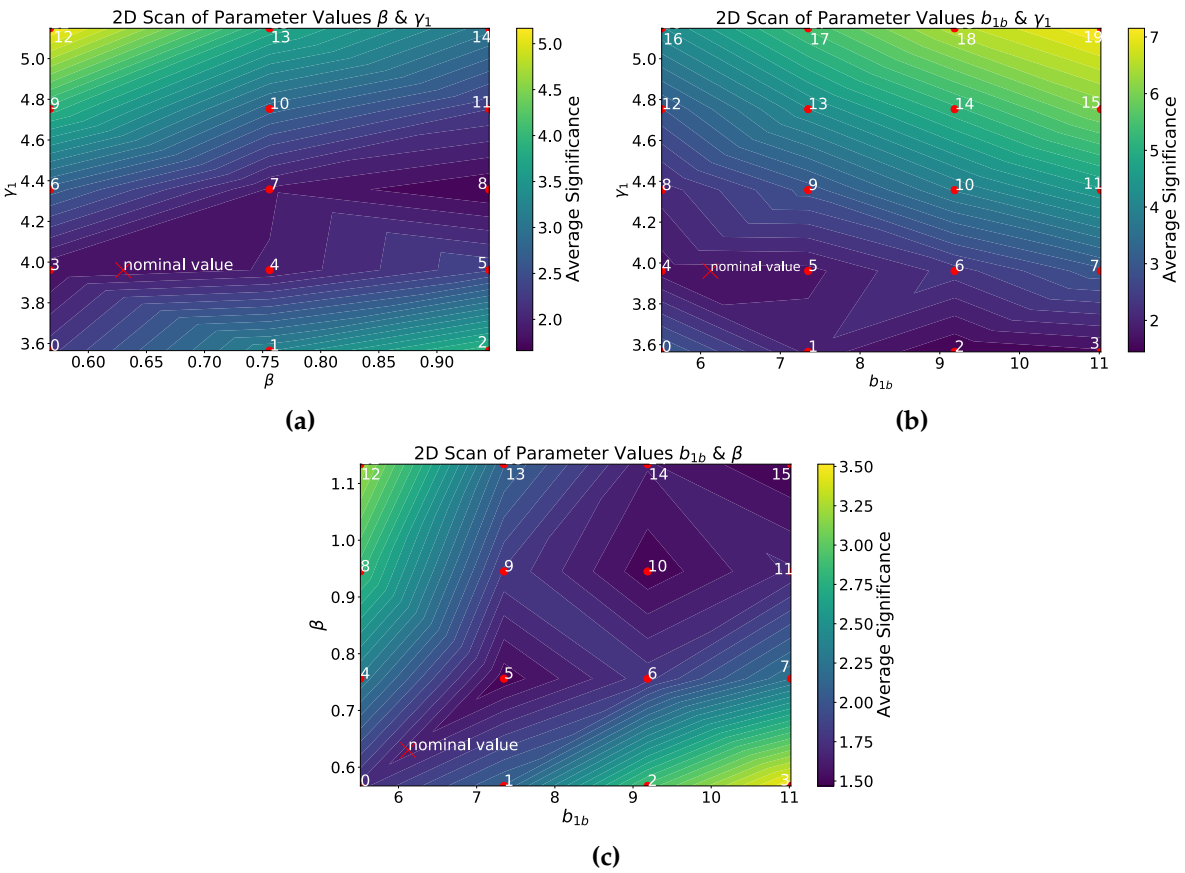


Figure 4.5.: 2-D scans of the average significance parameter space for values of (a) (β, γ_1), (b) (b_{1b}, γ_1), and (c) (b_{1b},β). The parameter combinations which are used for this scan are indicated by the red points. Each point has an associated average significance value, and the plotting tool used here interpolates the average significance between these points. The location of the nominal value (updated to the new, tuned values) is indicated in each figure.

the points in the figure. Figure 4.5 (b) and Figure 4.5 (c) show the parameter space for values of (γ_1 ,b_{1b}) and (b_{1b},β), respectively. The values of γ_1, b_{1b} = (9.18, 3.565) result in an average significance lower than that of the nominal value. However, in comparing the zenith and energy distributions, the difference between the nominal distribution and this parameter set are within the error margins associated with the distributions. Any improvement in describing the data is marginal and within the realms of statistical fluctuations. Similarly, b_{1b},β = (11.02, 1.134) results in a parameter combination which yields a lower average significance than that of the nominal parameters. Again, there is little difference between the zenith angle and energy distributions for this parameter set and the nominal values. As such, the nominal MUPAGE parameters are not updated following the results of this scan.

4.4. Applying the MUPAGE Tuning to KM3NeT/ORCA-6

The updated MUPAGE values of Table 4.1 following this tuning procedure are used to simulate the atmospheric muons for some runs of the ORCA-6 data set. A total of 9 data runs are simulated with the MUPAGE default parameter values and also the tuned parameter values. In this way the performance of the tuning is checked over a larger time scale, to see if it improves the data and simulation agreement for a larger KM3NeT/ORCA-6 data set.

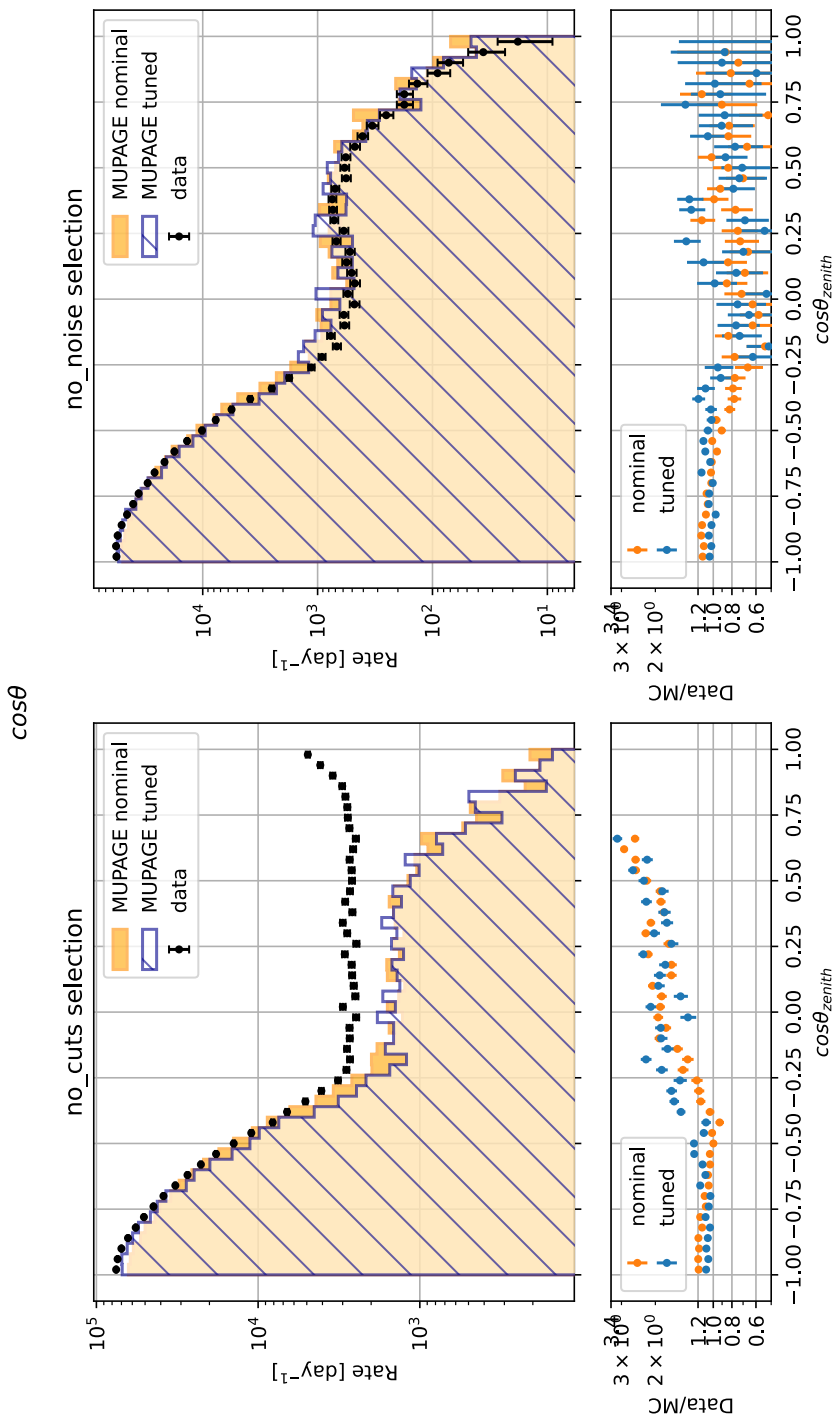


Figure 4.6.: The cosine of the zenith angle distribution for a number of ORCA-6 data runs. The atmospheric muon simulation for the nominal MUPAGE parameters and the tuned parameters are compared to the data. On the left: no cuts, on the right: an anti-noise cut is included.

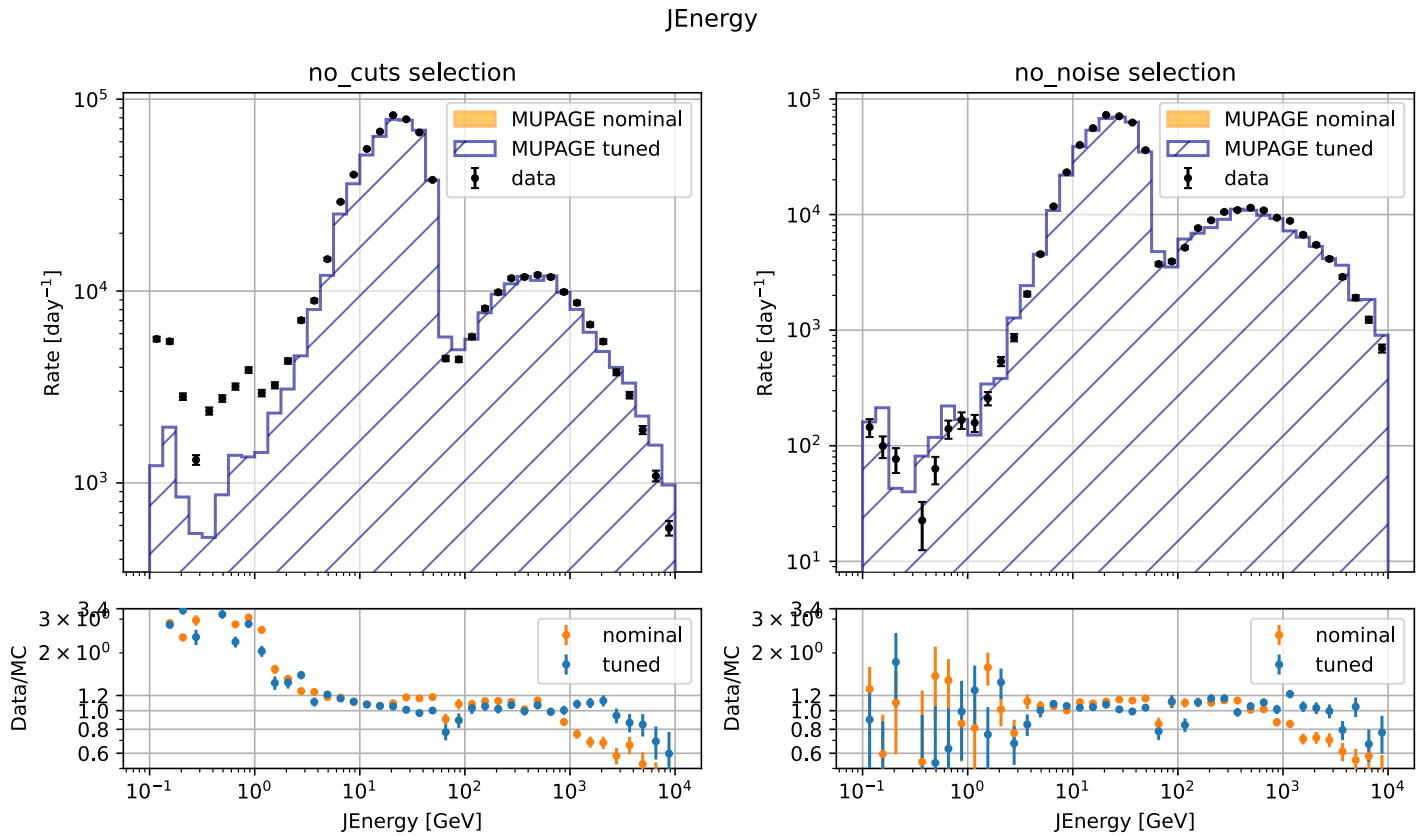


Figure 4.7.: The reconstructed energy distribution for a number of ORCA-6 data runs. The atmospheric muon simulation for the nominal MUPAGE parameters and the tuned parameters are compared to the data. On the left: no cuts, on the right: an anti-noise cut is included.

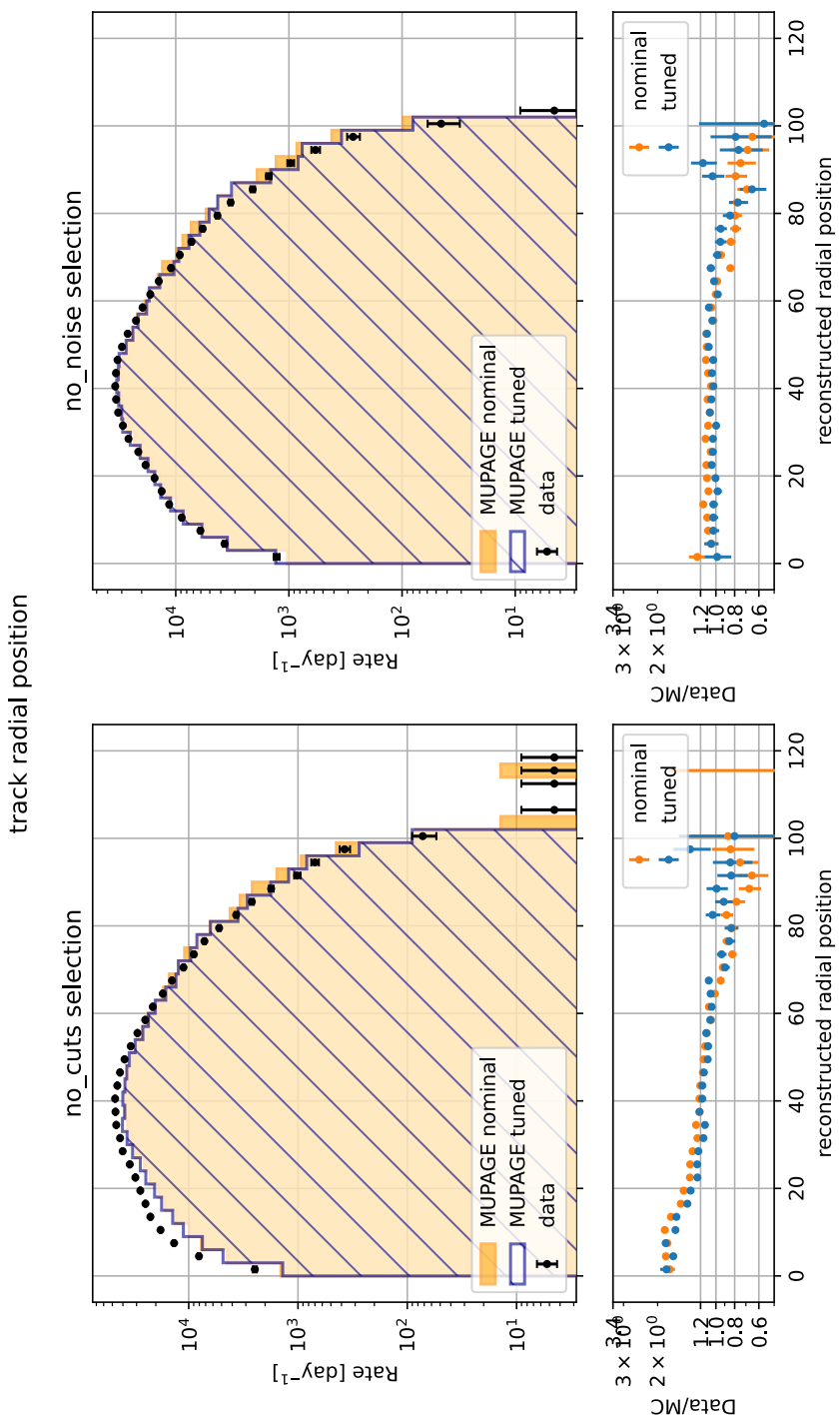


Figure 4.8.: The radial track position distribution for a number of ORCA-6 data runs. The atmospheric muon simulation for the nominal MUPAGE parameters and the tuned parameters are compared to the data. On the left: no cuts, on the right: an anti-noise cut is included.

The results are shown in Figures 4.6 to 4.8. The reconstructed zenith angle, reconstructed energy distribution, and track radial position distribution (an observable often used in KM3NeT/ORCA analyses) are shown, with no cuts (left figures), and with an anti-noise cut used in KM3NeT/ORCA analyses, which requires that the number of triggered hits > 15 (right figures). The values obtained with data divided by the values from the simulation are also shown. The data-simulation agreement improves when the anti-noise cut is added. The tuned parameters improve the description of the data compared to the nominal MUPAGE values. The tuning procedure is again shown to be successful for a larger data set.

4.5. Outlook & Discussion

This method of tuning MUPAGE parameters to KM3NeT data has been proven successful, with the new parameters resulting in an improved description of the data, in particular for the reconstructed zenith angle and reconstructed energy distributions. Using the significance test as a metric by which to determine the level of agreement between the distributions from simulation and data has also been proven successful. The selected parameters to be tuned - only 6 out of a potential 40 - are varied one by one, and bring about an improved data-MC agreement. There is potential for a multi-dimensional scan of the parameter space, as has been shown by carrying out a 2-D scan. The iterative scan of the parameter space is quite coarse and further probing of the significance landscape as a function of the different parameters could yield even further improvement in describing the data.

Noticeably, there is some remaining discrepancy between the observables from simulation and data; the observables are not in 100% agreement following the tuning procedure. There is a correlation between the high energy muons and those which are mis-reconstructed in direction, they end up in the upward-going direction regime (specifically $\cos(\theta_{\text{zenith}}) \simeq$ [-0.2, 0.5]). The two-dimensional scan attempted to address this relation by tuning parameters which affect the slope of the energy spectrum. This correlation is an outstanding one and requires further investigation in the context of parametric tuning of observables in simulation to data.

As alluded to already, the tuning procedure itself can be improved. In addition to this, the MUPAGE parameterisation itself may be incomplete in its description of the atmospheric muon; it is merely a parameterisation after all. There may be aspects of the

flux which are not simulated correctly or at all. The number of overlays was used as a proxy for the multiplicity, a quantity we have no direct access to in the data.

Tuning these MUPAGE parameters on one ORCA-6 data run results in an improved description of the data for the ORCA-6 data set spanning much longer time periods than a single run. Applying the tuned parameter values to KM3NeT/ARCA are the natural next step of this research, though this falls outside the realms of this thesis. This tuning procedure is carried out for one ORCA-6 data run, however there is potential to perform a tuning across both of the KM3NeT/ARCA and KM3NeT/ORCA detectors simultaneously.

A simulation which describes the recorded KM3NeT data to a high degree of precision and accuracy is vital for the physics analyses which can be performed by the KM3NeT experiment. As discussed, the atmospheric muon signal forms the primary background for the neutrino physics research in KM3NeT/ARCA and KM3NeT/ORCA, be it the determination of the neutrino mass ordering or the identification of astrophysical neutrino sources. Also, cosmic ray air shower simulations which encapsulate air shower interaction models can be tuned to both data and to the MUPAGE atmospheric muon simulation, as means of determining the interaction models and atmospheric muon flux which best describe the data.

Chapter 5.

Improving the Track Reconstruction for the KM3NeT/ORCA Detector

In this chapter, the inclusion of the track length in the track direction estimate of KM3NeT events is described. This is shown to improve the performance of the reconstruction of low-energy events, relevant for the KM3NeT/ORCA detector. The angular resolution is shown to be improved for the case of muons produced by ν_μ charged-current (CC) interactions. The energy resolution is also shown to be marginally improved for these low-energy muons. Simulated ν_μ-CC events for the 115-line KM3NeT/ORCA detector are used in this study, in order to quantify the performance of the reconstruction for the complete detector.

5.1. Inclusion of the Track Length in the Track Direction Reconstruction

5.1.1. Motivation

As noted in Section 3.7.3, the hits used for the various track reconstruction stages are considered within a cylinder around the hypothesised track, of user-defined radius and of infinite length — implying an infinitely-long muon track. Considering an infinitely-long cylinder aims to include as many hits as possible in the fitted direction of the track. Once the direction (and also position and time) of the track has been found, the track length is then estimated as described in Section 3.7.3. Reconstructing the direction of track-like events in the KM3NeT detectors is aided by the 'lever arm' effect. Typically, the greater the distance between hits measured along the length of the track, the better

the estimate of its direction. A longer lever arm (and track length) contributes to a better constraint on the direction.

A reminder of the different track reconstruction chain sequences between the KM3NeT/ARCA and KM3NeT/ORCA detectors can be found in Section 3.7.3. Using a cylinder of infinite length is done throughout the track reconstruction chain for events detected by KM3NeT/ARCA. This same approach (and reconstruction chain ordering) was also used to reconstruct events for KM3NeT/ORCA.

Thanks to the following investigation, the track length is now used to constrain the length of the cylinder used in the advanced direction fit stage of the track reconstruction chain, for KM3NeT/ORCA events. This direction fit stage is carried out by the `JMuonGandalf` program (see Section 3.7.3). For the neutrino energies detected by the KM3NeT/ORCA detector, the muon tracks can have short track lengths of a few metres; including hits many metres beyond the physical muon track length are of little help to the direction estimate and serves to introduce more background hits to the fit procedure.

5.1.2. Constraining the Hits

The track length is determined by a program named `JMuonStart`, and is estimated as described in Section 3.7.3. The beginning and end point of the track are determined in this fit stage. Here, Z_{min} and Z_{max} denote the respective start and end points of the track, measured along the track.

The selection of hits to be used in the advanced direction fit has been updated to include the track length; if Z_{min} and Z_{max} have been determined, they are used to define the length of the cylinder in which to select hits. The coordinate system of Figure 3.1 is used throughout the track reconstruction software and allows us to define Z_{min} and Z_{max}. As a reminder, the coordinate system is defined such that the track direction is along the positive z-axis.

The condition to be satisfied is the following (c.f. Figure 3.1):

$$\begin{aligned} z_{\text{hit}} - z_{\text{track}} &\geq Z_{\text{min}}\,; \\ z_{\text{hit}} - z_{\text{track}} - \frac{R_{\text{hit, track}}}{\tan\theta_C} &\leq Z_{\text{max}}\,. \end{aligned}$$

where z_{hit} and z_{track} are the z-coordinates of the hit and the track, respectively. The first condition ensures z_{hit} is downstream (towards positive values along the z-axis) of or

equal to the track start point. The second condition ensures that z_{hit} is upstream (towards negative values along the z-axis) of the track end point, and includes the distance in the z-direction in which direct (un-scattered) Cherenkov light may travel from the track z-position to the hit (i.e. PMT) position.

So, if the track start and end points have been determined, they are used to define the length of this cylinder around the track in which hits are considered for the direction fit. If not, the values of Z_{min} and Z_{max} are set to those of an infinitely-long track. In this way the same software can be used for track reconstruction in the KM3NeT/ARCA and KM3NeT/ORCA detector.

5.1.3. Comparing Muon Direction and Energy Estimates

Extensions to the Start & End Point

Before evaluating the impact of this addition to the track reconstruction, the range of $[Z_{\text{min}}, Z_{\text{max}}]$ in defining the track length is investigated, in terms of its resulting angular resolution. The conditions outlined above in defining Z_{min} and Z_{max} only account for direct Cherenkov radiation from the track. Upstream of the track start point, some Cherenkov light from the muon can be scattered in the backwards direction which can be detected by PMTs. Downstream of the track end point, the Cherenkov radiation emitted by the muon track may still reach some PMTs through scattering.

A few test cases are investigated, where the length of the cylinder in which hits are considered for the direction fit is extended upstream of the track start point and downstream of the track end point. This is carried out for a sample of simulated ν_μ-CC events for the ORCA-115 line detector, in the range 1–10 GeV. The neutrinos are generated using the gSeaGen program, and the simulation of the number of detected photo-electrons is carried out by `JSirene`. The events are propagated through the trigger stage and the reconstruction chain sequence of

$$\texttt{JMuonPrefit} \rightarrow \texttt{JMuonSimplex}$$
$$\rightarrow \texttt{JMuonStart} \rightarrow \texttt{JMuonGandalf} \ .$$

In quantifying how well the muon direction is reconstructed, the angular resolution is determined, a metric described in Section 3.7.6. The angle between the true muon direction (known from the simulation) and the reconstructed muon direction is found,

denoted by $\Delta\theta_{\text{true}\,\mu,\,\text{reco}\,\mu}$. The median angular resolution is shown in Figure 5.1 for these test cases. The muon energy is shown up to a value of 9 GeV, since beyond this point there are limited statistics in the simulated event sample. For example, [-3,+3] m implies the length of the cylinder is defined as $Z_{\text{min}} - 3$ and $Z_{\text{max}} + 3$, in metres. From these examples, the best angular resolution is achieved by including hits 3 metres upstream of the track start point and +9 metres downstream of the track end point. These values allow for the inclusion of hits which are likely caused by the muon track emitting Cherenkov radiation. These values of $[Z_{\text{min}} - 3,\, Z_{\text{max}} + 9]$ are used in comparing to the case of an infinite track length in the following section.

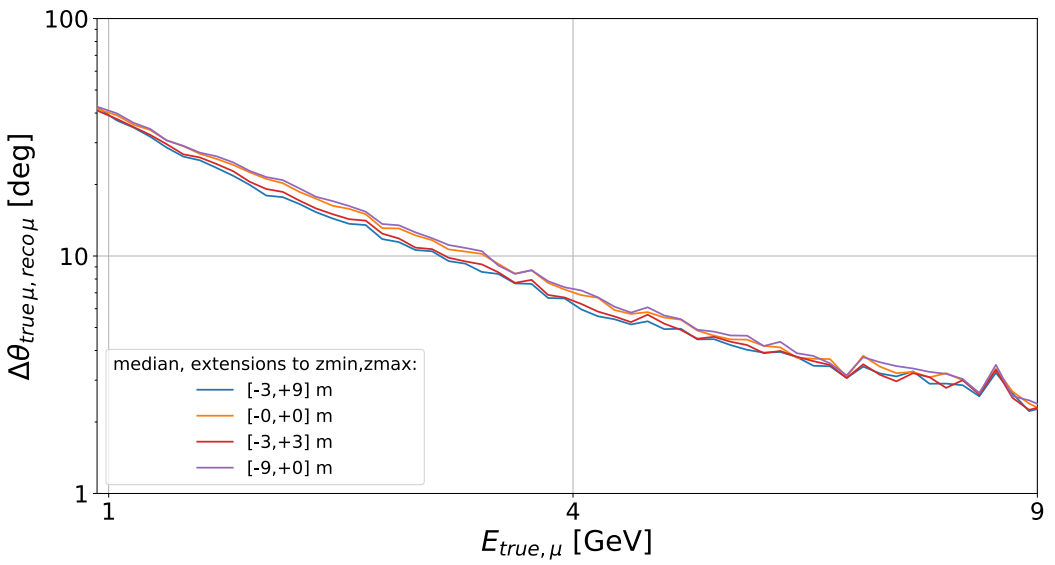


Figure 5.1.: The angular resolution relative to the true muon direction is shown as a function of the true muon energy. The length of the cylinder used to constrain the hits in the direction fit is extended a few metres upstream and downstream of the track start and end point, respectively, for a few different values.

Angular Resolution

To investigate whether the inclusion of the track length in the track direction reconstruction procedure yields an improvement, two cases are considered. In the first case, the direction is reconstructed according to the sequence of direction prefit - intermediate direction fit - advanced direction fit, i.e. following the sequence

$$\texttt{JMuonPrefit} \rightarrow \texttt{JMuonSimplex} \rightarrow \texttt{JMuonGandalf}\ [\text{no track length inclusion}]\,.$$

The second case carries out the same direction fit, with the track length estimated before the advanced direction fit stage, following the aforementioned sequence of

$$\texttt{JMuonPrefit} \rightarrow \texttt{JMuonSimplex} \\ \rightarrow \texttt{JMuonStart} \rightarrow \texttt{JMuonGandalf} \; [\text{track length included}] \, .$$

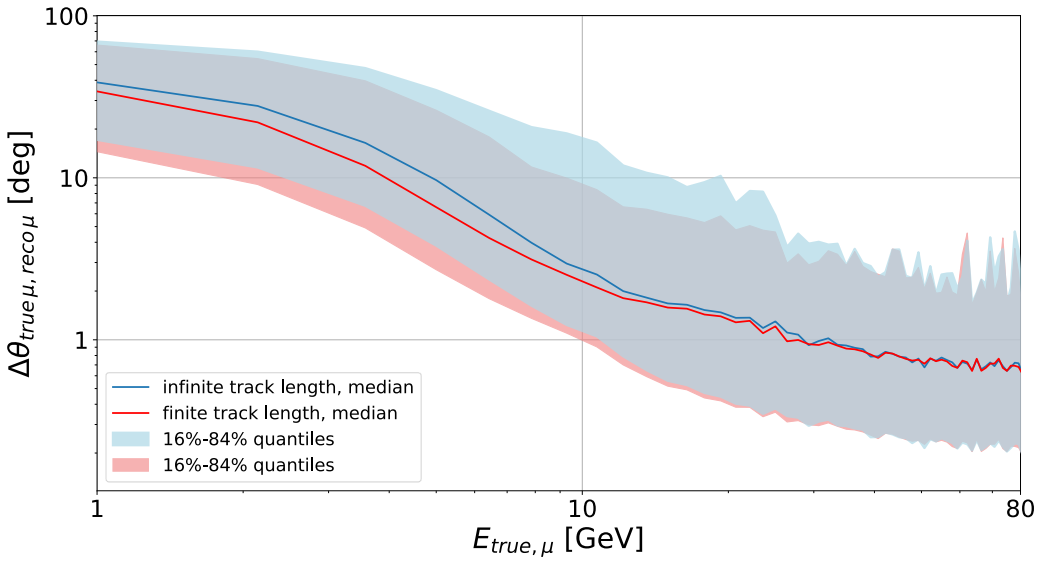
The impact on the full 115 line KM3NeT/ORCA detector is evaluated for these two cases, for ν_μ-CC events propagated through the simulation chain as before, in the neutrino energy range of 1–100 GeV.

Again, $\Delta\theta_{\text{true}\,\mu,\,\text{reco}\,\mu}$ is determined to illustrate the angular resolution. Since the direction of the neutrino is the ultimate quantity to be determined - pertaining to how well KM3NeT/ORCA can determine the neutrino mass ordering - the angular deviation between the true neutrino direction and reconstructed muon direction is also evaluated, denoted by $\Delta\theta_{\text{true}\,\nu,\,\text{reco}\,\mu}$. The quantities $\Delta\theta_{\text{true}\,\mu,\,\text{reco}\,\mu}$ and $\Delta\theta_{\text{true}\,\nu,\,\text{reco}\,\mu}$ are determined as a function of the true muon and true neutrino energy, separately. The median of both distributions of the angular deviation as a function of the respective energy are used to illustrate the angular resolution as a function of energy.

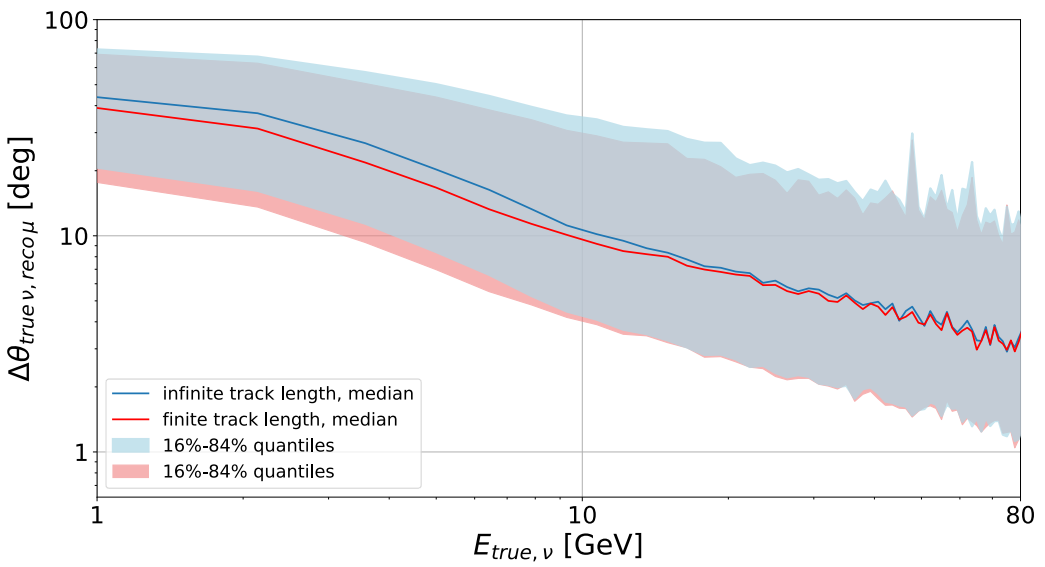
The angular resolution of reconstructed ν_μ-CC events, with and without the inclusion of the track length in the direction reconstruction, is shown in Figure 5.2. The median angular resolution of the reconstructed muon direction with respect to the true muon and true neutrino direction is shown as a function of the true muon and true neutrino energy, respectively. In addition to the median (the 50% quantile), the 16% and 84% quantiles are also shown. The muon energy is shown up to a value of 80 GeV, since beyond this point there are limited statistics in the simulated event sample.

Especially at lower energies, using the track length to constrain the length of the cylinder in which hits are considered yields an improvement in the angular resolution. The band from 16% – 84% shifts downwards on the plot, indicating improved values of the angular resolution. No selection criteria are incorporated here; this gives an overview of the performance of the reconstruction alone.

At higher energies, the inclusion of the track length in the direction estimate brings about an improvement for muon energies up to $\sim$25 GeV. It is apparent that the resolution is better when comparing the true and reconstructed muon direction, as opposed to the neutrino direction. The reconstructed muon direction only serves as a proxy for the true neutrino direction, with an inherent scattering angle existing between the two. In addition, for ν_μ-CC events, there is a hadronic shower created in this interaction which



(a)



(b)

Figure 5.2.: The angular resolution is compared between the cases of excluding (blue line) and including (red line) the track length in the direction reconstruction. Also indicated as a band for both cases is the 16% – 84% region. The angular deviation between the true muon and reconstructed muon as a function of muon energy is shown in (a), whilst the angular deviation between the true neutrino direction and the reconstructed muon direction as a function of neutrino energy is shown in (b).

is not reconstructed by the track reconstruction algorithms. To determine the neutrino direction accurately, the full kinematics of the event, muon track and hadronic shower, should be considered.

Energy Estimation

Given the improvement in the angular resolution when using the track length estimate as a prefit to the advanced direction fit, the results of the track energy estimate are also investigated. If the track direction has been determined more accurately, the energy estimate of the track should also improve.

The same two cases as before are investigated, with the addition of a second estimate of the track length (i.e. determining the track length now with the improved track direction estimate), and the track energy estimate. The track energy is estimated by fitting the energy, as described in Section 3.7.3. In terms of the reconstruction chain ordering, these two cases are

$$\texttt{JMuonPrefit} \rightarrow \texttt{JMuonSimplex} \rightarrow \texttt{JMuonGandalf} \rightarrow \texttt{JMuonStart} \rightarrow \texttt{JMuonEnergy},$$

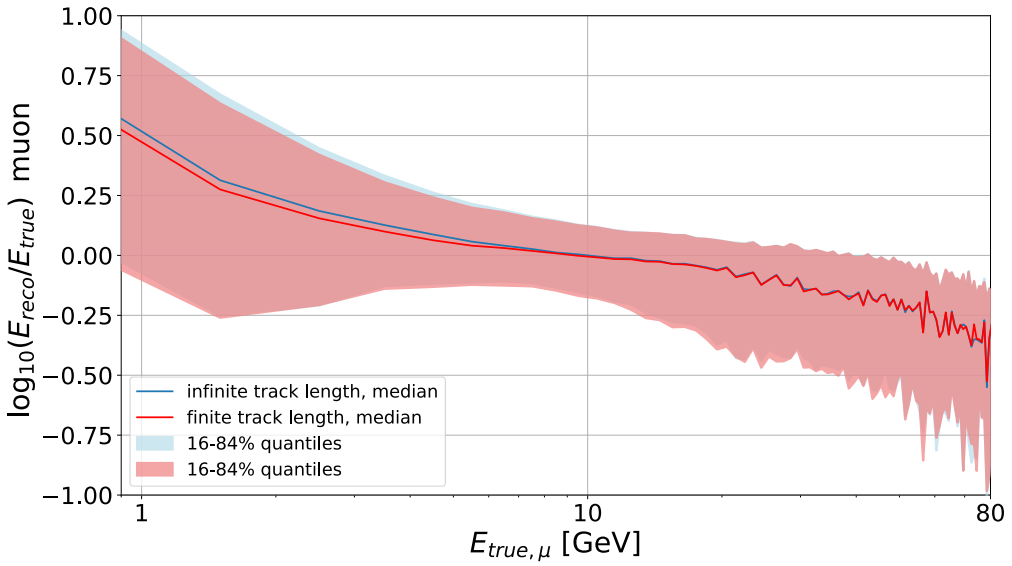
and

$$\texttt{JMuonPrefit} \rightarrow \texttt{JMuonSimplex} \rightarrow \texttt{JMuonStart} \rightarrow \texttt{JMuonGandalf} \rightarrow \texttt{JMuonStart} \rightarrow \texttt{JMuonEnergy}.$$

In the latter sequence, the track length is included in the advanced direction fit stage.

The quantity $\log_{10}\left(E_{\text{reco}}/E_{\text{true}}\right)$ as a function of the true muon energy, is determined. This means of quantifying the energy bias and resolution is mentioned in Section 3.7.6. The same two cases described above are compared from the same simulation of ν_μ-CC events for the 115-line KM3NeT/ORCA detector. This is shown in Figure 5.3. Again, all events are included here, no selection criteria are used to select events. The median and 16% – 84% quantiles are indicated. Again, the muon energy is shown up to a value of 80 GeV, since beyond this point there are limited statistics in the simulated event sample.

At lower energies, there is a marginal improvement from including the track length in the direction fit, with a slight reduction of the spread around the median value (the resolution) and a slightly reduced bias (the median line is closer to $\log_{10}\left(E_{\text{reco}}/E_{\text{true}}\right) = 0$). At higher energies, the performance of the reconstruction in determining the energy



(a)

Figure 5.3.: The quantity $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ as a function of the true muon energy is compared between the cases of excluding (blue line) and including (red line) the track length in the direction reconstruction. Also indicated as a band for both cases is the 16% – 84% region. The effect of introducing the track length on the direction fit is marginal in improving the energy estimation.

is almost the same between the two cases, with almost no change in the resolution and bias. In addition to this, there are some noticeable features of these figures. Below $\sim$5 GeV, there is a large spread on the values of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$, indicating a poor energy resolution. At these energies, the muon tracks are quite short and difficult to distinguish, not producing many hits in the detector. Beyond $E_{\text{true muon}} \sim 5$ GeV, the energy estimate improves. From approximately 5–15 GeV, the median of the distribution moves closer to $\log_{10}(E_{\text{reco}}/E_{\text{true}}) = 0$, as the true muon energy increases. In this energy regime, the muon has a sufficiently-long track length and energy to be reconstructed well, with the track length proportional to the energy as exhibited by a minimum ionising particle [3]. The resolution becomes worse beyond $\sim$15 GeV (the spread on the median value becomes wider). This is likely to be caused by events with sufficiently-long track lengths to leave the detector. For these events, the number of hits induced by the track and the reconstructed energy from these hits does not equate to the true track energy.

5.1.4. Discussion

As a result of this investigation, the reconstruction procedure for track events in KM3NeT/ORCA has been updated. The estimated track length is now used in the advanced direction fit stage, to constrain the length of the cylinder in which hits are included in determining the track direction. This improvement is implemented in the reconstruction of events from the deployed KM3NeT/ORCA detection units which are currently taking data, and also in sensitivity studies to the neutrino mass ordering and neutrino oscillation parameters, where this reconstruction is applied to simulated events for the full KM3NeT/ORCA 115-line detector [88].

The angular and energy resolution of KM3NeT/ORCA to muon track events, presented in this chapter, are valuable performance checks of the track reconstruction. In the following chapter, we will attempt to improve upon these resolutions further through a novel reconstruction technique.

Chapter 6.

A Novel Approach: Simultaneous Track & Shower Reconstruction

6.1. Motivation

The standard procedure to reconstruct events in the KM3NeT detectors is to apply both a track reconstruction and shower reconstruction algorithm to them. As mentioned in Chapters 2 and 3, two separate algorithms are used because track events and shower events produce distinctive signatures in the detectors, which can be treated separately in their reconstruction. In fact, the same procedure is applied in other Cherenkov-based neutrino detectors such as IceCube [118, 119]. The reconstruction of either of these event signatures is already a non-trivial task.

Applying both a track and shower reconstruction to the same event demands additional time and computational resources, but also necessitates another step in physics analyses - termed '*particle identification*' - which tries to identify the reconstructed events as shower-like or track-like. For high-energy neutrino interactions, as detected by KM3NeT/ARCA, the muon track signature alone is an incredibly clearly-defined one, with tracks spanning hundreds of metres (the track length of a muon is proportional to its energy), and resulting in many hits in the detector which comply with the hypothesis of Cherenkov radiation from a muon track. Particle showers also grow with increased neutrino energy, and since they expand in the longitudinal and radial direction, they can be distinguished from very long tracks. This makes track identification in KM3NeT/ARCA a manageable task. For KM3NeT/ORCA however, with extremely localised events at low (< 10 GeV) neutrino energies, muon tracks can be produced with a low energy and hence a short track length, and may not be so easily distinguishable from particle showers — in Chapter 5 we already encountered the need for special treatment of track events at

low energies, where the reconstructed track length was needed to better reconstruct the muon direction.

As discussed in Section 1.2.2, neutrinos with energies above 2 GeV will almost always create a hadronic shower in their interactions, following the hadronisation of particles created by resonant production processes or when a quark is freed from the nucleus following deep inelastic scattering processes. All neutrinos and anti-neutrinos which result in track-like events (i.e. of the muon or tau lepton flavour) will thus be accompanied by a hadronic shower. A reconstruction algorithm which can reconstruct *both* the track and shower components of such neutrino interactions would describe these events more completely.

When assuming only a muon or shower event in the reconstruction procedure, the kinematics of the event is not fully described, including information on the total energy of the interaction, the fractional energy transfer to the hadronic shower y (c.f. Equation (1.3)), and the actual neutrino direction. The reconstructed shower or muon direction and energy are currently used as a proxy for that of the neutrino direction and energy (the difference between the muon direction and the actual neutrino direction has already been illustrated in Figure 5.2). Although this standard procedure results in competitive angular and energy resolutions for KM3NeT/ARCA and KM3NeT/ORCA events [17,88], a more complete description of the event in its reconstruction should result in more precise estimates of the neutrino direction and energy. This would improve the ability of the KM3NeT experiment in determining the neutrino mass ordering, for example — the neutrino oscillations for normal and inverse ordering (see Section 1.2.4) are defined in terms of the neutrino energy and direction.

Ideally, a track and shower reconstruction procedure would also be able to reconstruct shower-like events even when the track component is not present (e.g. for neutral-current interactions or electron neutrino interactions). This track and shower reconstruction would also allow for a direct reconstruction of the fractional energy transfer to the hadronic shower, mentioned above. Reconstructing this physical parameter offers a means of fully understanding and describing the neutrino interaction. As described in Section 1.2.2, reconstructing y can allow for a discrimination between neutrinos and anti-neutrinos, based on the shape of the distributions. As an additional point, the dedicated electromagnetic shower reconstruction procedure (described in Section 3.7.4) is the one applied to events in KM3NeT/ORCA. There is no dedicated reconstruction technique used at the moment for the reconstruction of hadronic showers, for example a reconstruction algorithm making use of hadronic shower photo-electron distribution tables.

In this chapter, the conceptual description, technical implementation and performance of a novel **track+shower** reconstruction algorithm is described. The focus of this reconstruction is for events detected by the KM3NeT/ORCA detector. The angular and energy resolution of the reconstructed track and shower components is investigated, as well as the resolution of the neutrino direction and energy. This is carried out for charged-current muon neutrino interactions (which represent an event with a track and shower component), and charged-current electron neutrino interactions, simulated in the complete ORCA-115 line detector.

6.2. Event Model

The event model to be used in the reconstruction procedure is defined here. This model will be fitted to data, by maximising a likelihood function defined later in this chapter. The photo-electron distribution tables for muon and shower light described in Section 3.2 will be used to calculate the likelihood values for the event model parameters.

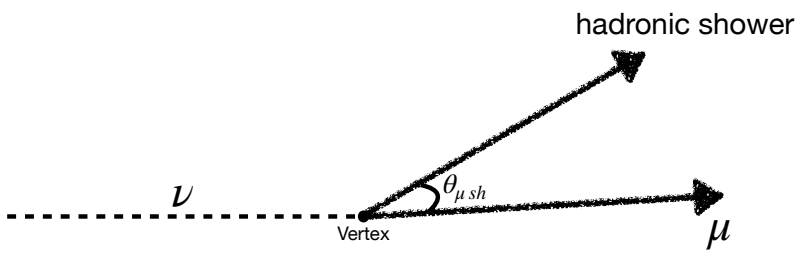


Figure 6.1.: The event model describing the track+shower event to be reconstructed. The neutrino ν creates a muon μ and hadronic shower at the interaction vertex. This vertex is denoted by the position (x_0, y_0, z_0) at a time t_0. Both the track and shower components of this event have an associated energy and direction. The angle between the two components is given by $\theta_{\mu\,sh}$.

The neutrino interaction point is termed the '*vertex*', which has an associated position (x_0, y_0, z_0) at a time t_0. At the vertex position, the particles created by the neutrino interaction are formed, and hence the vertex is the beginning point of both the track and particle shower. In the separate track and shower reconstruction techniques described in Chapter 3, the muon is estimated at its vertex position, while the shower is estimated

at the point of shower maximum (see Section 2.3). In this model, the muon and shower components both have an associated energy E_μ and E_{sh} respectively, as well as an associated three-dimensional direction: which can be defined using the angles (θ_μ, ϕ_μ) for the muon and (θ_{sh}, ϕ_{sh}) for the shower. An illustration of this event model is shown in Figure 6.1. These model parameters, the vertex position and time, and the component energies and direction, are the ones to be reconstructed. The muon is finite in length, beginning at the vertex and ending at the point at which it no longer produces detectable Cherenkov light. The end point of the track is not explicitly estimated, it is defined by being proportional to the muon energy (as explained in Section 2.3.1).

There is already an approximation used in this model. A particle shower consists of many particles created at the vertex (which then re-interact further) with their own directions and energies. Here, a total energy of the shower is assumed, as well as one total direction of the shower. For a hadronic shower, the charged particles within can produce Cherenkov radiation. The neutral particles within the shower produce no Cherenkov radiation, but can re-interact or decay and result in charged particles. Only the hits resulting from the Cherenkov radiation of charged particles in a shower allow the shower to be reconstructed. The true direction of the shower can thus not be fully reconstructed as the neutral particles go undetected, carrying some of the directional information with them, leading to an approximate estimate of the direction. In defining the true direction and energy of the shower component of this event, the visible shower energy E_{vis} will be used throughout this chapter unless stated otherwise. This concept is previously described in Section 2.3.2. The visible energy is the detectable energy from the shower due to Cherenkov light emission; for electromagnetic showers this is equivalent to the total energy, whilst for hadronic showers some of the energy is transferred to particles which produce no detectable light. The visible energy for hadronic showers is thus defined as the electromagnetic-equivalent energy produced by a hadronic particle, following the formulations in Dentler (2012) [120]. The true shower direction will be defined by the visible shower direction; it is the averaged direction of all particles in the hadronic shower, weighted by the visible shower energy.

6.2.1. Model Constraints

The beginning and end point of the track are defined when evaluating the expected number of photo-electrons along the track. Though the vertex position and time are included in the fit procedure described later, the end point is not included as a fit parameter. This is in contrast to what is done in the track reconstruction software, where the beginning

and end point are determined by a particular algorithm in the track reconstruction chain, by finding the points along the track where the signal probability exceeds a pre-defined threshold value (as described in Section 3.7.3). For the track+shower reconstruction algorithm, photons are not taken into account from points along the track where the track energy is below the muon threshold energy to emit Cherenkov radiation. This condition indirectly defines a track length.

A choice is made in this track+shower algorithm to make use of the pre-defined tables of the expected number of photo-electrons given the model parameters of an *electromagnetic* shower. A description of these tables can be found in Section 3.2. These tables for electromagnetic shower light are readily available and used in many parts of the KM3NeT simulation software, as described in Chapter 3. There are no equivalent tables for the photo-electron distributions of light from a hadronic shower. To avoid extensive simulations in order to produce such hadronic shower light tables, the electromagnetic shower ones are used in this track+shower reconstruction procedure.

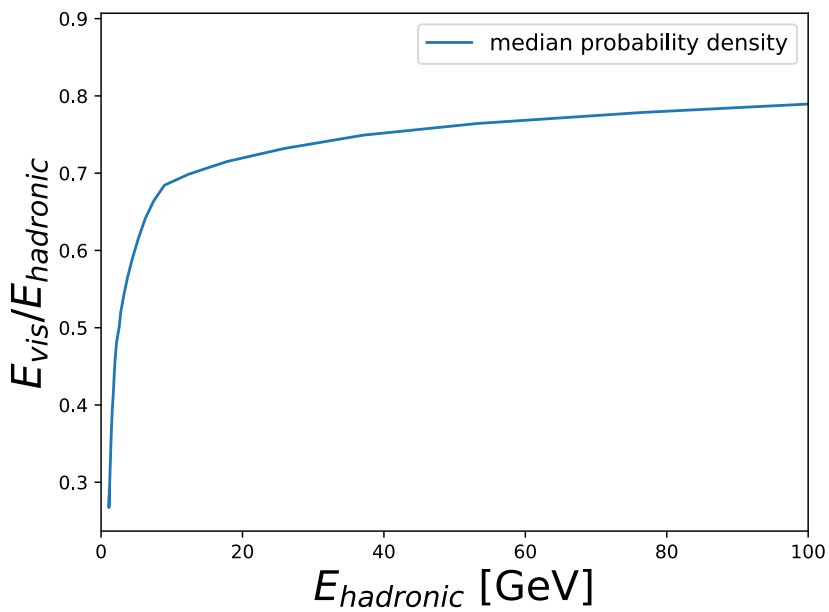


Figure 6.2.: The fraction of the visible shower energy over the hadronic shower energy $E_{vis}/E_{hadronic}$ as a function of the hadronic shower energy. The median of the probability density of this function is shown. The values have been taken from Figure 11.4 in Melis (2021) [69], using the WebPlotDigitizer tool [121].

The use of the photo-electron distribution tables for electromagnetic shower light is a glaring approximation used in reconstructing hadronic showers with this track+shower reconstruction. Making use of tables for hadronic shower light is beyond the scope of this work and remains to be addressed in future studies. In Section 2.3.2, the difference

between the light emitted by electromagnetic showers and the light emitted by hadronic showers is discussed. Figure 6.2 shows the ratio between the hadronic shower energy $E_{hadronic}$ and the visible shower energy E_{vis}, as a function of the hadronic shower energy, as extracted from simulations [69]. As the hadronic energy increases, the hadronic and visible energy become closer in value, as the electromagnetic component of the hadronic shower increases with energy. At low, few-GeV energies, the hadronic and visible shower energy can be quite different, implying a large difference between the true shower energy and the detectable light produced.

The emission of light from the vertex position is an assumption used to define the photo-electron distribution tables for electromagnetic shower light. There is however an offset z_{offset} along the longitudinal z-direction that needs to be accounted for if the shower profile is considered (i.e. the shower is not considered to be a point-like emission of light from the vertex, but it has structure). As depicted in Figure 2.5, the maximum point of shower light emission is not at the vertex, and this point differs between electromagnetic and hadronic showers. The average energy deposition profile for a hadronic shower can be parameterised by a Gamma function of the form:

$$F(E_{hadronic}, z, a, b) = E_{hadronic} \cdot \frac{z^{a-1} \cdot e^{-z/b}}{b^a \cdot \Gamma(a)}\,, \tag{6.1}$$

where z is the distance along the longitudinal direction, Γ is the Gamma function, and the parameters a and b are [69, 122]

$$a = 4.26 + 0.364 \cdot \ln\left(\frac{E_{hadronic}}{[\text{GeV}]}\right),$$
$$b = 0.56\,[\text{m}]\,.$$

To define an expression for the maximum point of hadronic shower light emission, the derivative of Equation (6.1) can be found and equated to zero. Doing so yields

$$z_{offset} = (a-1) \cdot b\,,$$

where z_{offset} is the offset z-position from the vertex.

An expression for the true hadronic energy is then needed, as it appears in the a term of Equation (6.1). The function $E_{vis}/E_{hadronic}$ in Figure 6.2 is parameterised and used to find a parametric relationship between the hadronic shower energy $E_{hadronic}$ and the visible shower energy E_{vis}. This estimated value of $E_{hadronic}$ is used to calculate

the value of z_{offset}. This value is then used when estimating the expected number of photo-electrons from the particle shower in the coordinate system shown in Figure 3.1.

6.2.2. Expression for the Reconstructed Neutrino Direction & Energy

Given the event model described above, the neutrino direction and energy can be estimated from the track and shower directions and energies.

The four-momentum (energy and direction vector) of the hadronic shower $\mathbf{p}_h = (E_h, \vec{p}_h)$ can be determined. Following the convention of Adrián-Martínez et al. (2017) [123]:

$$\mathbf{p}_h = \mathbf{p}_\nu - \mathbf{p}_l \,, \tag{6.2}$$

where $\mathbf{p}_\nu$ is the four-momentum of the original neutrino, and $\mathbf{p}_l$ is the four-momentum of the out-going lepton (in this case a muon). In general, $\mathbf{p}_h$ is a broad definition of whatever particles are produced in the interaction, which can include the momentum of the target nucleus, and sub-nuclear effects [123]. Here, the momentum of the target nucleus is not considered, an approximation used to simplify things. Additionally, hadronic final states with the same $\mathbf{p}_h$ can result in different numbers of particles and particle types, resulting in fluctuations in reconstructed quantities like direction and energy. By excluding the target four-momentum, variations between interactions with the same $\mathbf{p}_h$ will be included in the fluctuations due to different hadronic states [123]. Note that the hadronic shower direction can be defined as $\vec{\theta}_h^{reco} = \vec{p}_h / |\vec{p}_h|$.

The reconstructed neutrino energy E_ν^{reco} and the normalised, reconstructed neutrino direction $\theta_\nu^{\vec{reco}}$ can be expressed respectively as [123]

$$E_\nu^{reco} = E_\mu^{reco} + E_h^{reco} \,, \tag{6.3}$$

$$\vec{\theta}_\nu^{reco} = w_\mu \frac{E_\mu^{reco}}{E_\nu^{reco}} \vec{\theta}_\mu^{reco} + w_h \frac{E_h^{reco}}{E_\nu^{reco}} \vec{\theta}_h^{reco} \,. \tag{6.4}$$

The reconstructed neutrino energy is the sum of the reconstructed muon energy E_μ^{reco} and the reconstructed hadronic shower energy E_h^{reco}. The reconstructed muon direction $\vec{\theta}_\mu^{reco}$ and hadronic shower direction $\vec{\theta}_h^{reco}$ are weighted by their energy divided by the total neutrino energy. In addition, these contributions can have a respective associated weight w_μ and w_h. Decreasing w_h and increasing w_μ reduces the larger fluctuations in the shower direction, which in general leads to a better estimate of the reconstructed

neutrino direction $\vec{\theta}_\nu^{reco}$. The values which may be used for these weights are discussed in Adrián-Martínez et al. (2017) [123]; they will not be studied in depth in this chapter.

6.3. Event Likelihoods

The photo-electron distribution tables described in Chapter 3 readily provide the expected number of photo-electrons for a certain photon arrival time, given a muon track or electromagnetic particle shower. The number of photo-electrons can then be used in calculating the probability of detecting hits from such events, which in turn are used in likelihood functions describing the event. The separate muon track and particle shower reconstruction procedures used in KM3NeT/ORCA maximise such likelihood functions to obtain the best estimates of the model parameters describing the event. The formulation of (separate) muon track and particle shower likelihood functions in this manner are described in Chapter 3.

For the track+shower event model, the maximum likelihood estimation technique is also chosen to find the model parameter best describing this complex event. The likelihood function describing the probability of obtaining hits or no hits from a muon track and hadronic shower event is defined below.

Likelihood Formulation

The first hit, hit PMT, and non-hit PMT likelihood definitions are given in Section 3.7.2. All three likelihoods are combined to form a total likelihood expression for the track+shower event. As outlined in Section 6.3.3 below, all three likelihood components are useful in constraining the model parameters associated with this event.

The total likelihood of recording the first hits, hits and no hits on all the PMTs in the detector for a particular track and shower event is

$$\mathcal{L}(x;\vec{\theta}) = \prod_{i}^{N_{\text{first hit PMTs}}} P_i^{\text{first hit}} \cdot \prod_{i}^{N_{\text{hit PMTs}}} P_i^{\text{hit}} \cdot \prod_{i}^{N_{\text{non-hit PMTs}}} P_i^{\text{non-hit}}, \tag{6.5}$$

where i denotes the i-th PMT under consideration, x is the measurement of a number of photo-electrons, N is the total number of PMTs, and P is the probability of obtaining the hit following the definitions used in Section 3.7.2. As mentioned, these probability definitions depend on x and $\vec{\theta}$, these arguments are omitted from the probability functions

for the sake of brevity. Here $\vec{\theta}$ denotes the parameters to be estimated which describe the track and shower event model, which are

$$\vec{\theta} = [x_0, y_0, z_0, t_0, \theta_\mu, \phi_\mu, \theta_{sh}, \phi_{sh}, E_\mu, E_{sh}] \,.$$

Re-writing this as the negative log-likelihood gives

$$-\log \mathcal{L}(x;\theta) = -\sum_{i}^{N_{\text{hit PMTs}}} \log\left(P_i^{\text{first hit}}\right) - \sum_{i}^{N_{\text{hit PMTs}}} \log\left(P_i^{\text{hit}}\right) - \sum_{i}^{N_{\text{non-hit PMTs}}} \log\left(P_i^{\text{non-hit}}\right) . \tag{6.6}$$

6.3.1. Hit Probability Evaluation

The expected number of photo-electrons are determined for each event, for the track and shower component using the pre-defined photo-electron distribution tables, and for the background. The number of photo-electrons are then used to calculate the first hit, hit PMT and non-hit PMT probabilities. The number of photo-electrons are estimated *separately* for the track and shower since the coordinate system used for the tables is defined using the respective track or shower direction. The expected number of background photo-electrons can be set as a default value of $\sim$ 7 kHz (the expected rate observed in KM3NeT/ORCA, see Chapter 2), or can come from the measured rates in the data, available per PMT. For the track, shower and background, we obtain the expected rate of photo-electrons as a function of time $n(t)$, the combined cumulative number of photo-electrons up to a given time, and the total number of photo-electrons within the time window. These are all the ingredients necessary to determine the value of the likelihood defined in Equation (6.6).

First hit probability calculation

The probability of observing the expected number of photo-electrons for first hits (the earliest hits recorded on the PMT in the time window under consideration) from a track and a particle shower are determined in a similar manner. The hit position and direction are rotated according to the track direction and the shower direction, respectively, and the hit is rotated into the $x-z$ plane. The hit orientation is defined by (ϑ, φ).

For the track, the values of the instantaneous, cumulative and total number of photo-electrons are estimated from the pre-defined muon photo-electron distribution tables for the parameters (R, E_μ^{emission}, ϑ_μ, φ_μ, Δt_μ). E_μ^{emission} is the energy of the muon at the

emission point in the coordinate system of Figure 3.1. The time residual Δt_μ between the hit time t_{hit} and the expected arrival time of light from the muon t_μ^{expected},

$$\Delta t_\mu = t_{\text{hit}} - t_\mu^{\text{expected}}, \tag{6.7}$$

is used, with t_μ^{expected} defined as in Equation (3.32).

For a particle shower, the distance D from the emission point of shower light to the PMT is determined. The time residual Δt_{shower} is estimated from

$$\Delta t_{\text{shower}} = t_{\text{hit}} - t_{\text{shower}}^{\text{expected}}, \tag{6.8}$$

$$t_{\text{shower}}^{\text{expected}} = t_0 + \frac{D \cdot n}{c}, \tag{6.9}$$

where n is the refractive index in seawater. An additional parameter used is the angle δ between the shower direction and PMT position. The energy of the shower at the vertex E_{shower} is used. The values of the instantaneous, cumulative and total number of photo-electrons are then determined for the parameters (D, $\cos\delta$, E_{shower}, $\vartheta_{\text{shower}}$, φ_{shower}, Δt_{shower}) from the pre-defined tables for the photo-electron distribution tables for electromagnetic shower light.

Hit and non-hit probability calculation

For the hit and non-hit probabilities, no timing information is included, so the time residuals defined above for the muon and shower are not used. The total number of photo-electrons within the considered time window is estimated from the muon photo-electron distribution tables for the parameters (R, E_μ^{emission}, ϑ_μ, φ_μ), and from the electromagnetic shower photo-electron distribution tables for the parameters (D, $\cos\delta$, E_{shower}, $\vartheta_{\text{shower}}$, φ_{shower}). The probability values can then be found following the expressions in Section 3.7.2.

6.3.2. Combined Photo-Electron Distribution Example

An example is shown in Figure 6.3 of the photo-electron arrival time distribution when combining the expected number of photo-electrons from a track+shower event model. The number of photo-electrons from background sources are not included here. The parameters used to define this event and the PMT position relative to the event vertex are provided. For this illustration, the time residual $\Delta t = t_{hit} - t_{expected}$ is chosen such

that 0 corresponds to direct Cherenkov light from the muon. A time window of [-10, +500] nanoseconds is chosen around this time residual. The expected number of photo-electrons is calculated separately for both the track and shower as outlined in Section 6.3.1. For both cases, the same PMT orientations as described in Section 3.2 are used, with 'West' and 'East' indicating that the PMT faces towards or away from the track/shower direction, respectively. The separate contributions from muon and shower light are visible in this example. The energy of both the track and shower is 10 GeV, and there is an angle of $\sim 29.5^{\circ}$ between the track and shower directions in this example.

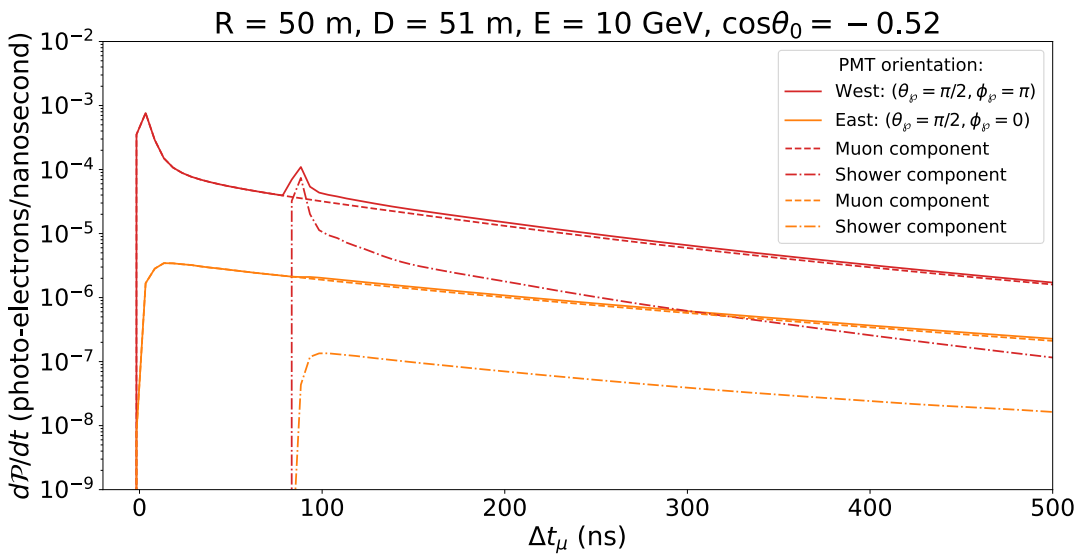


Figure 6.3.: An example of the combined electromagnetic shower and muon photo-electron distributions as a function of the photon arrival time is shown here for a track and shower event model, relative to a PMT. These distributions are shown for two different PMT orientations.

6.3.3. Event Likelihood Landscapes

In order to test whether the event likelihood of Equation (6.6) can be used to estimate the model parameters describing a track+shower event, the value of the likelihood for a range of the model parameter values is investigated for simulated charged-current muon neutrino (ν_μ-CC) events in the complete ORCA-115 line detector. These likelihood 'landscapes' should indicate a maximum likelihood value - equivalently, a minimum negative log-likelihood value - at the true value of the model parameters for these events.

A set of ν_μ-CC events in the range of 10–100 GeV are chosen to study the event likelihood and event model. These events are propagated through the generation, light and trigger simulation stages described in Chapter 3; the true energy distributions of the muons and showers for these events are shown in Figure 6.4. The majority of muons have energies from about 5–30 GeV, while the showers tend to have less energy, with most showers having energies of 1–10 GeV. A sample of events are then selected in order to study the behaviour of the event likelihood. The criteria chosen for these events are that there is at least one muon or shower particle (photon, electron, hadron) per event, and that each event results in at least 200 hits in the detector. Only events with a vertex that lies within the instrumented volume of the detector are considered. The events which meet these criteria are used to perform the likelihood scans described in this section.

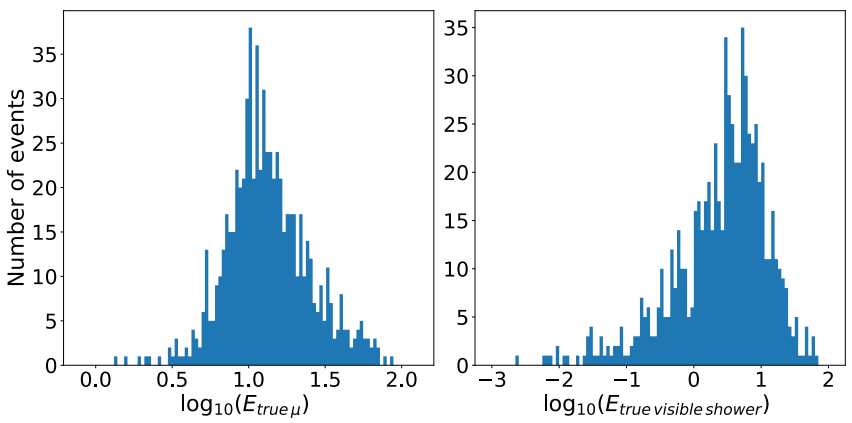


Figure 6.4.: The true energy of the muon and showers used to perform the likelihood scan, as described in the text. Some showers have a visible energy less than 1 GeV.

The impact of the likelihood components of Equation (6.6) - the first hit likelihood, the hit PMT likelihood and the non-hit PMT likelihood - in evaluating the model parameters is investigated here. The value of the negative log-likelihood $-\log\mathcal{L}$ is calculated for a range of values of the model parameters — the vertex positions x, y, z, the track and shower directions (θ_μ, ϕ_μ) and (θ_{sh}, ϕ_{sh}), respectively, and the track and shower respective energies E_μ and E_{sh}. For each parameter, the parameter value is varied across a range of values and $-\log\mathcal{L}$ is calculated for each value. A time window of [-1000, +1900] ns (relative to the expected arrival time of muonic Cherenkov radiation) is used to evaluate the expected number of photo-electrons for each event; these wide time ranges are used to ensure the track and shower signal is found for each event (i.e. ensuring the maximum of the photo-electron distributions can be found for the track and shower). The background signal due to potassium-40 decays and thermal PMT noise is simply set to the average expected value of 7 kHz.

For the selected ν_μ-CC events, $-\log\mathcal{L}$ is calculated for each event, varying around the true value of one parameter while keeping the other parameters fixed at their true values. The minimum value of the likelihood for all events is subtracted from the calculated likelihood value, this quantity is denoted by $\Delta\log\mathcal{L}$ (to give a minimum value close to zero). The median value of $\Delta\log\mathcal{L}$ in this parameter space is then illustrated for the sample of events.

Figure 6.5 shows the results of these likelihood scans for different values of the vertex-z position, muon direction, and muon energy. On the left are the results for the first hit likelihood, the first hit likelihood + hit PMT likelihood, and the total $-\log\mathcal{L}$. The figures on the right show $\Delta\log\mathcal{L}$ for the track+shower model compared to the case where the presence of a muon track only is assumed (i.e. only the expected number of photo-electrons from a muon and background are considered, the presence of a shower is not considered).

When varying the vertex z-position and the muon direction, each additional likelihood component offers extra constraint to the parameter, resulting in a more clearly defined minimum value of the likelihood. The minimum of $\Delta\log\mathcal{L}$ is located at or close to the true model parameter. When varying the three model parameters, the track+shower likelihood is very similar to that of the likelihood from a muon-only model. This would suggest a similar performance in estimating these track parameters with either event model). When varying the muon energy, the total likelihood is necessary in determining a minimum at or close to the true parameter value, shown in Figure 6.5 (e) where the total likelihood gives a more constraining estimate of the parameter compared to the other likelihood formulations.

The likelihood scans of the shower model parameters result in many events for which the likelihood landscape is completely flat. This implies little or no sensitivity to the parameters since a minimum is difficult to find for these parameters. The majority of events showed this flat behaviour of $\Delta\log\mathcal{L}$ and hence the median over events is not illustrated for the likelihood scan over shower model parameters. Instead, Figure 6.6 shows some example events for which the minimum value of $\Delta\log\mathcal{L}$ are close to or far from the true shower energy and true shower direction. The visible energy of the hadronic shower and the total number of hits from the shower (i.e. the hits due to all hadronic particles, electrons and photons in the event) are indicated in the figures. The minimum value of $\Delta\log\mathcal{L}$ is also shown. These likelihood landscapes suggest the difficulty in reconstructing the shower direction and energy alongside the track. The likelihood values do not change much across the entire range, compared to those of the muon depicted in Figure 6.5.

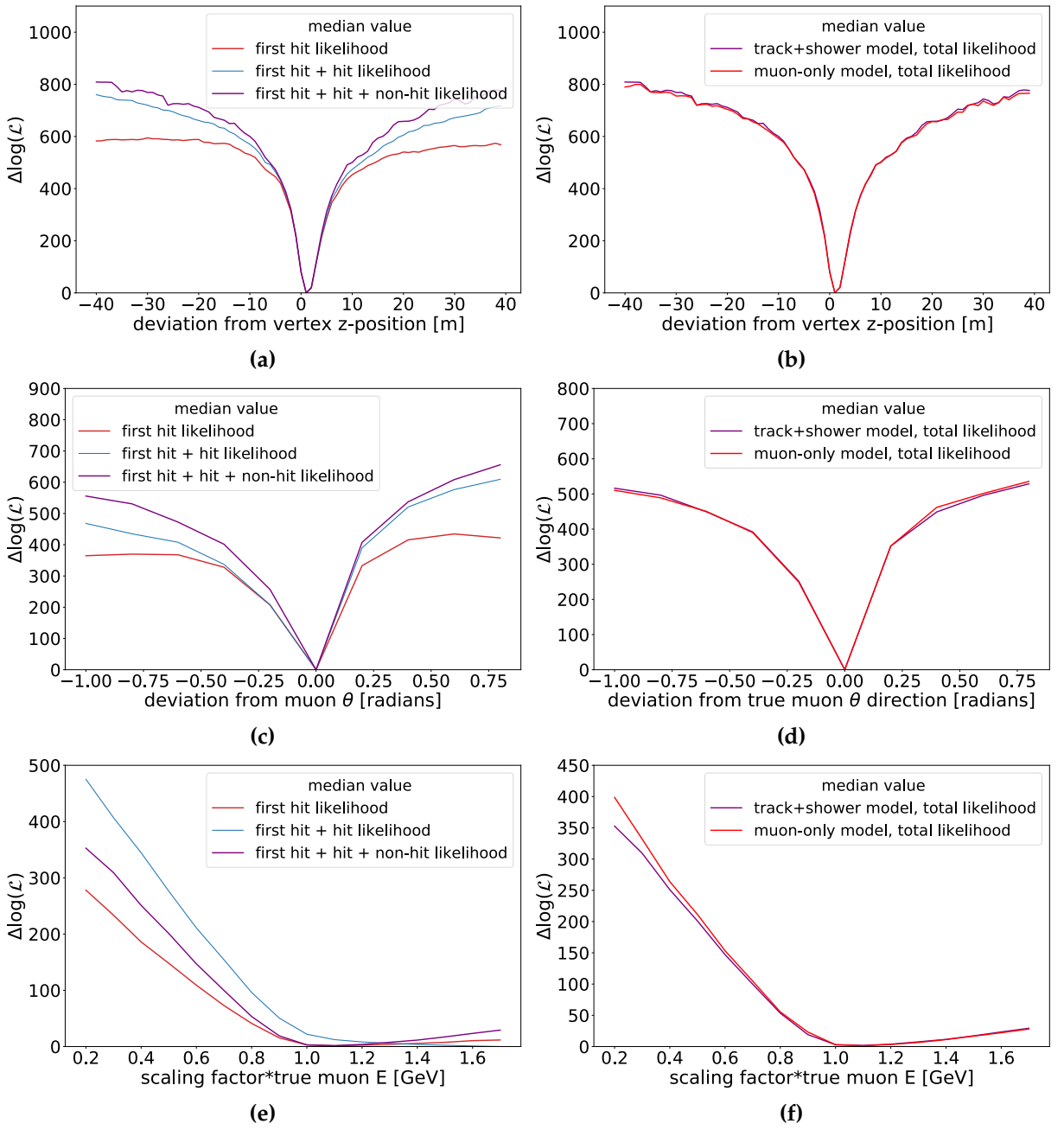


Figure 6.5.: The likelihood landscape for an example set of ν_μ-CC events, as a function of the track+shower model parameter values. The median value of $\Delta \log \mathcal{L}$ is used to illustrate the likelihood, with a minimum value of zero. On the left, the figures show the cases for the likelihood landscape with various likelihood formulations as function of deviations from the true model parameter values. On the right, the track and shower total $\Delta \log \mathcal{L}$ is shown compared to a muon-only event model.

These likelihood scans suggest that the event likelihood and model can be used to estimate the vertex position and muon direction and energy, though estimating the shower direction and energy will be quite difficult. The majority of the hadronic showers in the sample of selected events have an energy from 0.1–10 GeV as displayed

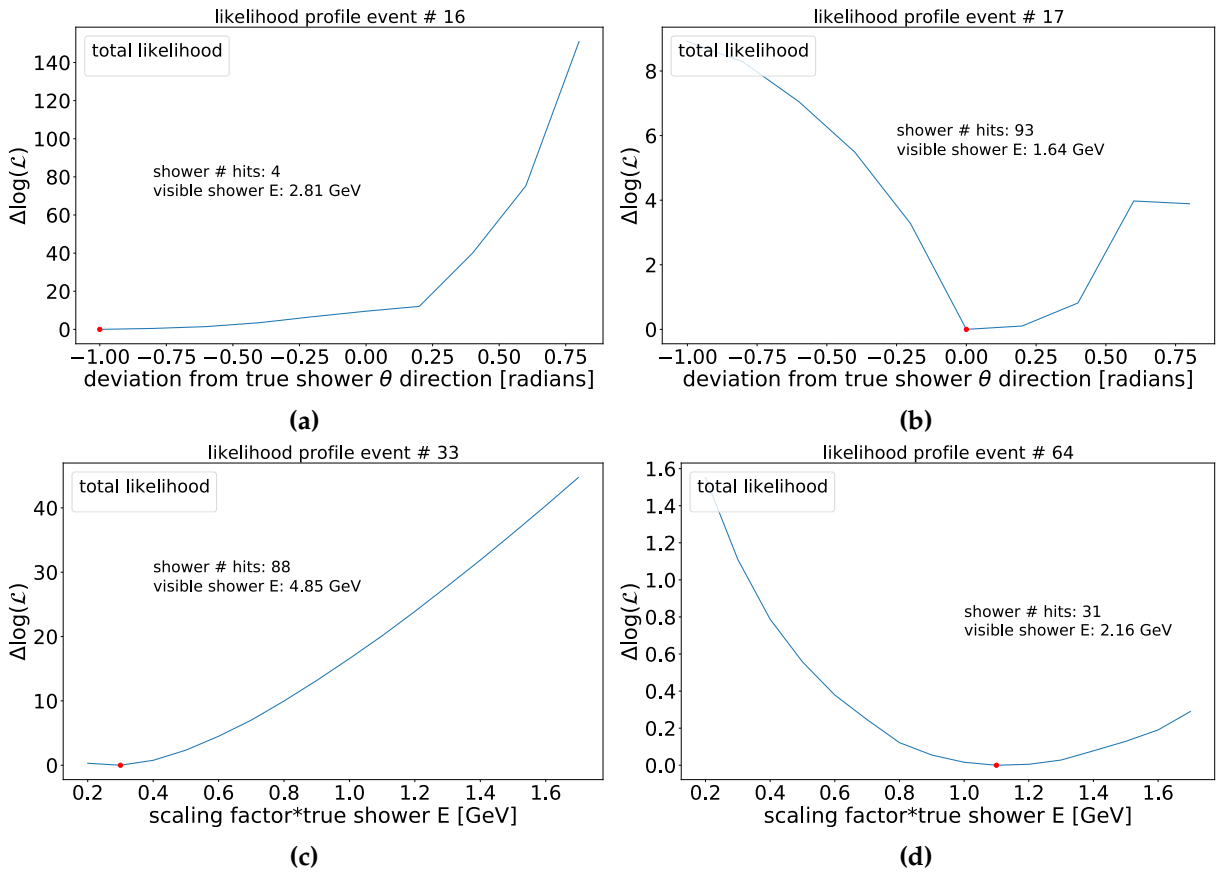


Figure 6.6.: The likelihood landscape for an example set of ν_μ-CC events, as a function of the track+shower parameter values, namely the shower direction and energy. (a) and (b) show deviations from the true shower direction, showing a minimum far and at the true value, respectively. (c) and (d) show scans of the likelihood when varying the true shower energy, again for a minimum likelihood far from and close to the true value, respectively. These figures illustrate the difficulty in fitting a hadronic shower component in the selected neutrino events.

in Figure 6.4. The lower the shower energy, the fewer hits should be recorded in the detector. These hits should also be spread out in the detector across different PMTs in order for the event (and in particular the shower component) to be reconstruct-able. Not all events will be reconstructed well, especially those with few to no hits, though the events which find a minimum of Equation (6.6) near the true value of the model parameter can be identified. This will be discussed further in Section 6.5.

6.4. Technical Implementation

6.4.1. Choice of Time Window

In the standard track or shower reconstruction algorithms used for KM3NeT/ORCA, a time window around the expected arrival time of Cherenkov light from a muon track or shower is used. In order to avoid the use of a time window relative to the photo-electron distributions (in which the photo-electrons from the track and shower arrive at the PMT at unknown times relative to each other), the full event time is used as the range in which to evaluate the hits in each event. This range is in fact the first and last hit in the timeslice defining the event time window, as described in Section 2.5.

6.4.2. Minimisation Procedure

The choice was made to use MINUIT2 [124], a library of algorithms used in minimising a user-defined function. With this library widely used in particle physics, MINUIT2 was the minimiser of choice for the track and shower likelihood definition of Equation (6.6).

The model parameters are not fitted all at once. This would result in too many degrees of freedom in which the function should be minimised. Instead, the parameters are fitted in a sequence in order to aid MINUIT2 in determining the function minimum. The parameters are fitted to the data in the following order (those that are not fitted are kept fixed):

1. The vertex parameters (x_0, y_0, z_0, t_0), by minimising the first hit negative log-likelihood only. This is done to give a quick estimate of these parameters.
2. The vertex parameters and the muon direction parameters (θ_μ, ϕ_μ) are fitted by minimising the full likelihood expression of Equation (6.6).
3. The vertex parameters and the shower direction parameters (θ_{sh}, ϕ_{sh}) are fitted by minimising the full likelihood expression of Equation (6.6).
4. All model parameters, including the muon and shower energies (E_μ, E_{sh}), are fitted by minimising Equation (6.6).

The ordering of the fit procedure described here resulted in the optimal performance of the algorithm, yielding the best track and shower angular and energy resolutions.

6.4.3. Starting Values

The starting value problem (mentioned in Section 3.7.1) is an issue for this track and shower model fit to data — where does one begin the fit procedure?

As mentioned at the beginning of this chapter, both the standard track and shower algorithms are applied to KM3NeT/ORCA data, the results of which are then used for physics analyses. In the following chapter, this track+shower reconstruction is applied to data from the KM3NeT/ORCA-6 detector. The events are reconstructed twice with this algorithm, once using the standard track reconstruction procedure as a prefit and once using the standard shower reconstruction procedure as a prefit. The results are compared between the two cases. A detailed study of the reconstruction performance for many neutrino interaction types using both prefits can thus be found in Chapter 7.

A case study is presented of the performance of the reconstruction in the following section, for simulated data, using only *one* of the standard reconstruction algorithms to provide starting points to the track+shower reconstruction. For ν_μ-CC events, the track reconstruction algorithm is used as a prefit for the track+shower fit parameters. For ν_e-CC events, the electromagnetic shower reconstruction algorithm is used as a prefit for these events, for both track and shower fit parameters. The results from the track+shower reconstruction algorithm are compared to those from the prefits used.

6.5. Performance of the Track & Shower Reconstruction Algorithm

The performance of the track+shower reconstruction is described in this section. The reconstruction is evaluated for simulated ν_μ-CC and ν_e-CC events in the KM3NeT/ORCA 115-line detector, for neutrino energies ranging from 1–100 GeV. These events are simulated via the generation, light propagation and detection, and trigger algorithms described in Chapter 3. There are many events for which the low-energy particle showers induce few or no hits in the detector. It is thus important to focus the evaluation of the reconstruction on events which can be reconstructed well. A selection of events is made in the following sections; it is important to make this selection using reconstructed variables and not Monte Carlo truth information, since the latter is not available in the recorded data.

6.5.1. Charged-Current Muon Neutrino Events

The true energy distributions for the tracks and showers associated with these simulated ν_μ-CC events are presented in Figure 6.7. The angular and energy resolutions presented in this section reflect the events with true energies distributed as shown.

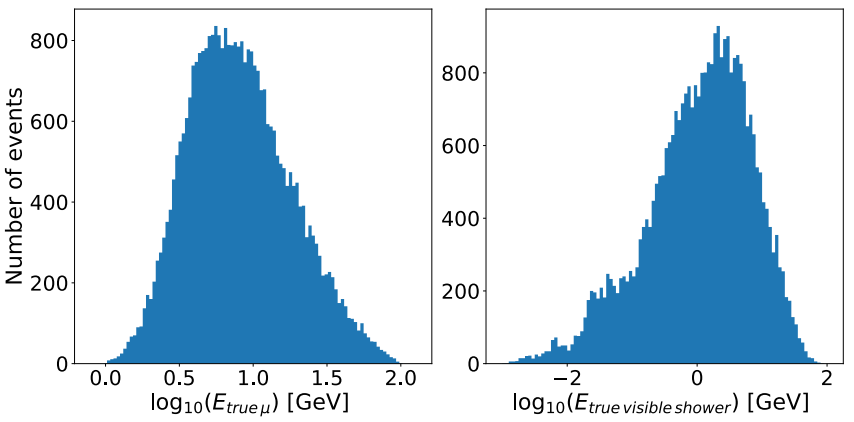


Figure 6.7.: The true energy of the muon and showers for the ν_μ-CC events used to evaluate the track+shower reconstruction performance.

In order to eliminate the events with zero hits and/or which are reconstructed quite poorly, reconstructed variables are used as selection criteria ('cuts'). The selected events have a reconstructed muon energy from 1–80 GeV, a reconstructed shower energy from 0.3–80 GeV, and a hit PMT likelihood value from 1720–2200. Figure 6.8 shows the difference in the true and reconstructed angle and energy of the muon as a function of the hit PMT likelihood. The chosen range removes the events at low values of the likelihood with large differences between the true and reconstructed track direction and energy.

Although the reconstruction algorithm is applied on all events for which there is a prefit, only the events which comply with the cuts are shown in this section. The effect of these cuts on the events and their resolution is presented in Table 6.1. These are merely an example of some cuts that can be applied to these neutrino events in order to evaluate the reconstruction performance, they are not definitive and are used here to convey the ability of the reconstruction to estimate the track+shower model parameters. More rigorous selection criteria are applied in physics analyses of KM3NeT/ORCA data, e.g. in Nauta (2021) [85], Carretero (2023) [91]. The cuts noted in Table 6.1 improve the angular and energy resolution by reducing the number of events with a large angular difference between the true and reconstructed direction of the track and shower, and by reducing the standard deviation (the width) of the distributions of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$.

These quantities have been mentioned before in Section 3.7.6 when quantifying the angular and energy resolution from a reconstruction algorithm.

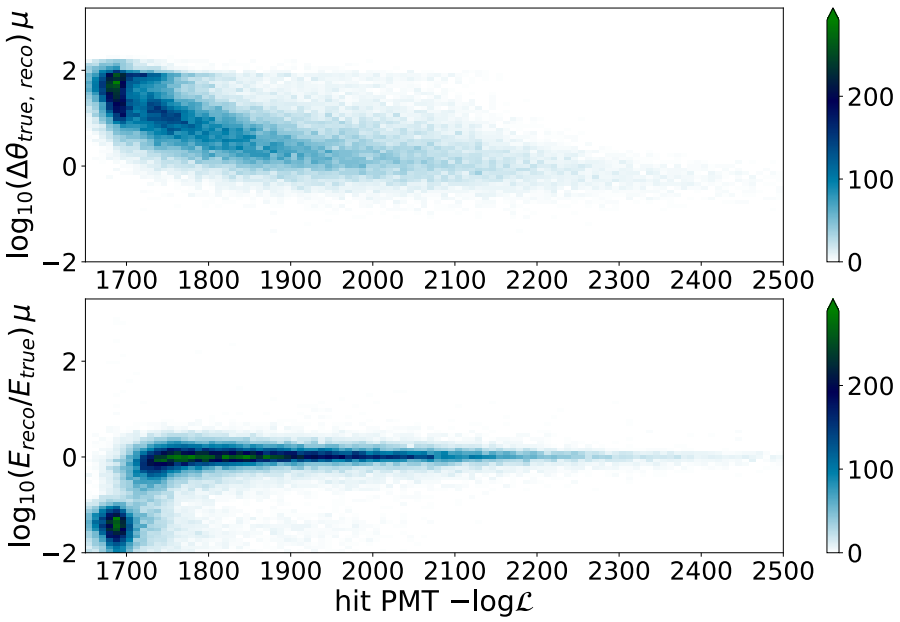


Figure 6.8.: The angular difference $\Delta\theta_{\text{true}\,\mu,\,\text{reco}\,\mu}$ and $\log_{10}(E_{\text{reco}}/E_{\text{true}}\mu)$ are shown as a function the hit PMT negative log-likelihood. A range of $-\log\mathcal{L}$ from 1720–2200 is selected to remove some poorly reconstructed events. The colour-bar indicates the number of events.

Quantity	Before Cuts	After Cuts
Number of events with fit result	74,787	33,664
Median $\Delta\theta_{\text{true}\,\mu,\,\text{reco}\,\mu}$	6.7°	3.4°
Median $\Delta\theta_{\text{true}\,sh,\,\text{reco}\,sh}$	41.7°	35.7°
$\log_{10}(E_{\text{reco}}/E_{\text{true}})$ muon (standard deviation)	0.8	0.2
$\log_{10}(E_{\text{reco}}/E_{\text{true}})$ shower (standard deviation)	1.4	0.8

Table 6.1.: The effect of some example cuts, as described in the text, used for the charged-current muon neutrino events on which the track+shower reconstruction performance is evaluated.

Vertex reconstruction

The distance between the true neutrino vertex position and the reconstructed vertex position for both the track+shower reconstruction and the standard track reconstruction is shown in Figure 6.9. Both reconstruction procedures result in similar estimates of the

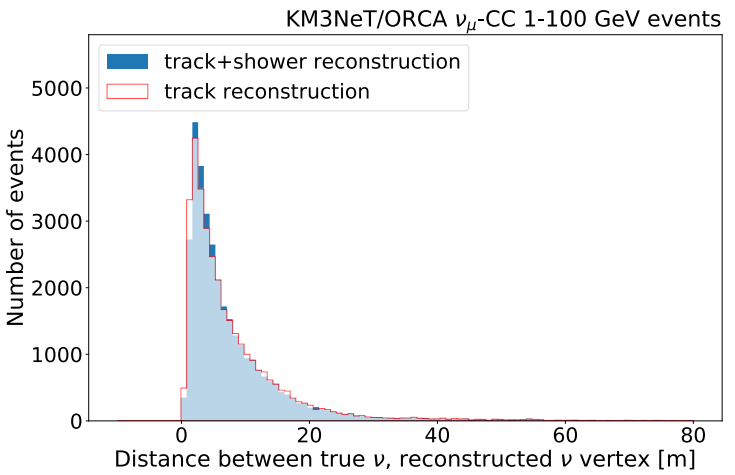


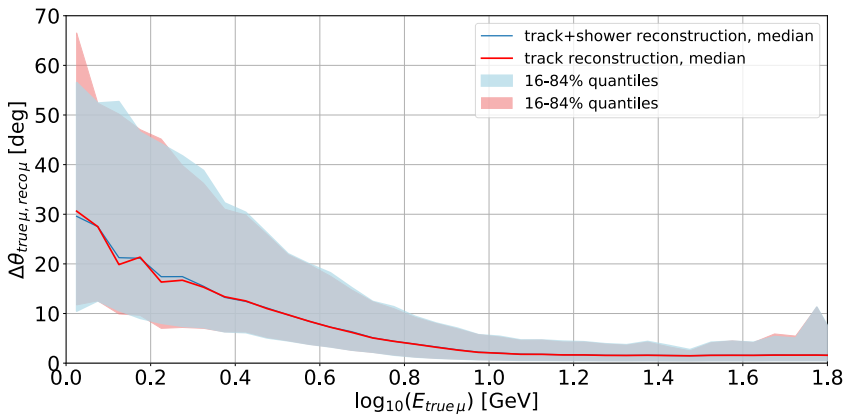
Figure 6.9.: The distance between the true neutrino vertex and reconstructed vertex position, for the track and shower reconstruction (blue histogram) and track reconstruction (red). With the same number of events in each distribution, the track+shower reconstruction succeeds in reconstructing more events < 10 m from the true vertex position.

vertex position, most importantly with many events being reconstructed very close to the true position.

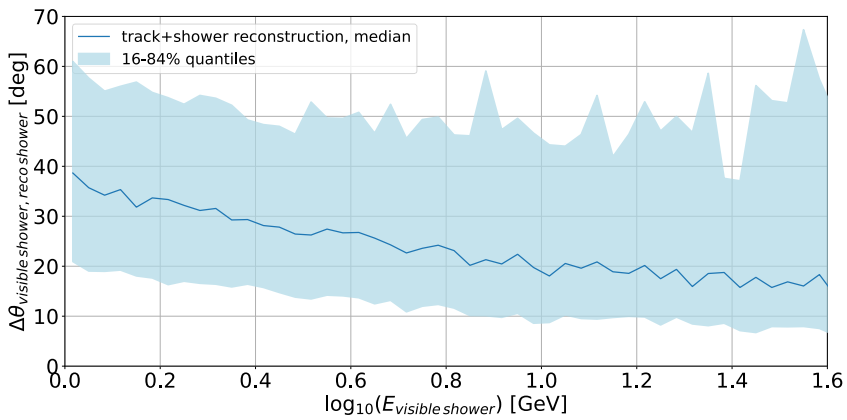
Direction reconstruction

The angular resolution (as described in Section 3.7.6) as a function of the true muon and shower energy respectively is determined to quantify how well the directions of the track and shower components are reconstructed. The angle between the true and reconstructed muon $\Delta\theta_{\text{true }\mu,\,\text{reco }\mu}$ is used to show the angular resolution of the muon direction as a function of the muon energy. The angle $\Delta\theta_{\text{visible shower, reco shower}}$ between the true, visible shower direction and the reconstructed shower direction is estimated and used to convey the angular resolution of the shower direction. The angular resolution for the muon and shower case are shown in Figure 6.10. In addition to the median of the distributions (the 50% quantile), the 16% and 84% quantiles are indicated.

For the track direction, the angular resolution from the track+shower reconstruction is very similar to that from the standard track reconstruction, shown in Figure 6.10 (a). The reconstruction succeeds in determining the muon direction, likely aided by the starting value from the track direction (which may already be close to the intrinsic limit for these events). The 16% – 84% quantiles are slightly more narrow in the case of the track reconstruction. This is expected as with another component to reconstruct, the muon direction in the track+shower case has a larger degree of freedom to vary during



(a)



(b)

Figure 6.10.: The angular resolution for the (a) muon component and (b) shower component, as estimated by the track and shower reconstruction. For both cases, the median and 16%–84% quantiles of the distributions are indicated, for the angle between the true and reconstructed component direction as a function of the true energy. In the case of the muon direction, the angular resolution from the track+shower reconstruction is compared to the result from the track reconstruction.

the fit procedure. Both reconstructions show an angular resolution of less than 2° above 20 GeV.

For the hadronic shower direction, the median angular resolution falls below 20° above 10 GeV. This is in agreement with the results from the study of the intrinsic limits on the shower resolution in KM3NeT/ORCA [123]. Due to limited statistics, the angular resolution is not shown beyond a shower energy of 40 GeV.

Given that the direction of both the muon and shower is reconstructed, an attempt is made to reconstruct the neutrino direction itself, using Equation (6.4). Naively, the values of the weights w are set to unity. The angular deviation between the true and reconstructed neutrino direction is calculated, and compared to the angular deviation

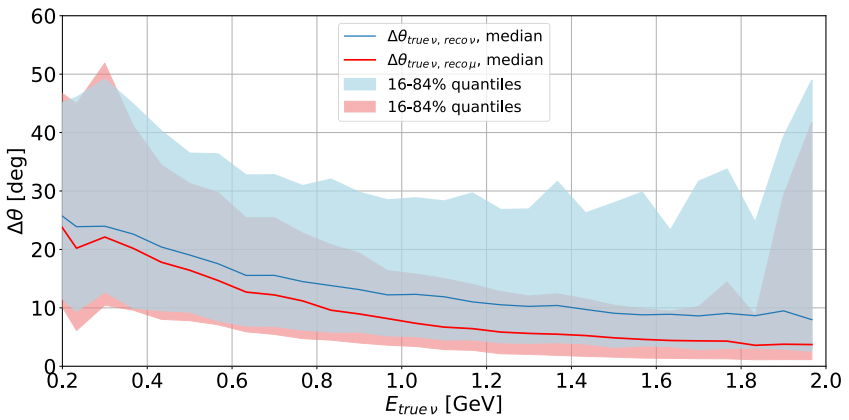


Figure 6.11.: The angular deviation between the true and reconstructed neutrino direction and the angular deviation between the true neutrino and reconstructed muon direction is presented as a function of the neutrino energy, in terms of the median and quantiles of the distributions. The reconstructed directions of the neutrino and muon are those evaluated from the results of the track+shower reconstruction.

between the true neutrino and reconstructed muon direction, as a function of the neutrino energy. The median of both of these distributions is presented in Figure 6.11, alongside the associated 16%–84% quantiles. Although the results are similar, it is apparent from this formulation of $\vec{\theta}_{\nu}^{reco}$ that the reconstructed muon direction acts as a better proxy to the true neutrino direction than the reconstructed neutrino direction. This is as result of the determination of the hadronic shower direction, a difficult quantity to determine, which has a larger associated error compared to that of the reconstructed muon direction. The reconstructed direction errors are not accounted for in the definition of the above neutrino direction. A study of the effect of the values of the weights in Equation (6.4) on the reconstructed neutrino angular resolution is not included in this work.

Energy reconstruction

The reconstructed muon energy is presented in Figure 6.12 in the form of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ and the reconstructed muon energy as a function of the true muon energy. The quantity $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ is compared between the track+shower reconstruction and the standard track reconstruction. Both results are quite similar. The width of the distributions can indicate the resolution to the energy; it can be quantified using the standard deviation σ. For the track reconstruction and track+shower reconstruction, $\sigma = 0.23 \pm 0.0009$ and 0.22 ± 0.0008, respectively. This indicates a slightly better resolution from the track+shower reconstruction. Figure 6.12 (b) shows that many events from the track+shower reconstruction are reconstructed with an energy close to that

of the true muon energy, as indicated by the diagonal line. The equivalent plots are presented for the reconstructed shower energy. The quantity $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ for the shower energies is shown in Figure 6.12 (c). The standard deviation of this distribution is 0.77. The majority of events are reconstructed with an energy close to that of the true visible energy, with a spread of events with under- and over-estimated energies. This is reflected in Figure 6.12 (d), where the reconstruction shows sensitivity to reconstructing the shower energy above $\simeq 3$ GeV, and the events are distributed around the diagonal line where $E_{\text{reco}} = E_{\text{true}}$.

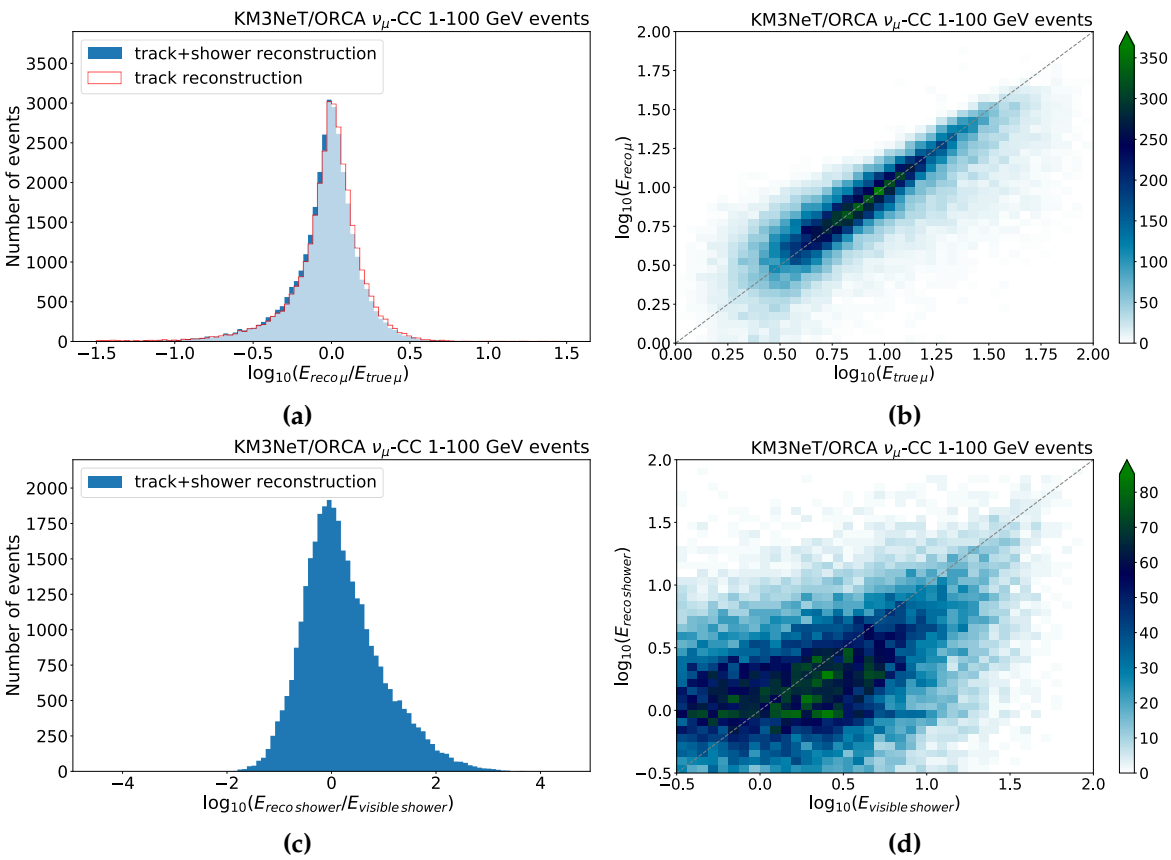


Figure 6.12.: (a) The distribution of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ for the muon components of the track+shower reconstruction (blue), compared to that of the track reconstruction (red). (b) The reconstructed muon energy from the track+shower reconstruction, as a function of the true muon energy. (c) The quantity $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ for the shower component of the track+shower reconstruction. (d) The reconstructed shower energy from the track+shower reconstruction, as a function of the true, visible shower energy.

The quantity $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ can be presented as a function of the energy, as is illustrated in Figure 6.13 (a) for the reconstructed muon energy and Figure 6.13 (b) for the reconstructed shower energy. The behaviour of the energy resolution and bias for the track energy is described in Section 5.1.3. The standard deviation σ of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ for the muon and shower case are presented in Figure 6.13 (c) and (d) respectively. At $E_{\text{true}\,\mu} = 10$ GeV, $\sigma \simeq 0.16$ for both the track reconstruction and track+shower reconstruction. Beyond 10 GeV both reconstruction procedures continue to yield a very similar value of σ. Below 10 GeV, the resolution of the track energy is better from the track+shower reconstruction than that from the track reconstruction. In the case of the reconstructed shower energy, at $E_{\text{true shower}} = 10$ GeV, $\sigma \simeq 0.44$.

Finally, the reconstructed neutrino energy E_ν^{reco} of Equation (6.3) is presented in Figure 6.14 (a) as a function of the true neutrino energy. Figure 6.14 (b) shows the reconstructed muon energy from the track reconstruction (the prefit energy) as a function of the true neutrino energy. Using both the muon and shower energy components appears to improve upon the estimate of the true neutrino energy, compared to using the muon energy estimate alone as a proxy for the neutrino energy, with more events appearing along the diagonal line. Figure 6.15 (a) shows the median and 16 – 84% quantiles of $\log_{10}(E_{\text{reco}}/E_{\text{true}\,\nu})$ as a function of the true neutrino energy, while the width σ of this quantity is shown in Figure 6.15 (b). In both figures, $\log_{10}(E_{\text{reco}}/E_{\text{true}\,\nu})$ is presented for the case of the reconstructed track energy $E_{\text{reco}\,\mu}$ (in red, from the track+shower reconstruction) and the reconstructed neutrino energy $E_{\text{reco}\,\nu}$ (in blue, from the track+shower reconstruction), which is calculated using the addition of the reconstructed track and shower energies as given in Equation (6.3). Figure 6.15 (a) shows a similar behaviour of both distributions as a function of the neutrino energy, with a different bias, coming from the additional shower energy used to define $E_{\text{reco}\,\nu}$. In Figure 6.15 (b), the value of σ for $E_{\text{reco}\,\nu}$ is lower across almost the entire true neutrino energy range compared to that of $E_{\text{reco}\,\mu}$. At $E_{\text{true}\,\nu} = 10$ GeV, $\sigma \simeq 0.145$ and 0.161 when using the reconstructed neutrino energy and reconstructed muon energy, respectively. The energy resolution is better when using the reconstructed neutrino energy from this track+shower reconstruction, rather than the muon energy alone as a proxy for the true neutrino energy. Such a result justifies the use of this novel track+shower reconstruction algorithm.

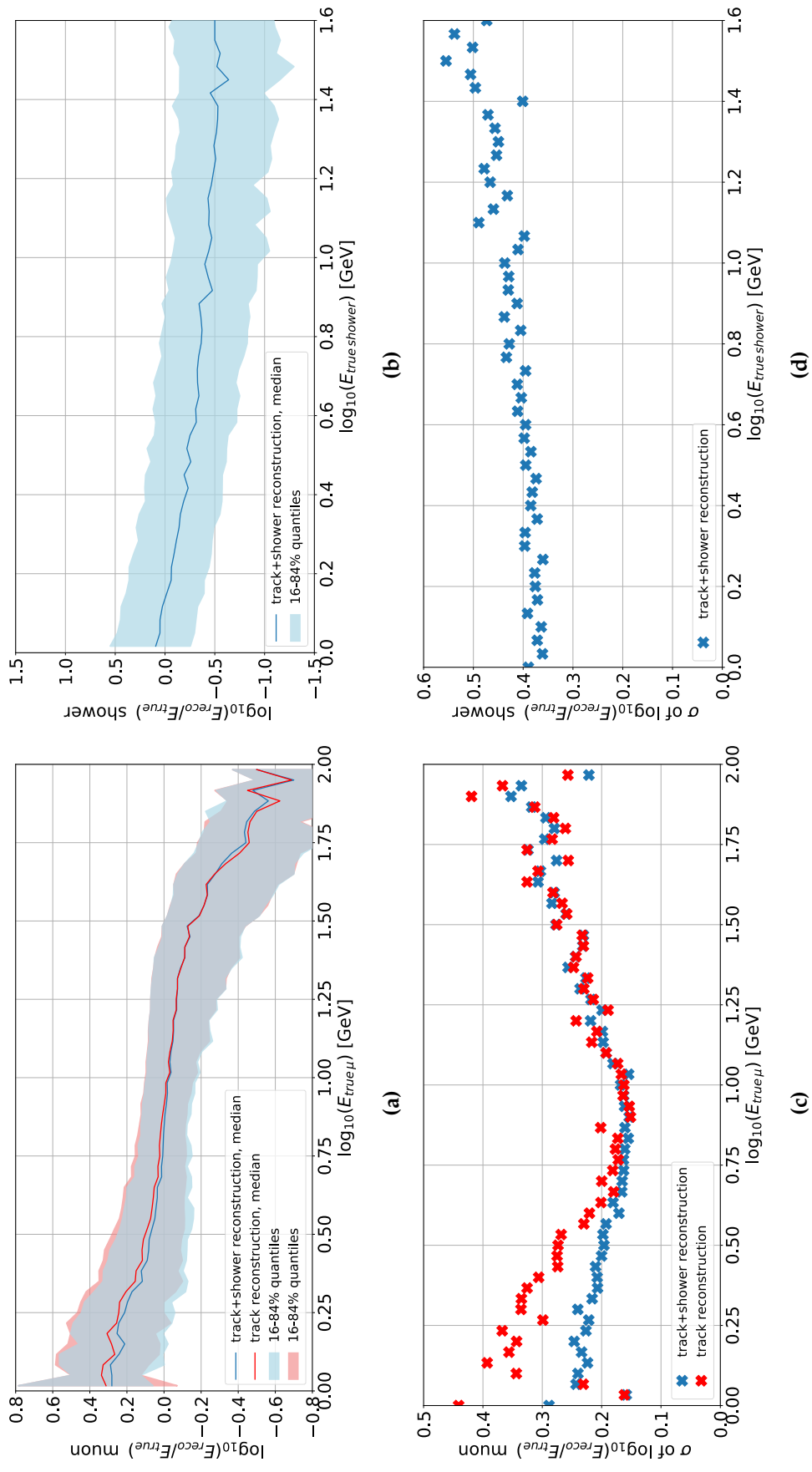
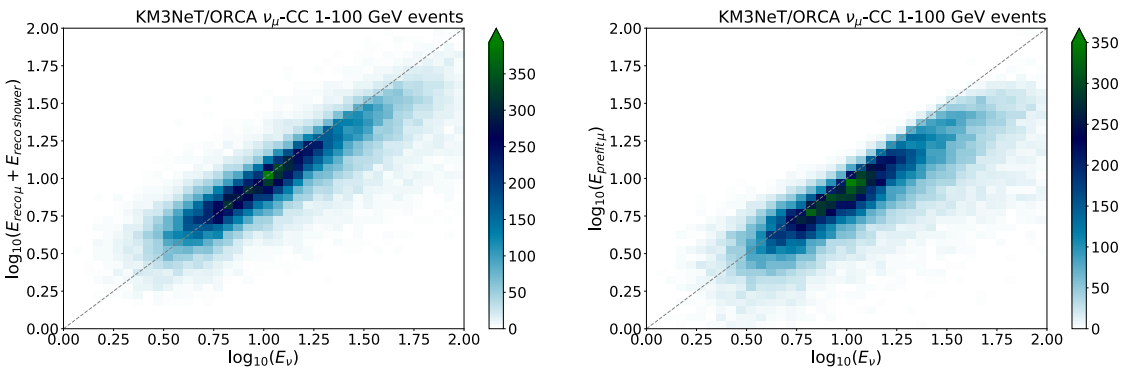
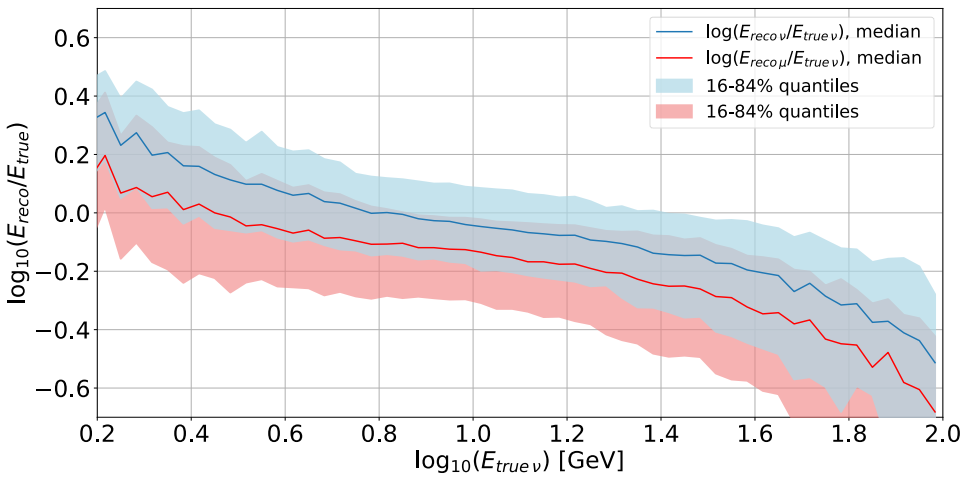


Figure 6.13.: (a) The distribution of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ for the muon component of the track+shower reconstruction (blue), compared to that of the track reconstruction (red). (b) The reconstructed shower energy from the track+shower reconstruction, as a function of the true shower energy. (c) and (d) show the width σ of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ for the distributions in (a) and (b), respectively.

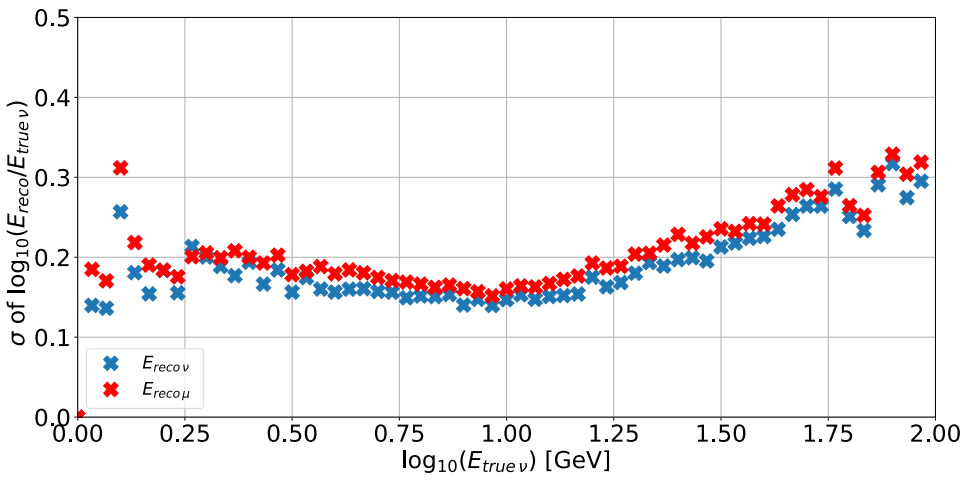


(a) **(b)**

Figure 6.14.: The true neutrino energy is shown in (a) vs. the reconstructed neutrino energy, and in (b) vs. the prefit track energy. The colour-bar indicates the number of events. The diagonal line indicates the region where the true and reconstructed energies are equal.



(a)



(b)

Figure 6.15.: (a) The median and quantiles of the distribution of $\log_{10}(E_{\text{reco}}/E_{\text{true}\,\nu})$ as a function of the true neutrino energy, using the reconstructed neutrino energy (blue) and the reconstructed muon energy (red), from the track+shower reconstruction. (b) The width of the distribution shown in (a) is shown (i.e. the energy resolution) when using the reconstructed neutrino energy (blue) and the reconstructed muon energy (red) as a proxy for the true neutrino energy.

Opening Angle & Bjorken-y

The opening angle between the muon and shower direction $\Delta\theta_{\mu,\,\text{shower}}$ is shown in Figure 6.16, for the true and reconstructed directions respectively, as a function of the neutrino energy. The median and 16%–84% quantiles of both distributions are presented. For the true opening angle, the difference between both directions decreases with energy, as expected, as with higher energies events the muon and shower components should become more co-linear along the neutrino direction axis.

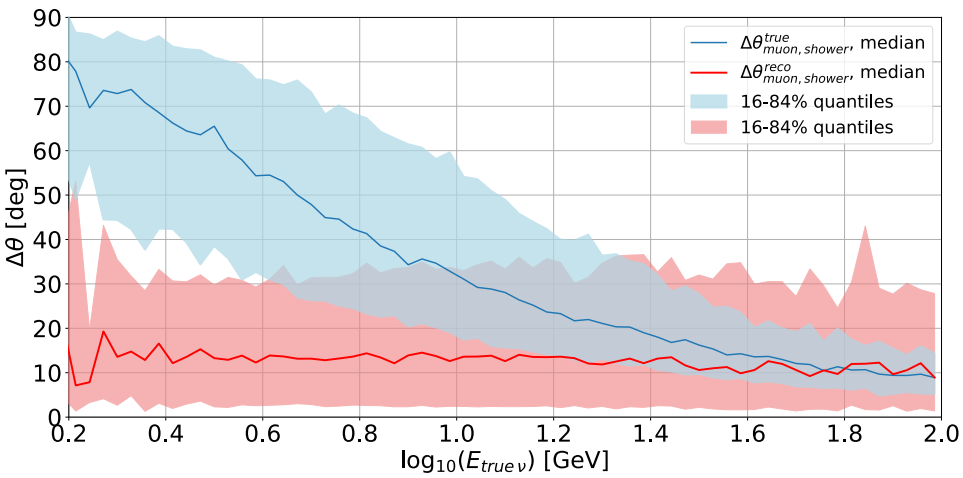
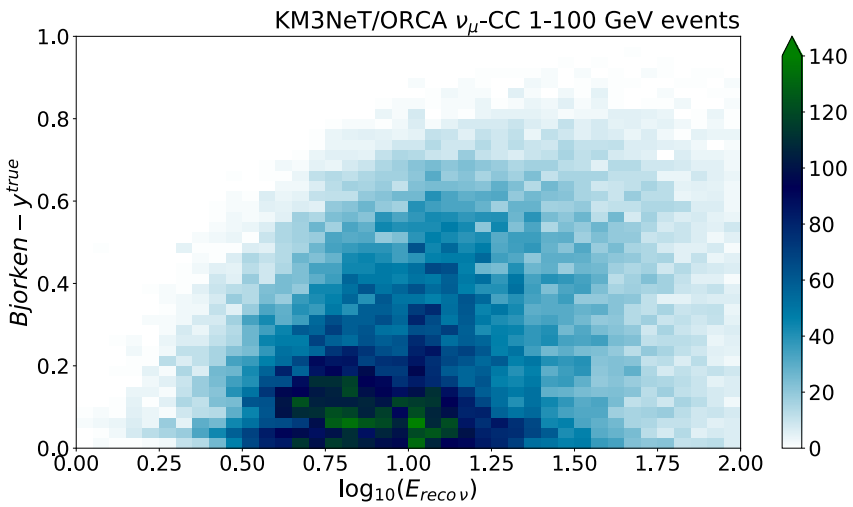


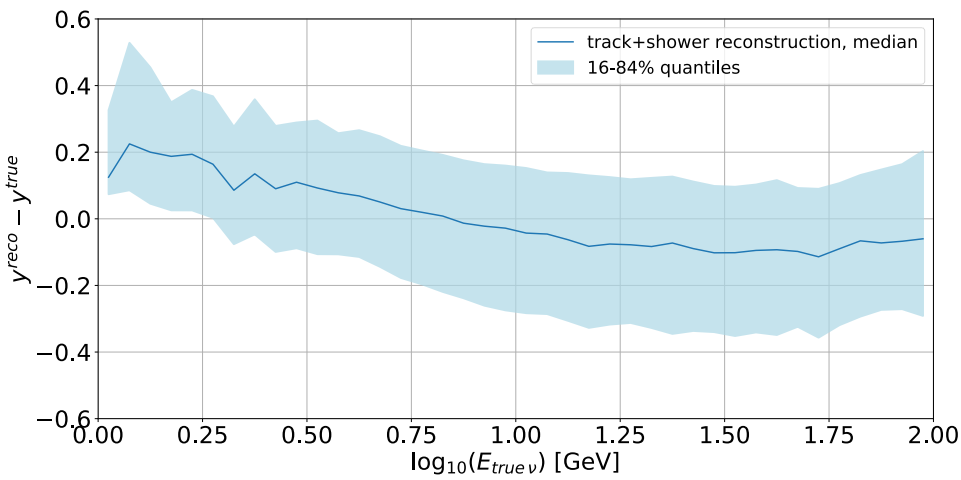
Figure 6.16.: The median and 16%–84% quantiles of the distributions for the true (blue) and reconstructed (red) opening angle between the muon and shower directions, are shown as a function of the neutrino energy. The muon and shower are reconstructed within 15° of each other, smaller than the true difference.

The reconstructed opening angle however does not change with the energy. In fact for the entire energy range, $\Delta\theta_{\text{reco}\,\mu,\,\text{reco}\ \ \text{shower}} \approx 15^\circ$. The reconstructed directions are more co-linear than the true directions are, indicating a bias in the direction estimate of the track+shower reconstruction. It is important to note that the shower direction uses the track prefit direction as a starting point, and the fit result has not diverged much from the muon direction. It is likely that the shower direction, which is difficult to estimate, is reconstructed along the same direction as the track, which tends to have more energy and produce more hits in the detector than the shower. The larger spread on the reconstructed opening angle indicates the difficulty in resolving the directions. Beyond 40 GeV however, the true and reconstructed opening angle become closer in value.

The Bjorken-y of the events, the fraction of the energy transferred to the associated hadronic shower, is also reconstructed. The true value of the Bjorken-y is not available from the simulation software version used to simulate these events. The ‘true‘ Bjorken-y



(a)



(b)

Figure 6.17.: (a) The true Bjorken-y is shown as a function of the true neutrino energy. (b) The median and 16% – 84% quantiles of $y^{\text{reco}} - y^{\text{true}}$ is shown as a function of the neutrino energy.

is instead calculated using its definition in Equation (1.3). The reconstructed Bjorken-y is approximated as

$$y^{\text{reco}} = \frac{E_\nu^{reco} - E_\mu^{reco}}{E_\nu^{reco}}, \tag{6.10}$$

where E_ν^{reco} is defined in Equation (6.3) as the sum of the reconstructed muon and shower energies. Figure 6.17 (a) shows the distribution of the true Bjorken-y value as a function of the true neutrino energy. For events from approximately 3—30 GeV, the true Bjorken-y value is ≤ 0.2, for the majority of events. This energy range contains the majority of

events in the event sample (c.f. the true energy distribution of Figure 6.7). Figure 6.17 (b) shows the median and quantiles of $y^{\text{reco}} - y^{\text{true}}$ as a function of the neutrino energy. For true neutrino energies of approximately 10–100 GeV, the median difference between the true and reconstructed Bjorken-y is about 0.1. The quantiles indicate the resolution to the quantity, which at 10 GeV shows a total spread of 0.4 between the maximum and minimum value of the quantiles band. Note that this quantity has only been investigated for neutrinos and not anti-neutrinos. The difference between the reconstructed Bjorken-y for neutrinos and anti-neutrinos is explored in Chapter 7.

This is the first time the Bjorken-y has been reconstructed directly for events in the KM3NeT experiment, showing a median difference of $y^{\text{reco}} - y^{\text{true}} \sim 0.1$. For comparison, in Adrían-Martínez et al. (2016) [17], the Bjorken-y is estimated from the time residual distribution of hits from either a track or shower hypothesis — certainly not a direct estimate. In that case, the reconstructed Bjorken-y is estimated for 4 different intervals of the true Bjorken-y value. Again, the use of such a track+shower reconstruction algorithm is justified from its ability to reconstruct the Bjorken-y parameter.

6.5.2. Charged-Current Electron Neutrino Events

The performance of the reconstruction is also examined for charged-current electron neutrino (ν_e-CC) events. These interactions result in the creation of both an electromagnetic and a hadronic shower. The results from the electromagnetic shower reconstruction for KM3NeT/ORCA events, described in Chapter 3, are used to provide starting points to the reconstruction for these events. The fit procedure is applied to a number of simulated ν_e-CC events, simulated in the same manner as for the muon neutrino events described above, with energies of 1–100 GeV. The energy spectrum of these events is shown in Figure 6.18, with a comparison between the true visible shower energy and the energy of the electron which initiates the shower. The visible energy definition accounts for the hadronic shower particles which have more (electromagnetic-equivalent) energy than the initial electron alone. In this section, the electron energy is chosen to represent the true energy of the shower when presenting the resolution in the following pages.

During the fit procedure the track direction and energy parameters are not fixed, they are free to converge on the best estimated value. In this way we check how the track+shower reconstruction responds to the absence of a track in the neutrino event. Again, to select better reconstructed events for this performance evaluation, a simple cut is made on the total negative log-likelihood value of $-\log\mathcal{L} > 22500$ & $-\log\mathcal{L} < 28000$. This range can be shown in Figure 6.19, where the angular resolution and

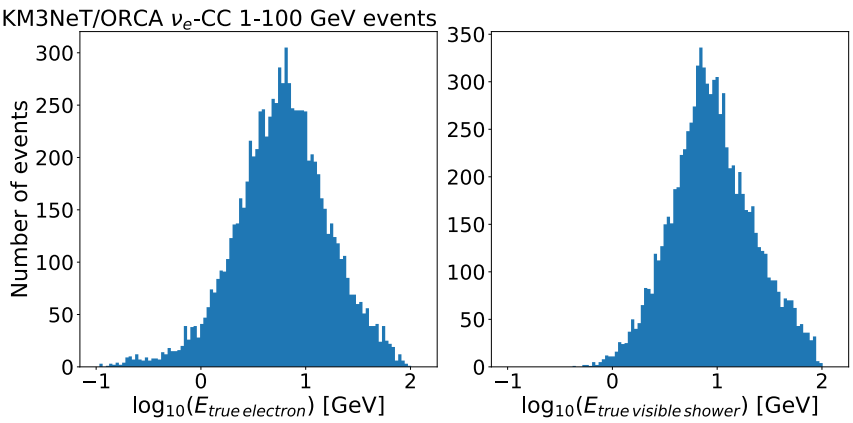


Figure 6.18.: The true energy of the particle showers for the ν_e-CC events used to evaluate the track+shower reconstruction performance. A comparison is shown between the true visible and the initial electron energy.

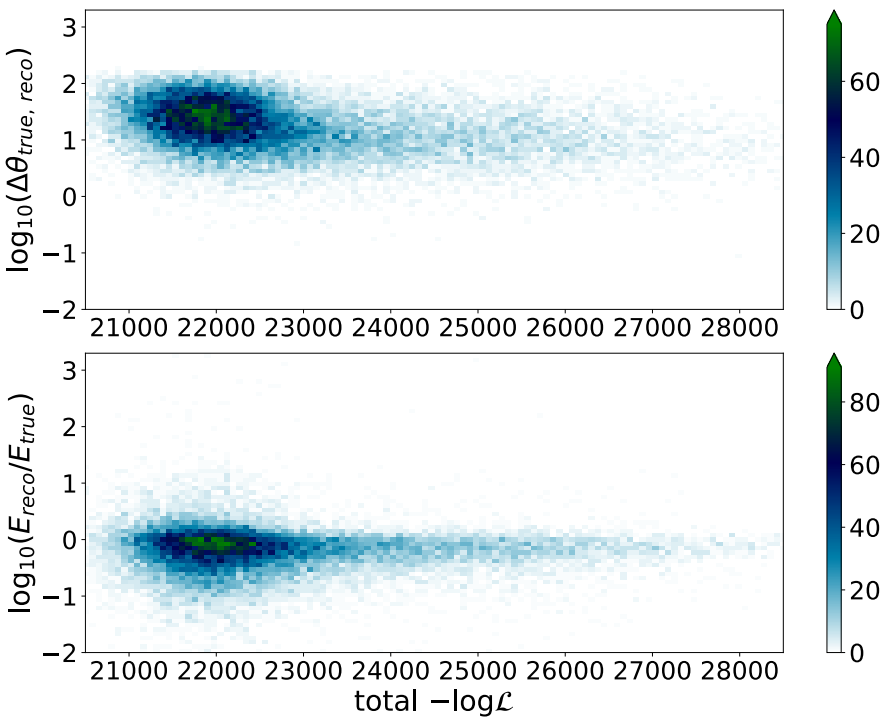


Figure 6.19.: The angular difference $\Delta\theta_{\text{true, reco}}$ and $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ are shown as a function the total negative log-likelihood, comparing the true and reconstructed shower directions and energies. A range of $-\log\mathcal{L}$ from 22500–28000 is selected to remove some poorly reconstructed events. The colour-bar indicates the number of events.

$\log_{10}(E_{\text{reco}}/E_{\text{true}})$ are shown as a function of $-\log\mathcal{L}$. This simple cut is chosen to improve the angular and energy resolution. Many poorly-reconstructed events congregate around low values of $-\log\mathcal{L}$. This may be due to limited information to reconstruct

these events properly. The effect of this cut on the number of events and resolutions is presented in Table 6.2.

Quantity	Before Cuts	After Cuts
Number of events with fit result	23,389	8,917
Median $\Delta\theta_{\text{true electron,reco } sh}$	18.8°	13.0°
σ on $\log_{10}(E_{\text{reco shower}}/E_{\text{true electron}})$	0.5	0.4

Table 6.2.: Example cuts used for the charged-current electron neutrino events on which the track+shower reconstruction performance is evaluated.

The distance between the neutrino vertex and reconstructed vertex is shown in Figure 6.20 for both the track+shower and standard shower reconstruction algorithms. Most events are reconstructed < 5 metres from the true vertex position.

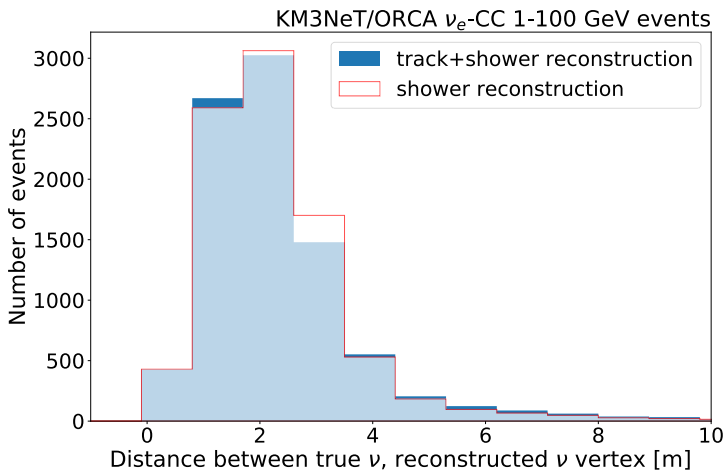
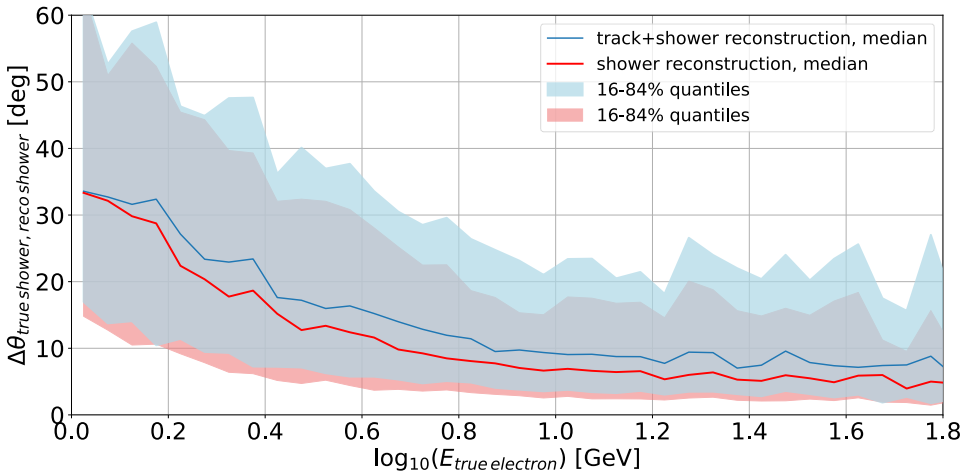


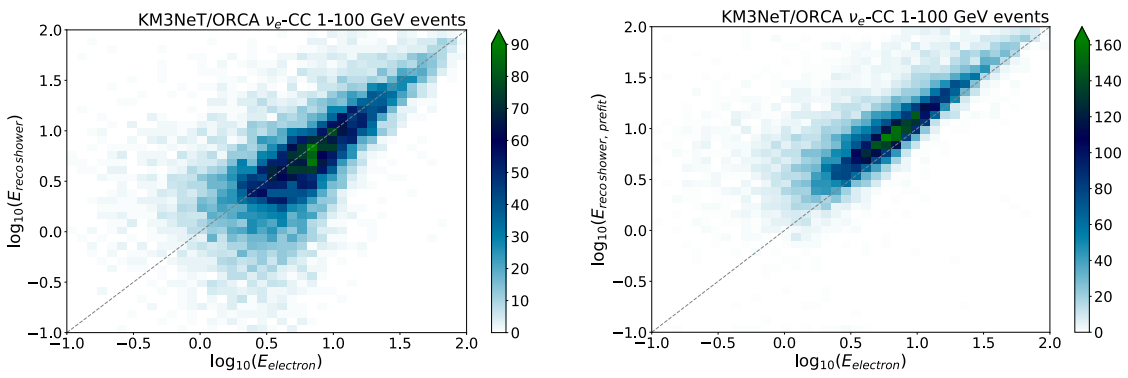
Figure 6.20.: The distance between the true neutrino vertex and reconstructed vertex position, from the track+shower reconstruction (blue) and standard shower reconstruction (red), for ν_e-CC events. Most events are reconstructed with a vertex position very close to the true neutrino position.

The direction resolution of the track+shower reconstruction in reconstructing showers is presented in Figure 6.21, with the resolution from the shower reconstruction also shown. The median of the distributions, alongside the 16% – 84% quantiles, are illustrated. The angular resolution for both reconstructions is similar, with the slightly better resolution from the shower reconstruction probably coming from the track+shower reconstruction attempting to also reconstruct a muon direction — perhaps treating the electromagnetic shower and hadronic shower as separate components. In comparing the median of both lines, the values are very similar, with the track+shower reconstruction providing an angular resolution of $\sim 10^\circ$ from the true direction at 10 GeV.



(a)

Figure 6.21.: The angular resolution of the shower direction as a function of the electron energy, for both the track+shower reconstruction (blue) and the standard KM3NeT/ORCA shower reconstruction (red). For both cases, the median and 16% – 84% quantiles of the distributions are indicated.

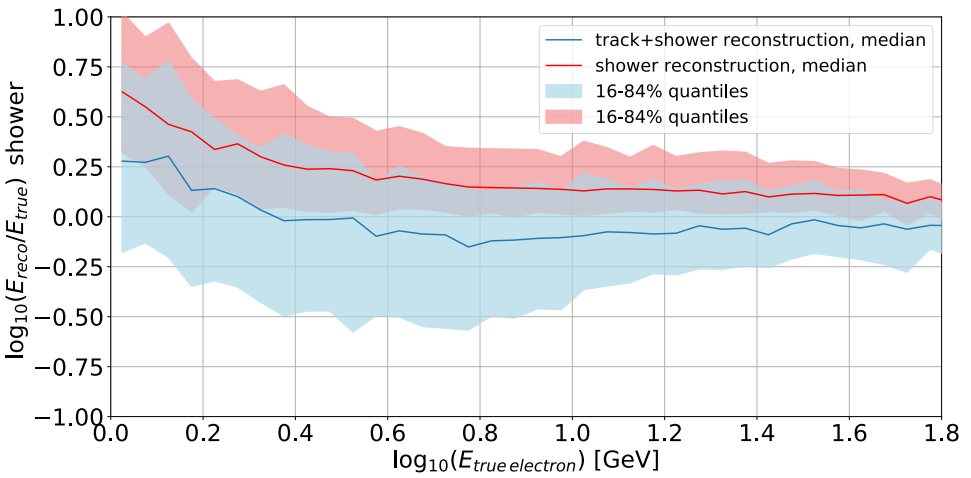


(a) **(b)**

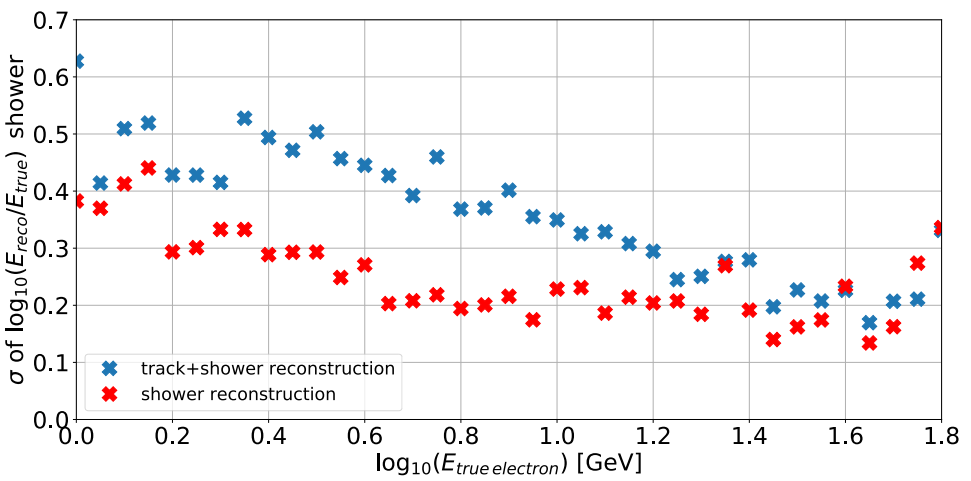
Figure 6.22.: (a) The reconstructed energy from the track+shower reconstruction, and (b) the reconstructed shower energy from the KM3NeT/ORCA shower reconstruction, both as a function of the true electron energy. The diagonal line in both figures shows where the reconstructed energy is equal to the true energy.

In Figure 6.22 (a) and (b), the reconstructed energy and prefit (from the standard shower reconstruction) shower energy as a function of the true electron energy (the electron which initiates the electromagnetic shower) are presented, respectively. Although the KM3NeT/ORCA-dedicated shower reconstruction somewhat overestimates the electron energy, many events lie close to the diagonal line with relatively little spread in the distribution, indicating a good reconstructed energy estimate. The reconstructed energy from the track+shower reconstruction also lies close to the true energy for many events, though with a larger spread around the diagonal line compared to the prefit

case. Though the track+shower reconstruction employs the pre-defined photo-electron distributions for electromagnetic showers, the position of light emission for the shower is assumed to be that of a hadronic one (c.f. Section 6.2.1). This may contribute to the difference between the results from the two reconstruction procedures. The fact that these events consist of an electromagnetic and hadronic shower may also impact the track+shower reconstruction performance; the model assumed in this reconstruction is of a track and shower. Whether these two showers are treated separately by the track+shower reconstruction should be investigated in a future study.



(a)



(b)

Figure 6.23.: (a) The median and quantiles for $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ as a function of the electron energy are illustrated for the track+shower reconstruction (blue) and the KM3NeT/ORCA-dedicated shower reconstruction (red). (b) The value of the standard deviation σ on $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ as a function of the electron energy, for the track+shower reconstruction (blue) and the KM3NeT/ORCA-dedicated shower reconstruction (red).

The distribution of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ is shown for the track+shower reconstruction and shower reconstruction as a function of the shower energy in Figure 6.23 (a). The median and quantiles of both reconstructions show the expected behaviour; the width of the distribution (i.e. the resolution) is worse (wider) at lower energies and improves (narrows) as the energy increases. Likewise, the bias reduces as the energy increases. Figure 6.23 (b) shows the width σ of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$, where E_{reco} is estimated by both the track+shower reconstruction and the standard shower reconstruction algorithm. At $E_{\text{true electron}} = 10$ GeV, $\sigma \simeq 0.23$ and 0.36 for the shower reconstruction and track+shower reconstruction procedures, respectively. For $E_{\text{true electron}} \geq 15$ GeV, the values from both reconstruction algorithms are quite similar. This may result from the hadronic shower scaling of the shower maximum point, as described in Section 6.2.1. The scaling between the true hadronic and visible energy becomes less prominent beyond $\sim$ 10 GeV, as shown in Figure 6.2. For low energies below $E_{\text{true electron}} \sim 10$ GeV, the difference in the energy resolution between the two reconstruction procedures is larger, with the dedicated electromagnetic shower reconstruction procedure resulting in a better resolution compared to that from the track+shower reconstruction.

As already mentioned, the track+shower reconstruction assumes an emission point in the longitudinal direction of a hadronic shower, which may have an impact on these resolutions. There are many events for which a non-zero track energy is reconstructed, as depicted in in Figure 6.24. During the reconstruction procedure, the starting point energy is set to 0.1 GeV for events with energies ≤ 0.1 GeV. This is why many events have a reconstructed muon energy of 0.1 GeV, no better value is found in the fit procedure. The reconstructed shower energy, meanwhile looks more similar to the true energy spectrum of Figure 6.18.

The fact that the muon parameters must be left free in fitting these particle showers with this reconstruction technique is likely affecting the performance compared to the KM3NeT/ORCA-dedicated shower reconstruction, however, when reconstructing data from the KM3NeT detectors this reconstruction cannot discriminate between neutrino flavours, and so both the track and shower parameters are left free in the fitting procedure.

6.6. Discussion

For the first time, a reconstruction algorithm has been developed and implemented with the goal to reconstruct both the track and (hadronic) shower components of neutrino-

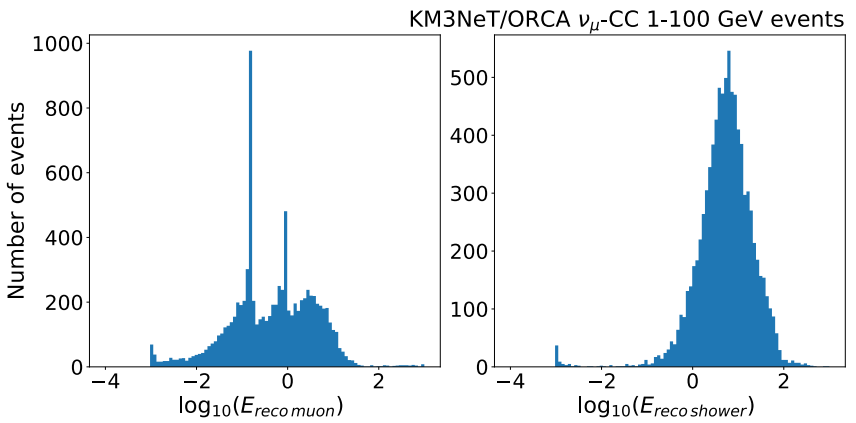


Figure 6.24.: The distributions of the reconstructed muon and shower energies determined by the track+shower reconstruction procedure for charged-current electron neutrino events.

induced events detected by the KM3NeT experiment. Such a reconstruction is especially relevant for ν_μ-CC events which can more accurately be described by this track and shower event model. The track+shower reconstruction procedure is applied to simulated sets of charged-current muon and electron neutrino events in the complete KM3NeT/ORCA detector. For the ν_μ-CC events, it is shown to be capable of reconstructing the vertex position, the direction and energy of the muon component, and the direction and energy of the shower component of the events. Furthermore, this allows for an estimate of the neutrino energy and direction, no longer requiring the shower or muon direction alone as a proxy for the neutrino direction and energy. For ν_e-CC events, the reconstruction procedure is capable of reconstructing the direction and energy of ν_e-CC showers, even when assuming a track in the event model. The emission profile for hadronic showers is implemented in this reconstruction procedure, as outlined in Section 6.2.1. Such an implementation may result in differences between the electromagnetic shower reconstruction algorithm and this track+shower algorithm. The performance of the track+shower algorithm will be studied for neutral-current neutrino interactions (i.e. pure hadronic shower events) in Chapter 7.

The likelihood function of Equation (3.23) has been implemented and used to scan the likelihood space for the model parameters describing this model. For the track model parameters, the likelihood minimum coincides with the true parameter value. For the case of the hadronic shower direction and energy, the likelihood landscape can be a difficult one for the minimiser to navigate, and the true value of the shower direction and energy cannot be found for all events in the likelihood scan carried out. As indicated in Section 6.3.3, the task of reconstructing the direction and energy of a hadronic shower is a challenging one, with the minimum value of the total likelihood

function at a different position in the parameter space from the true parameter value. For a vast majority of events, the total likelihood is flat for a scan of the shower direction and energy values. The results of the track+shower reconstruction, presented in Section 6.5, show a good performance of the reconstruction procedure — e.g. the shower direction is estimated within 20° of the true direction at shower energies of 10 GeV, and with an energy resolution of $\sigma = 0.4$ at shower energies of 10 GeV. These resolutions are likely aided by the resolution of the standard track reconstruction procedure, which provides the starting points to the fit.

In Figure 6.16, the reconstructed shower direction is shown to be estimated with a direction close to that of the muon direction, the two directions are reconstructed within 10 degrees of one another for all considered ν_μ-CC energies. The use of the muon direction as a starting point to the shower fit likely influences the result. Figure 6.16 shows the true muon and shower direction differ in angle by $\geq 15^\circ$ below 40 GeV; the shower direction may not vary enough in the fit procedure to find the optimal estimate. The shower energy, meanwhile, is estimated well enough to show a promising contribution in estimating the actual neutrino energy (e.g. as shown in Figure 6.15).

The visible shower direction is compared to the 'true' momentum direction of the shower, as can be calculated from Equation (6.2), for ν_μ-CC events in Figure 6.25. For many events, these two directions are within 7° of one another for neutrino energies above 10 GeV. At neutrino energies of about 6 GeV, these two direction definitions can differ by $\sim 11^\circ$. However, the momentum-direction definition should incorporate the neutral particles produced within the shower which carry some of the momentum, but produce no Cherenkov light. These particles simply cannot be reconstructed. The visible direction definition is maintained as a definition of the light which can actually be detected and reconstructed from the shower. Events which have neutrino energies < 5 GeV are excluded due to a lack of statistics in the sample.

There are means of improving the performance of this track+shower reconstruction algorithm. Different minimization algorithms and different means of minimising the likelihood can be explored for the track+shower reconstruction procedure, which may lead to further improvements on the performance. As mentioned multiple times in this chapter, electromagnetic shower photo-electron distribution tables are used in this reconstruction to estimate the hadronic shower model parameters. This is an outstanding approximation in this reconstruction technique, which is thus far not fully addressed. More accurate probability density functions, created through the simulation of many hadronic shower events, can be obtained and used to improve upon the estimation of the photo-electron yield. Such studies remain outside the scope of this particular thesis.

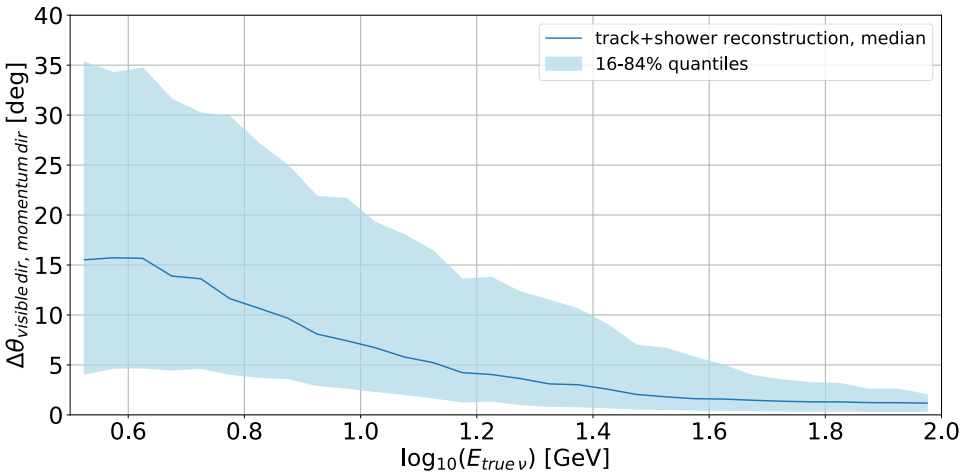


Figure 6.25.: The angular difference between the shower direction as defined by the visible direction and the momentum-defined direction, as a function of the neutrino energy. The median and 16–84% quantiles are indicated.

Another factor which can influence the evaluation of the performance of this track+shower reconstruction is the simulation of the events for the 115-line KM3NeT/ORCA detector, which is outdated. This simulation contains many known issues, an example being that the final state particles cannot be tagged, affecting the definition of the true visible shower direction/energy. As a result, all particles in the event are used to define the visible energy and direction. In addition, the Bjorken-y is not calculated and stored in this simulation version of the ORCA-115 line detector, and thus is calculated using Equation (1.3) instead. A new simulation of all-flavour neutrino/anti-neutrino and atmospheric muons is being carried out at the time of writing for the complete KM3NeT/ORCA detector. The performance of this reconstruction procedure will likely be affected by the updates in the software, but such a study will not be addressed in this thesis.

This reconstruction procedure paves the way for reconstructing the 'golden' quantity in neutrino interactions, the Bjorken-y parameter. This parameter allows for a complete description of the neutrino interaction. As shown in Figure 6.17 (b), events can be reconstructed by this novel algorithm for which the true and reconstructed Bjorken-y are within 0.15 in value, for neutrino energies of 1–100 GeV. Using this track+shower reconstruction as an input to a neutrino oscillation parameter analysis will be explored in the following chapter.

Chapter 7.

Towards an Oscillation Analysis

It has been demonstrated in Chapter 6 that the novel track+shower reconstruction algorithm, developed as part of the work for this thesis, can be used to estimate the vertex position, event time, and the directions and energies of the track and shower components of neutrino interactions. This reconstruction procedure is shown in Chapter 6 to improve the estimate of the neutrino energy for events simulated in the envisaged, 115-detection unit KM3NeT/ORCA detector, and for the first time to directly estimate the Bjorken-y parameter. In this chapter, the track+shower reconstruction method is applied to simulated data in the KM3NeT/ORCA-6 detector, and the potential of using this reconstruction algorithm for a future analysis is explored. Although a precise determination of the neutrino oscillation parameters (an '**oscillation analysis**'), and ultimately the determination of the neutrino mass ordering, are main goals of the KM3NeT experiment, many analyses require the same input, hence the studies carried out in this chapter are applicable to many possible analyses with the KM3NeT/ORCA-6 data.

A step in an analysis termed '*particle identification*' is usually employed to identify and discriminate between track- and shower-like (ν_μ-CC, ν_τ-CC), shower-like (ν_μ-NC, ν_e-CC) and atmospheric muon-like events. Machine learning algorithms are used to identify particular features of these different event types, and assign 'scores' to them based on how track-like, shower-like, or atmospheric muon-like they are. An example of this particle identification step, for the oscillation analysis of KM3NeT/ORCA-6 data, is described in Carretero (2023) [91]. The track+shower reconstruction procedure reconstructs some variables not provided by the separate track or shower reconstruction algorithms, including the directions and energies of the respective track and shower components and the reconstructed likelihood components defined in Equation (3.23). Distributions of the likelihood, resulting from the track+shower reconstruction, are presented in Section 7.2.1, for neutrino events containing only particle showers and

events comprising a track and shower. Differences between these distributions are found for the two event classes. As is explained in Section 7.2.1, these differences can complement and provide further discriminatory input to the particle identification step in an oscillation analysis.

As discussed in Chapter 6, the track+shower reconstruction algorithm results in an improved estimate of the neutrino energy for the envisaged KM3NeT/ORCA detector, compared to estimating only the track or shower energy as a proxy for the neutrino energy. In order to see if this improvement, and any other improvements to the resolution, also pertains to the ORCA-6 detector, the expected resolution on the neutrino direction, neutrino energy and Bjorken-y is determined and presented in Section 7.2.2. These resolutions are shown for the various neutrino interaction channels in the ORCA-6 dataset, comparing the results between track- and shower-like and shower-like events. The resolution performance is also compared between the ORCA-6 detector and the envisaged, complete ORCA-115 detector.

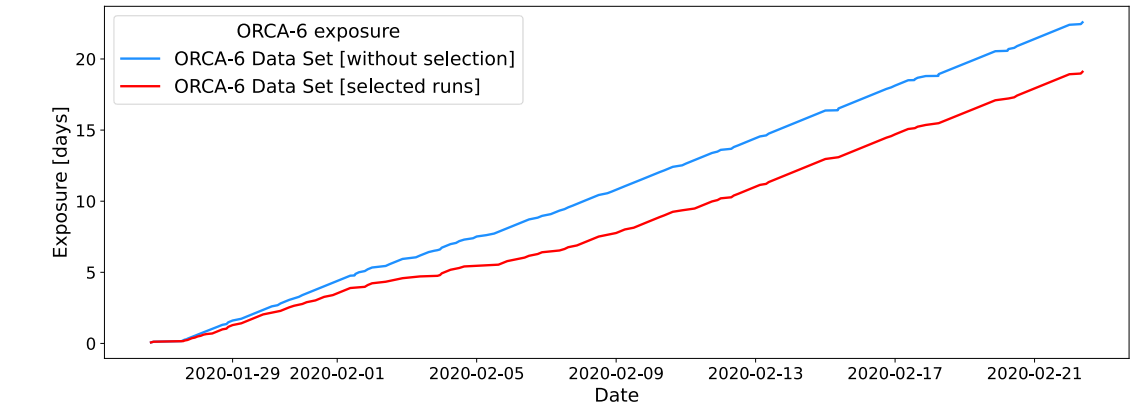
A summary and discussion of the results of this study are presented later in this chapter, followed by conclusion and future outlook for this work.

7.1. Data Set - ORCA-6

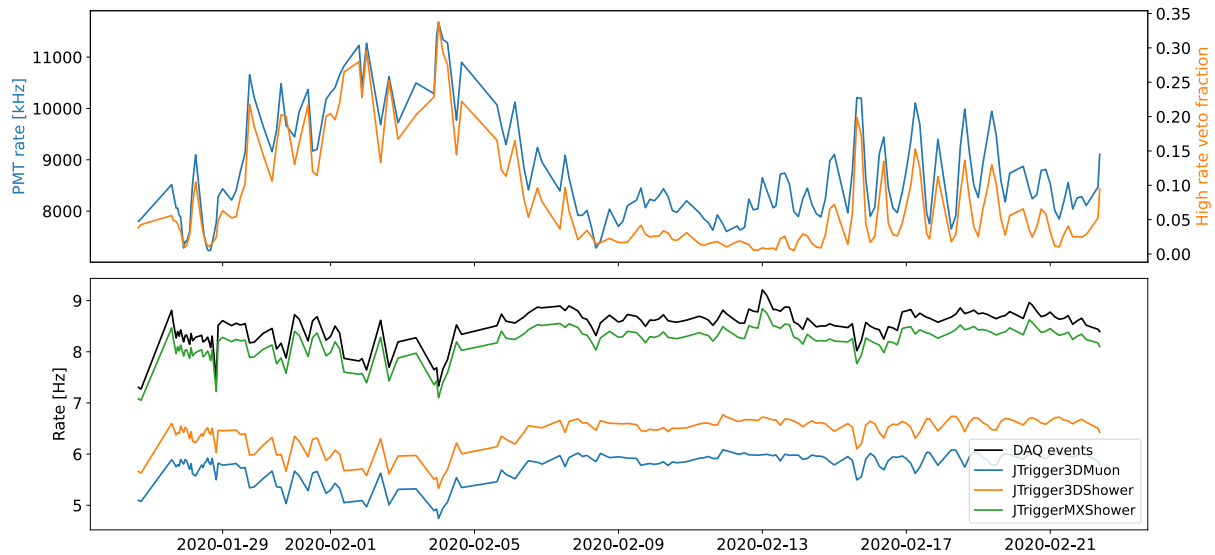
The data set taken by the ORCA-6 detector is used in this chapter in exploring the performance and potential of the track+shower reconstruction algorithm. This data set is well-understood and has been used in other physics analyses of KM3NeT data. As shown in Figure 2.12 and discussed in Carretero (2023) [91], a data set from the ORCA-6 detector amounting to 510 days of data-taking is used to perform an oscillation analysis. The data used in this chapter is merely a subset of that.

Figure 7.1 (a) presents the cumulative exposure (data-taking time period) of the data set used in this chapter from run numbers 7224 – 7495, as a function of the date. The data runs which are suitable for physics analysis must comply with a set of criteria, which include a sufficient run duration and minimal periods of high-rate veto (defined in Chapter 2). The data runs which comply with these conditions are referred to as the 'selected runs', their cumulative exposure is also presented in Figure 7.1 (a).

The triggered event rate (triggered events are referred to as DAQ Events, as described in Section 2.5) for the selected data runs is shown in Figure 7.1 (b) as a function of the time, alongside the rate of DAQ events from the individual trigger algorithms (see



(a)



(b)

Figure 7.1.: (a) The cumulative livetime of the ORCA-6 data set on which the track+shower reconstruction is applied is presented. The full data set (blue line) and selected quality data runs (red line) are indicated. (b) (Top) The mean PMT rate (blue line) and high-rate veto fraction (orange line) for the selected runs are shown as a function of the date. (Bottom) The trigger rate is shown for all triggered events (black line) and from the respective trigger algorithms, which are described in Chapter 2.

Section 2.5.1). As previously mentioned, the mixed trigger algorithm **MXShower** aims to identify events with both track-like and shower-like characteristics at lower neutrino energies. Clearly, the majority of triggered events are identified as such. The mean (per day) PMT rate and high-rate veto (see Chapter 2) fraction (i.e. the fraction of times one optical module reaches the high-rate veto) is illustrated in Figure 7.1 (b). The high-rate veto fraction and the PMT rate are clearly correlated, since the veto is implemented for high PMT rates. The trigger rate is anti-correlated with the high-rate veto fraction, since

more photons are rejected during these periods, and fewer events are identified by the trigger algorithms.

7.2. Simulated Neutrino Dataset

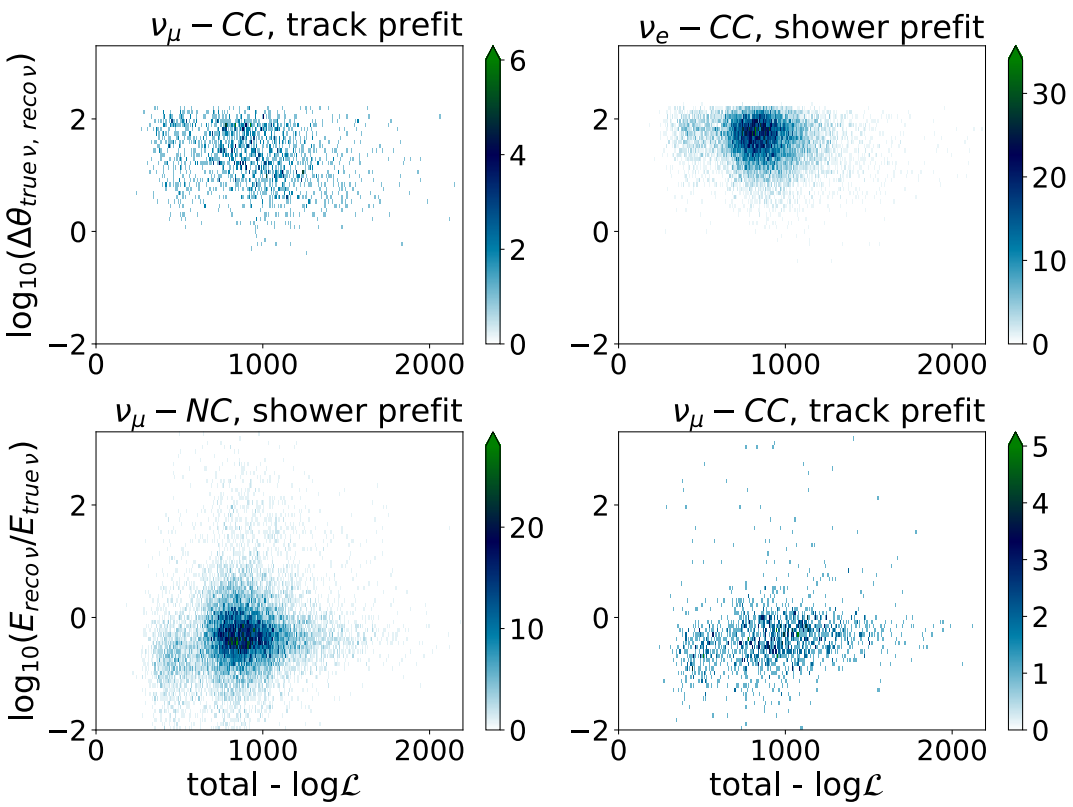
The track+shower reconstruction algorithm is applied to the data runs, simulated following the generation, light and trigger run-by-run software described in Chapter 3. The considered neutrino events are simulated with an energy from 1–100 GeV; this energy range is chosen as it is pertinent to neutrino oscillation studies. The true, simulated energy distributions for these simulated data sets can be found in Appendix E. Given the livetime of the data runs, the appropriate number of neutrinos and atmospheric muons are simulated to comply with the run livetime. The number of generated events for each run and the spectral index γ (of the energy spectrum $E^{-\gamma}$) according to which the events are generated, are given for the different channels in Table 7.1. Since the ν_μ-NC channel is used to represent neutral-current interactions of all flavours, more events are generated than for ν_e-CC events. For the ν_τ-CC and ν_μ-CC channels, fewer events end up being triggered on since they are simulated in a larger volume than the shower-like interaction channels. As a result there are less statistics available in reconstructed ν_μ-CC events compared to ν_μ-NC events, for example.

Type	Generated Neutrinos	Energy Range (GeV)	Spectral Index
ν_μ-CC	7.0×10^5	1–100	2.5
ν_e-CC	6.0×10^4	1–100	3.0
ν_μ-NC	2.0×10^5	1–100	2.5
ν_τ-CC	2.0×10^5	1–100	2.0

Table 7.1.: Number of generated neutrinos for each channel, according to a spectral index γ, for each simulated ORCA-6 run.

Currently in physics analyses of KM3NeT data, both the track and shower reconstruction algorithms are applied to the data and simulated data, as described in Chapters 3 and 6, followed by particle identification steps in order to identify track-like events, shower-like events, and background. In this section, the performance of the track+shower reconstruction algorithm is presented for each neutrino interaction channel, using the separate track and shower reconstruction algorithms to provide the starting points to the track+shower fit procedure. These are referred to as the 'track reconstruction prefit' and 'shower reconstruction prefit' in the below section, respectively.

In the track+shower fit procedure, the prefit direction is used as the starting directions for both the track and shower component, and likewise the prefit energy is used as the starting value of the energy for the track and shower component. Only muon-flavoured neutrinos are simulated to represent neutral current events for each data run.



(a)

Figure 7.2.: Distributions of the angular difference $\Delta\theta_{\text{true}\,\nu,\,\text{reco}\,\nu}$ and the logarithm of the energy ratio $\log_{10}(E_{\text{reco }\nu}/E_{\text{true }\nu})$ as a function of the negative log-likelihood $-\log\mathcal{L}$ from the track+shower reconstruction. These distributions are shown for various simulated interaction channels.

A cut on the total negative log-likelihood $-\log\mathcal{L}$ is used to improve the resolution to the track and/or shower direction and/or energy, and to show reconstructed distributions from the track+shower reconstruction, as presented later in this section. This selection criteria is used to select the well-reconstructed events resulting from the fit procedure. The $-\log\mathcal{L}$ which comes from the track+shower reconstruction is related to the quality of the angular and energy resolution. In general, events with a better resolution are reconstructed with a lower value of $-\log\mathcal{L}$, as also shown in Chapter 6. Some examples of the angular difference between the true and reconstructed neutrino direction and $\log_{10}(E_{\text{reco }\nu}/E_{\text{true }\nu})$ as a function of $-\log\mathcal{L}$ are shown in Figure 7.2, respectively. The definitions of the reconstructed neutrino direction and energy follow

	# of events	$\Delta\theta^{\text{median}}_{\text{true}\,\mu,\,\text{reco}\,\mu}$	$\Delta\theta^{\text{median}}_{\text{true}\,sh,\,\text{reco}\,sh}$	$\sigma^{\text{muon}}\,\log_{10}(E_{\text{reco}}/E_{\text{true}})$	$\sigma^{\text{shower}}\,\log_{10}(E_{\text{reco}}/E_{\text{true}})$
ν_μ-CC (track prefit)	1,773 966	8.6° 4.9°	35.4° 30.8°	0.4 0.4	1.4 1.3
ν_μ-CC (shower prefit)	1,746 972	33.5° 29.4°	41.3° 38.1°	0.8 0.8	1.1 0.9
ν_e-CC (track prefit)	13,151 5,009	- -	40.0° 36.6°	- -	0.8 0.7
ν_e-CC (shower prefit)	13,054 4,994	- -	39.1° 35.0°	- -	0.7 0.6
ν_τ-CC (track prefit)	7,237 3,281	6.8° 4.5°	39.9° 36.3°	0.4 0.4	1.3 1.2
ν_τ-CC (shower prefit)	7,065 3,201	33.9° 31.3°	33.8° 28.8°	0.8 0.9	0.8 0.6
ν_μ-NC (track prefit)	16,088 7,006	- -	41.5° 38.7°	- -	1.2 1.2
ν_μ-NC (shower prefit)	15,886 7,003	- -	37.2° 32.7°	- -	0.8 0.7

Table 7.2.: The effect of the selection criteria $-\log\mathcal{L} > 900$ on the reconstructed, simulated events in the ORCA-6 detector. The values in each column are given with no selection (left value) and with the selection criteria (right value). These values are presented for the neutrino channels resulting from both prefit cases, displaying the number of events and the angular and energy resolution values for the reconstructed track and shower compared to their true values.

from Equations (6.3) and (6.4). The distributions in Figure 7.2 are shown for various interaction channels. From Figure 7.2, a range of the likelihood from 900–2,200 is chosen to attempt to remove poorly-reconstructed events. At values of $-\log\mathcal{L} < 900$ the spread on the angular and energy deviations are large, and are therefore removed by the selected range. The effect of this cut, and its improvement on the resolution, is presented in Table 7.2, for the neutrino events (not anti-neutrinos).

Distributions which result from the track+shower reconstruction, relevant to particle identification, are compared below for the different neutrino interaction channels, followed by a discussion of the performance of the reconstruction. All of the following results use the cut on $-\log\mathcal{L}$ as described.

7.2.1. Providing Input for Particle Identification Methods

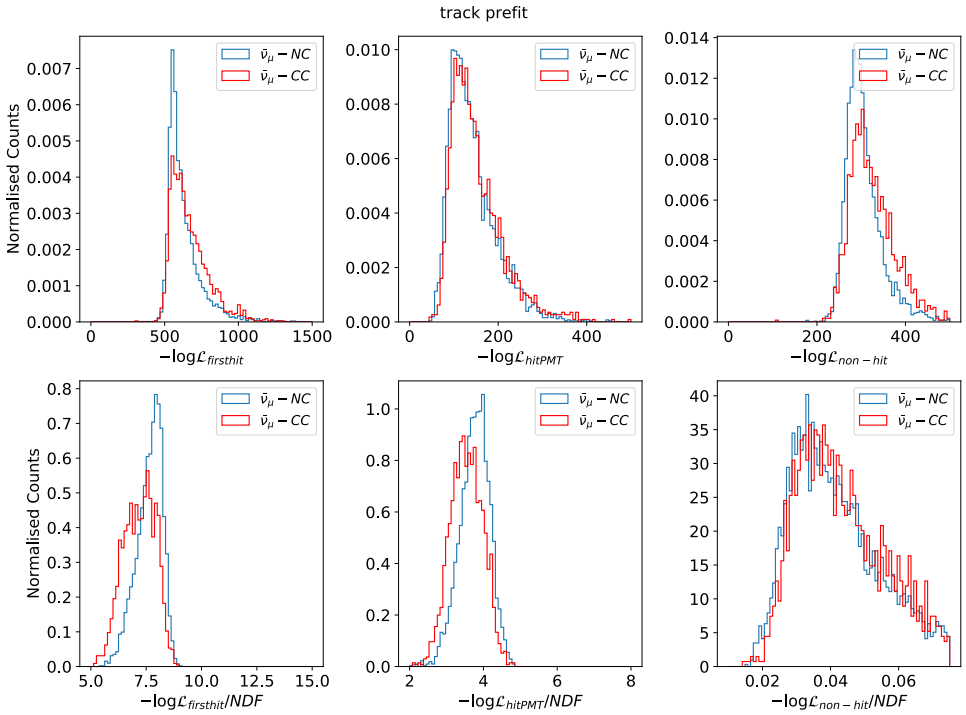
In this section, distributions of the likelihood are compared between events which contain only a hadronic shower (i.e. $\bar{\nu}_\mu$-NC events) and track- and shower-like events ($\bar{\nu}_\mu$-CC). The events from ν_τ-CC interactions and ν_e-CC interactions are omitted in this study. For ν_τ-CC interactions, fewer events are generated (for this data set) compared to ν_μ-CC interactions, meaning the reconstructed distributions are sparse in the number of statistics. For ν_e-CC interactions (i.e. electromagnetic and hadronic showers), the distributions are found to be quite similar (but not identical) to those reconstructed from ν_μ-NC events (i.e. hadronic showers only). The two channels compared here are intended to show the maximal difference between likelihood distributions. A comparison is not made between the reconstructed distributions for neutrino and anti-neutrino events, as these can be shown to be near-identical, except in comparing the reconstructed Bjorken-y parameter. The resolution on the Bjorken-y parameter with the ORCA-6 data will be illustrated later in this chapter.

Figure 7.3 shows the various likelihood components of Equation (3.23), for track- and shower-like events and shower-like events, as reconstructed using both prefits. The values of $-\log\mathcal{L}$ and $-\log\mathcal{L}/NDF$ are shown, respectively, where NDF denotes the number of degrees of freedom. The number of degrees of freedom for each negative log-likelihood component is defined as the number of hits (or, no hits) considered, and subtracting the number of fit parameters. The $-\log\mathcal{L}/NDF$ is used to normalise the events, removing the effects of those with very many or very few hits.

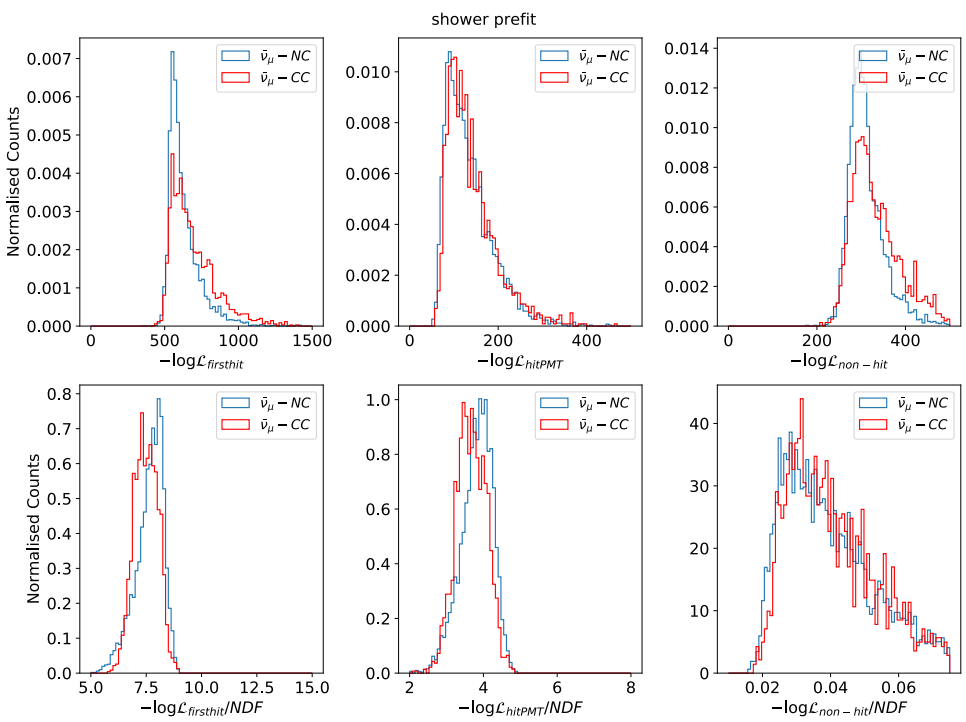
It is apparent that there are differences between these two event classes when these distributions are compared. Regardless of the prefit used, there is a difference be-

tween the $\bar{\nu}_\mu$-NC and $\bar{\nu}_\mu$-CC events, in the distributions of $-\log \mathcal{L}_{firsthit}/NDF$ and $-\log \mathcal{L}_{hitPMT}/NDF$. For example, there are more shower-like events reconstructed with $-\log \mathcal{L}_{firsthit}/NDF > 7.5$ and $-\log \mathcal{L}_{hitPMT}/NDF > 4$ compared to track- and shower-like events. There is a difference in the shape of these distributions for the two event classes.

Using both prefits to provide starting points to the track+shower reconstruction can be exploited to also find a difference between the two event classes. The difference in the values of the negative log-likelihood distributions between both the track and shower prefit case, for the two event classes, is shown in Figure 7.4. The difference in $-\log \mathcal{L}_{hitPMT}/NDF$ and $-\log \mathcal{L}_{non-hit}/NDF$ are chosen as examples. Again, the two event classes show different shapes in the distributions.



(a)



(b)

Figure 7.3.: The likelihood components resulting from the track+shower reconstruction in the ORCA-6 detector, for events containing only a hadronic shower ($\bar{\nu}_\mu$-NC) and events containing a track and shower ($\bar{\nu}_\mu$-CC). The bottom-row figures of (a) and (b) show the likelihood divided by the number of degrees of freedom.

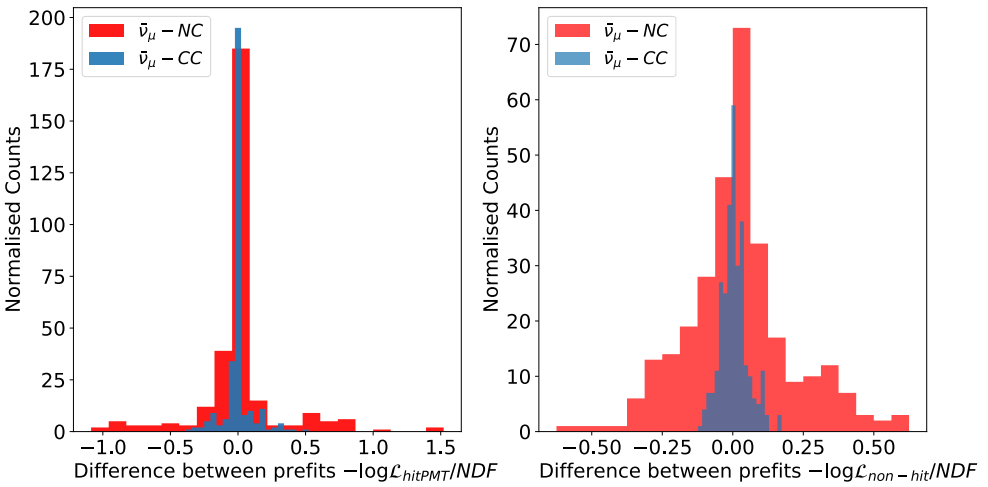


Figure 7.4.: The difference in the values of $-\log\mathcal{L}_{hitPMT}/NDF$ and $-\log\mathcal{L}_{non-hit}/NDF$, as resulting from the track and shower prefits, are shown for two different event classes.

7

Currently, the particle identification step used in analyses with ORCA-6 data takes a list of features (from the reconstructed data, such as likelihood values) as input, and makes use of machine learning techniques known as 'Boosted Decision Trees' [125] to classify how track-like, shower-like and atmospheric muon-like an event is. These Boosted Decision Trees are employed as they are powerful tools in identifying subtle differences between features for such classification, exactly like the subtle differences in the likelihood distributions shown in Figures 7.3 and 7.4. However, many features with differences between event classes are needed to train these Boosted Decision Trees such that they can discriminate effectively between different events. One feature currently used in ORCA-6 analysis to distinguish neutrino events from atmospheric muon background events, and track-like events from shower-like events, is the ratio of the likelihoods which come from the two separate track and shower reconstruction algorithms. The examples presented in this section, of differences between the negative log-likelihood components coming from the track+shower reconstruction, can clearly provide additional input to the particle identification step used in ORCA-6 analyses, alongside the variety of features already used. There are many other features from the track+shower reconstruction, such as the estimated vertex position or the opening angle between the track and shower component, that may differ between shower-like and track- and shower-like events and which can also be explored as input to the particle identification step in an analysis. Such studies are outside the scope of this thesis but should be carried out in future research.

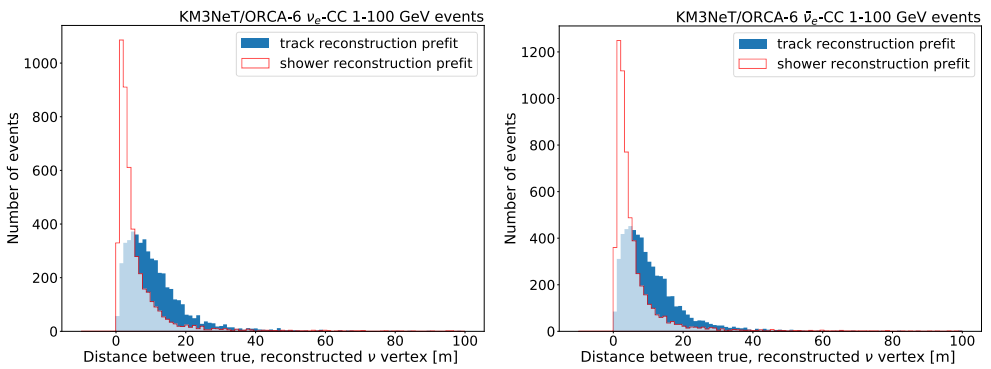
7.2.2. Resolution on the Neutrino Direction, Energy, Bjorken-y

The reconstructed angular and energy resolution are illustrated in this section, using the same quantities as in Chapter 6 to indicate the resolution to the direction or energy. The distance between the true and reconstructed vertex position is also presented, as well as an estimate on the reconstructed Bjorken-y. The results are shown separately for neutrinos and anti-neutrinos, in order to compare and contrast the reconstruction performance between both particle types. Though the resolution on the vertex, energy and direction should be almost equivalent between neutrinos and anti-neutrinos, a difference is expected between the Bjorken-y distributions. This is due to different neutrino-nucleon and anti-neutrino nucleon differential cross sections, as discussed in Section 1.2.2.

The reconstructed neutrino direction and energy is compared to the reconstructed track or shower directions and energies. This is done in order to give an idea of which reconstructed variable gives the best estimate for track- and shower-like (ν_μ-CC, ν_τ-CC) or shower-like (ν_e-CC, ν_μ-NC) events. Note that for the cases of the track direction and energy resolution presented below, the condition is enforced that there must be a true muon track in the event (in order to determine to angular and energy resolution to true track events only). As a result of this condition, there are less statistics available in e.g. the tau neutrino channel.

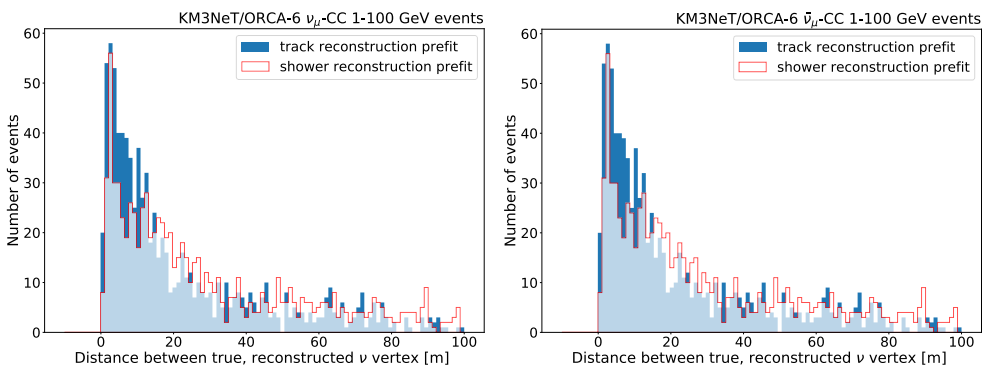
Vertex Resolution

The distance between the true and reconstructed vertex positions, for all the simulated neutrino interaction channels, is presented in Figures 7.5 to 7.8, for the two different prefits used. The shower prefit results in events which are reconstructed closer to that of the true neutrino vertex position for every event type, except for the ν_μ-CC interactions, for which the track and shower prefit give quite similar results (slightly more events are reconstructed closer to the true vertex position for the track prefit case).



(a) **(b)**

Figure 7.5.: The distance between the true and reconstructed vertex position for charged-current ν_e and $\bar{\nu}_e$ events simulated in the ORCA-6 detector. The results from the track (blue) and shower (red) reconstructions are shown, providing starting points to the track+shower fit procedure.

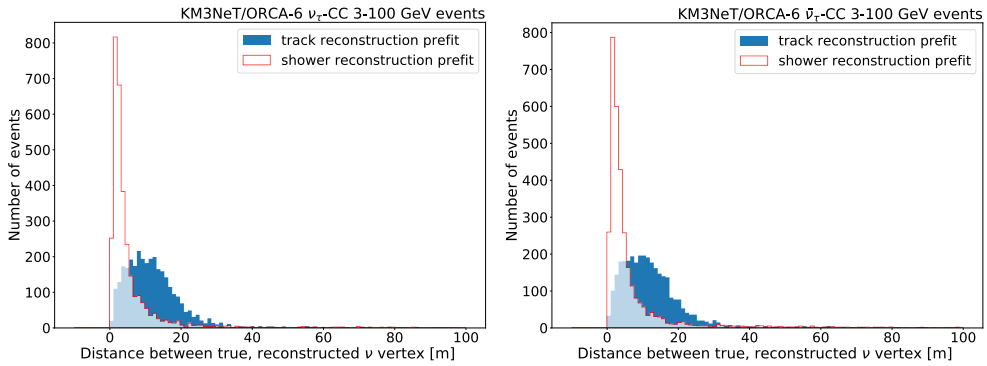


(a) **(b)**

Figure 7.6.: The distance between the true and reconstructed vertex position for charged-current ν_μ and $\bar{\nu}_\mu$ events simulated in the ORCA-6 detector. The results from the track (blue) and shower (red) reconstructions are shown, providing starting points to the track+shower fit procedure.

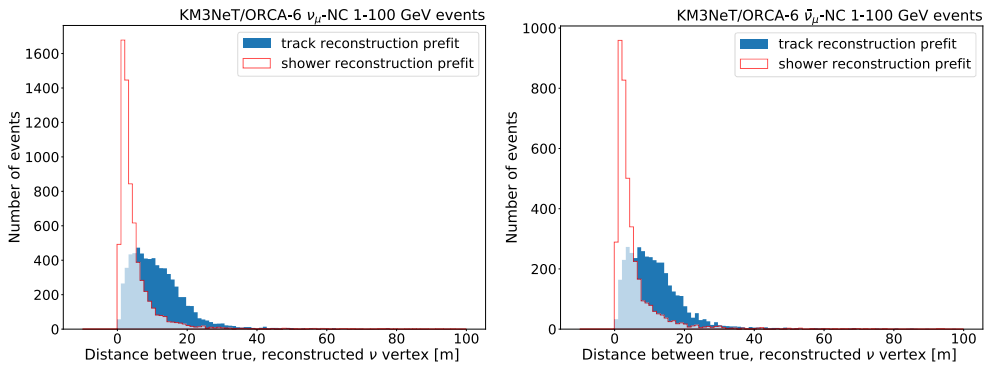
Considering that the tau lepton decays into a muon track $\sim$ 17% of the time and mostly creates a shower-like signature [3], the results for the tau neutrino channel show

the shower prefit giving the better vertex estimate. In general, these results suggest that the shower prefit should be used to provide the starting points in estimating the vertex position in the track+shower reconstruction. This in turn would lead to a better estimate of the other reconstructed parameters.



(a) **(b)**

Figure 7.7.: The distance between the true and reconstructed vertex position for charged-current ν_τ and $\bar{\nu}_\tau$ events simulated in the ORCA-6 detector. The results from the track (blue) and shower (red) reconstructions are shown, providing starting points to the track+shower fit procedure.



(a) **(b)**

Figure 7.8.: The distance between the true and reconstructed vertex position for neutral-current ν_μ and $\bar{\nu}_\mu$ events simulated in the ORCA-6 detector. The results from the track (blue) and shower (red) reconstructions are shown, providing starting points to the track+shower fit procedure.

Angular Resolution

The angular resolution as a function of the true neutrino energy is shown for track- and shower-like events in Figures 7.9 and 7.10 for neutrinos and anti-neutrinos. Using the track reconstruction as a starting point to the reconstruction procedure results in a

better angular resolution, which improves with increasing neutrino energy. The direction of the out-going track gives a better estimate of the true neutrino direction (i.e. a lower median value, and a more narrow spread in the quantiles), as opposed to the reconstructed neutrino direction definition of Equation (6.4). This observation is also made in Chapter 6 when evaluating the performance of the track+shower reconstruction for ν_μ-CC events in the ORCA-115 line detector. The track+shower reconstruction is capable of reconstructing both ν_μ and ν_τ events with a similar precision; the direction of the muon track is reconstructed with an angular resolution of $\sim 10^\circ$ at higher energies. Using the shower reconstruction as a prefit results in a much poorer angular resolution for the events, and the median values are similar when using the reconstructed track and reconstructed neutrino direction. Note that only the events which contain a muon are used to illustrate the median and quantiles of $\Delta\theta_{\text{true}\,\nu,\,\text{reco}\,\mu}$, i.e. there are far more statistics used in determining the median and quantiles for $\Delta\theta_{\text{true}\,\nu,\,\text{reco}\,\nu}$. For the track prefit case, the reconstructed muon direction alone provides a much better estimate on the true neutrino direction compared to the reconstructed neutrino direction. Using the shower prefit, the median and quantiles ranges of the angular resolution are more comparable in value.

For shower-like events, the angular resolution as a function of the true neutrino energy is shown in Figures 7.11 and 7.12 for neutrinos and anti-neutrinos. The shower prefit results in a better angular resolution compared to using the track reconstruction as a prefit. The improved resolution on the direction is likely aided by the better vertex position estimate. The resolution on the neutrino direction improves with increasing energy. Both the reconstructed shower direction and reconstructed neutrino direction from the track+shower fit result in very similar values of the median and quantile range of the angular resolution on the neutrino direction. The angular resolution is slightly better for the case of ν_e-CC events compared to neutral-current interactions. This may result from the use of the electromagnetic photo-electron distribution tables, described in Chapters 3 and 6. It is however clear that the track+shower reconstruction algorithm is capable of reconstructing hadronic shower events with an angular resolution within 20°–40° of the true neutrino direction, for the few-line ORCA-6 detector.

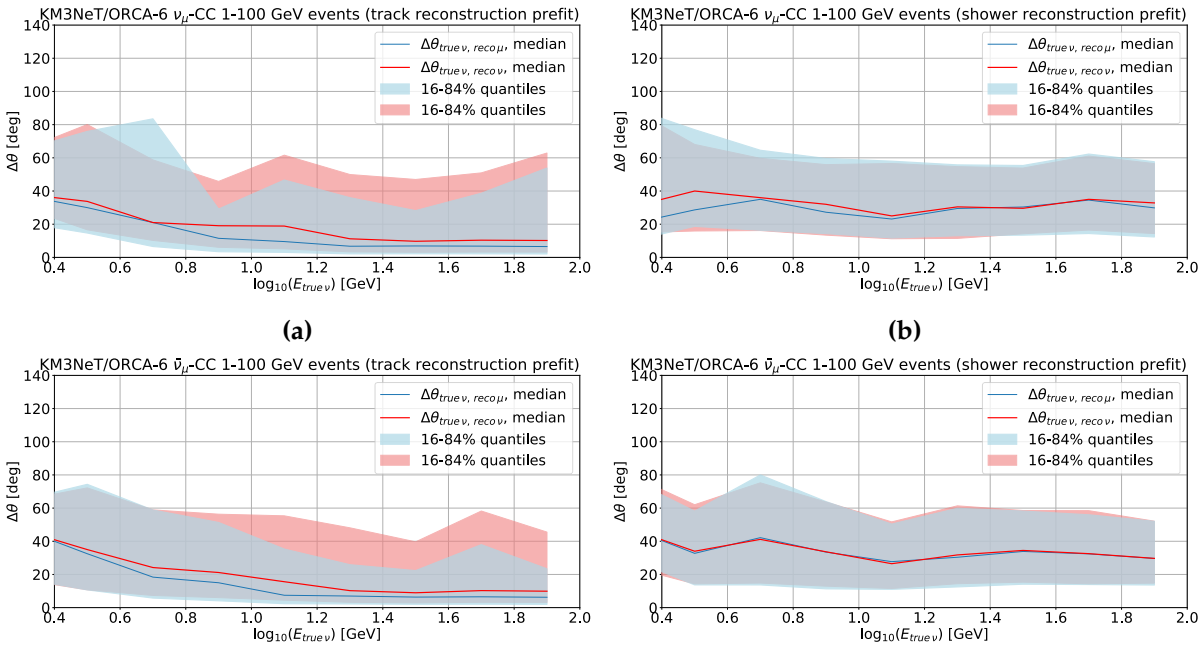


Figure 7.9.: The angular resolution as a function of the true neutrino energy for charged-current ν_μ and $\bar{\nu}_\mu$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure.

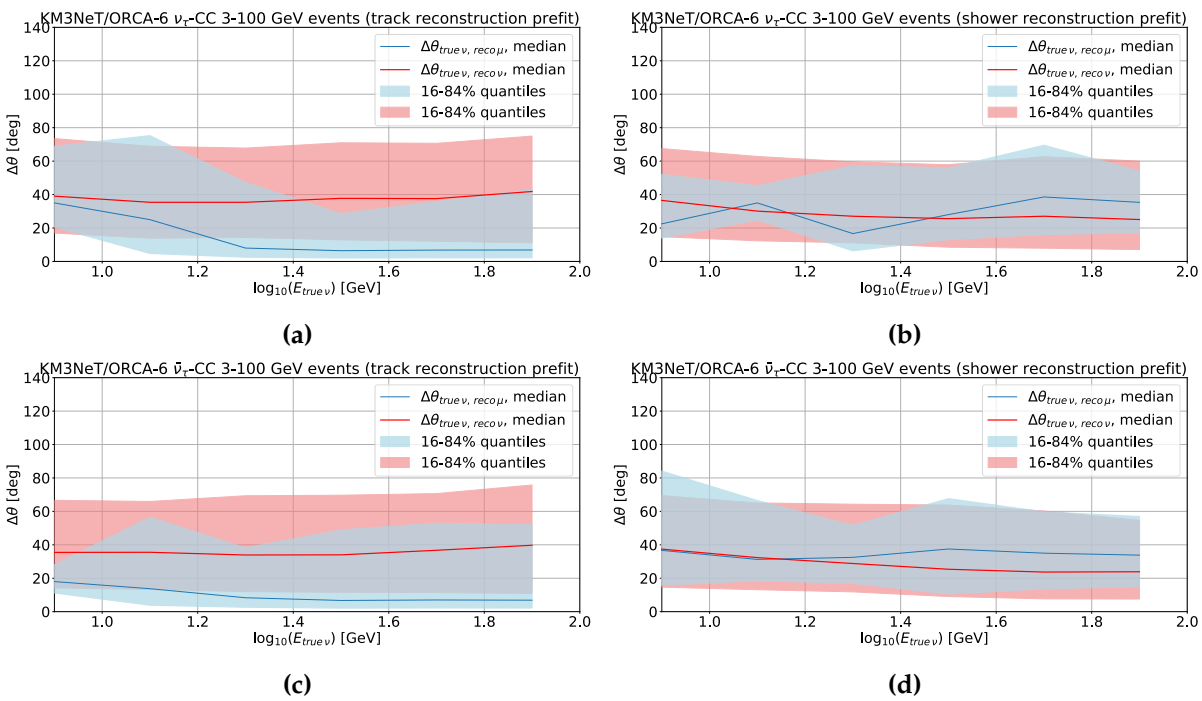


Figure 7.10.: The angular resolution as a function of the true neutrino energy for charged-current ν_τ and $\bar{\nu}_\tau$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure.

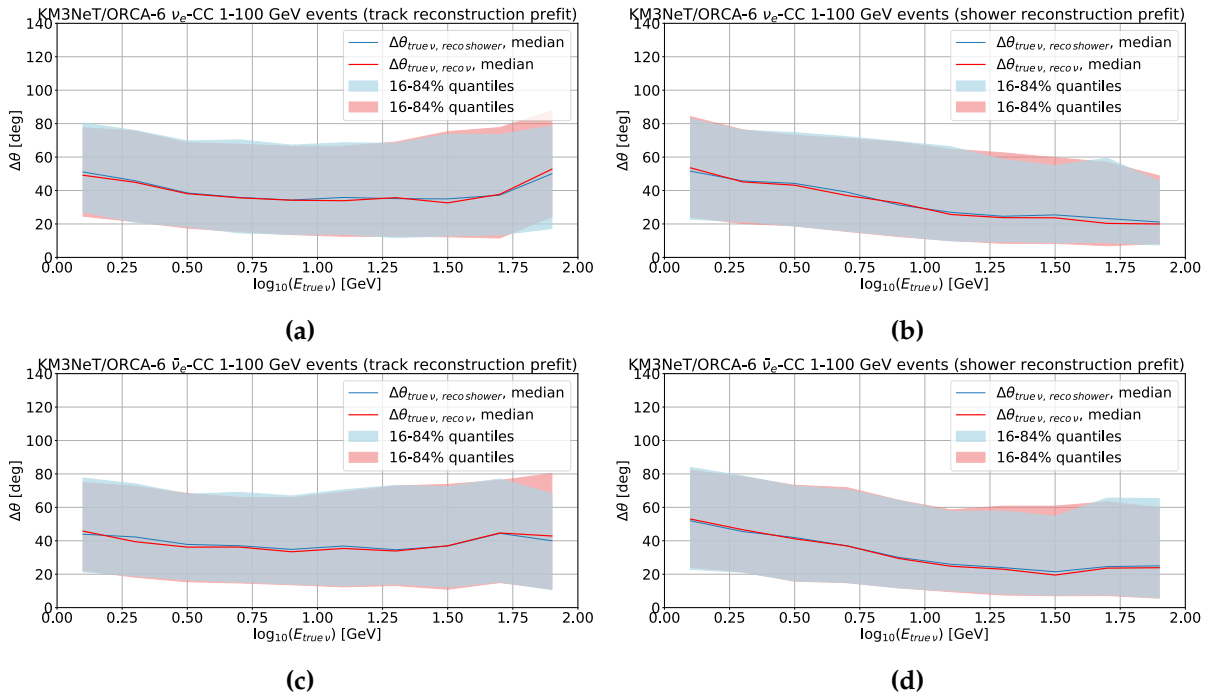


Figure 7.11.: The angular resolution as a function of the true neutrino energy for charged-current ν_e and $\bar{\nu}_e$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure.

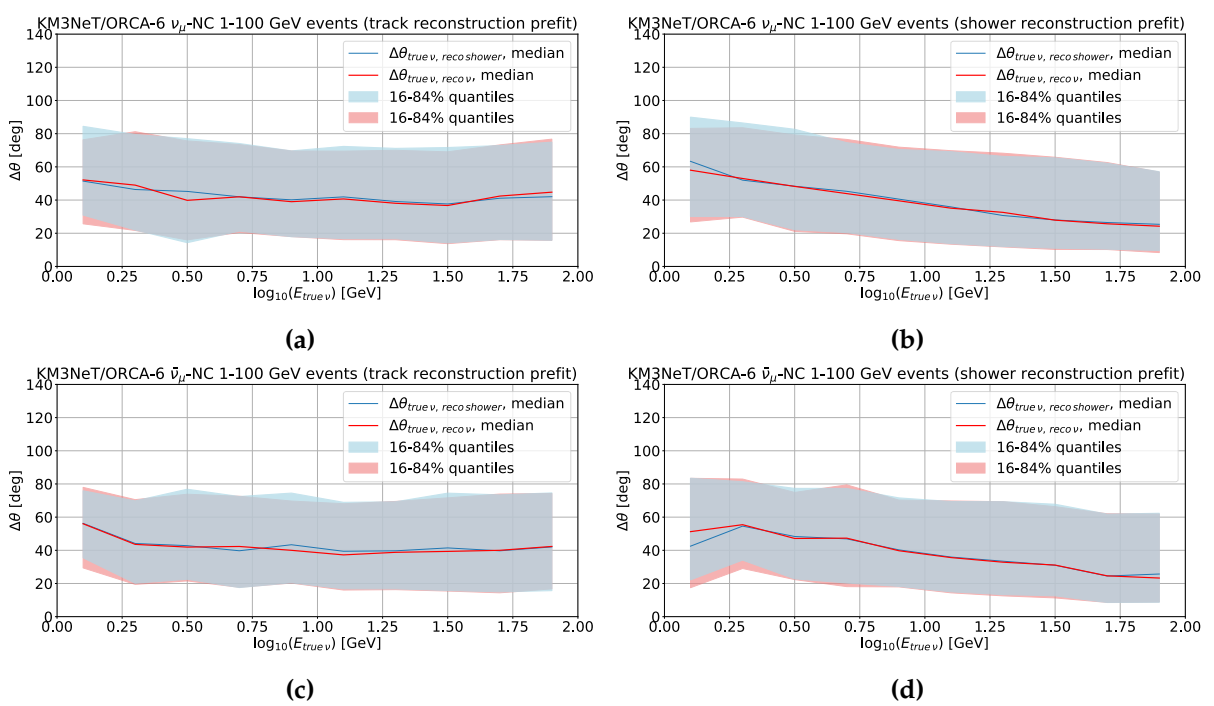


Figure 7.12.: The angular resolution as a function of the true neutrino energy for neutral-current ν_μ and $\bar{\nu}_\mu$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure.

Energy Resolution

The energy resolution for track- and shower-like events as a function of the true neutrino energy is shown in Figures 7.13 and 7.14. The reconstructed neutrino energy of Equation (6.3) is compared to using the track energy alone to determine the true neutrino energy. The track prefit gives a much better estimate of the neutrino energy compared to using the shower prefit. The reconstructed neutrino energy gives an improved (or very similar) estimate of the true neutrino energy compared to using the reconstructed track energy, with the improvement being quite apparent when the shower prefit is used. The worsening of the resolution at higher energies is likely from the track component leaving the detector (i.e. longer tracks at high energies), with the hits it produces not being fully contained.

The energy resolution for shower-like events as a function of the true neutrino energy is shown in Figures 7.15 and 7.16. Using the reconstructed neutrino energy of Equation (6.3) is compared to using the shower energy alone for these shower-like events. Using the reconstructed neutrino energy gives an improved estimate of the true neutrino energy for the case of the track prefit, for both neutrinos and anti-neutrinos. When the shower prefit is used, the reconstructed neutrino energy gives a slightly worse estimate of the true neutrino energy above $E_{\text{true}\,\nu} \sim 10$ GeV. Below this energy, the values are comparable between the two energy estimates.

For ν_e-CC events, the improvement in the energy resolution (comparing the reconstructed neutrino energy or shower energy) is greater than for ν_μ-NC events. This may be due to the hypothesis that the electromagnetic shower is treated as the track component, and the hadronic shower as the shower component (or vice versa, but this is in need of further investigation). However, the assumption of a track in the event model may lead to the worsening resolution at higher energies for ν_e-CC events. For the neutral-current events, which result in only a hadronic shower, ideally the track+shower reconstruction would reconstruct only the shower energy (with no reconstructed track energy), but currently there is a non-zero energy associated with the reconstructed track. The non-zero track energy in hadronic shower events is discussed again in Section 7.2.3.

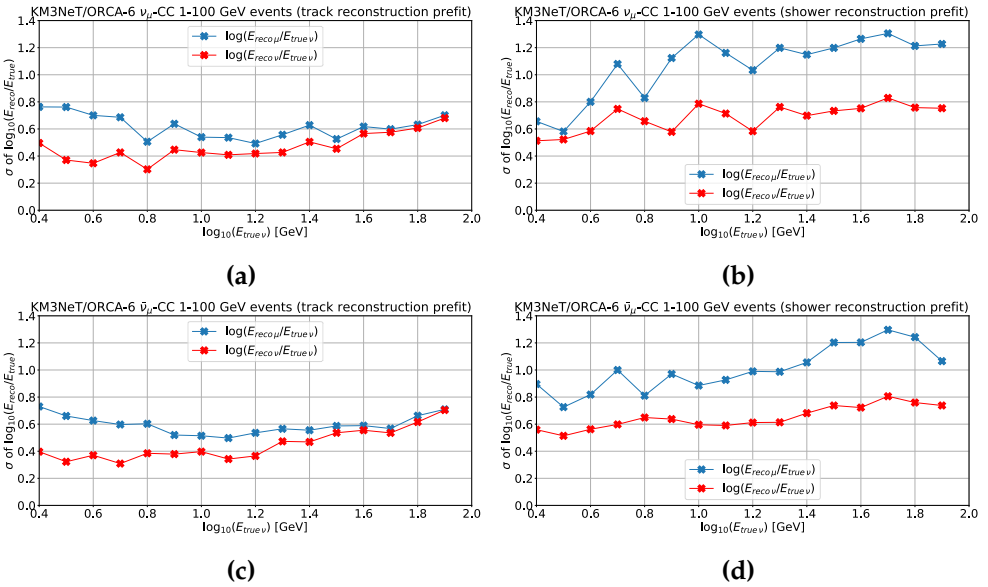


Figure 7.13.: The standard deviation σ of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ as a function of the true neutrino energy for charged-current ν_μ and $\bar{\nu}_\mu$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure. The values for the reconstructed track energy (blue) and reconstructed neutrino energy (red) are compared.

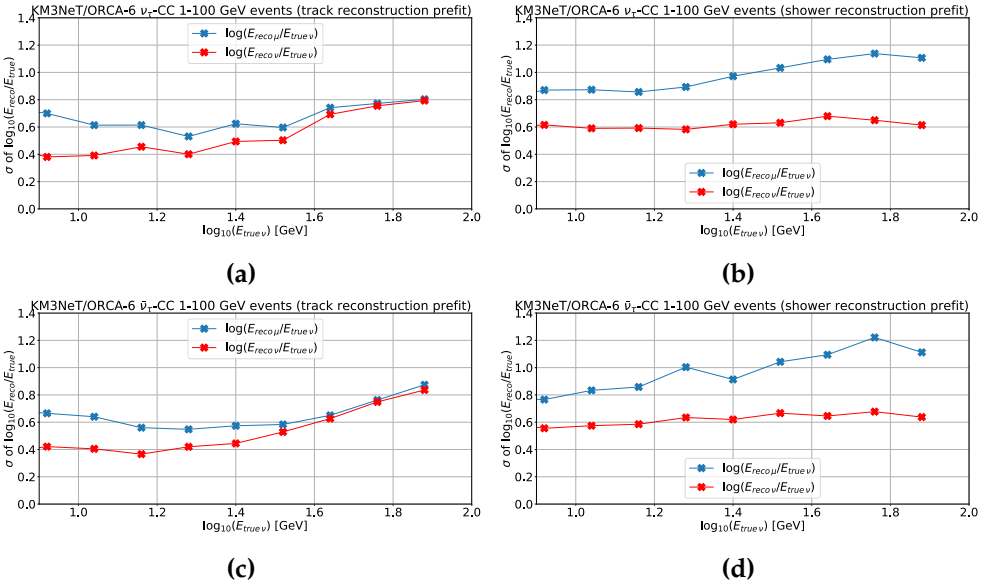


Figure 7.14.: The standard deviation σ of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ as a function of the true neutrino energy for charged-current ν_τ and $\bar{\nu}_\tau$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure. The values for the reconstructed track energy (blue) and reconstructed neutrino energy (red) are compared.

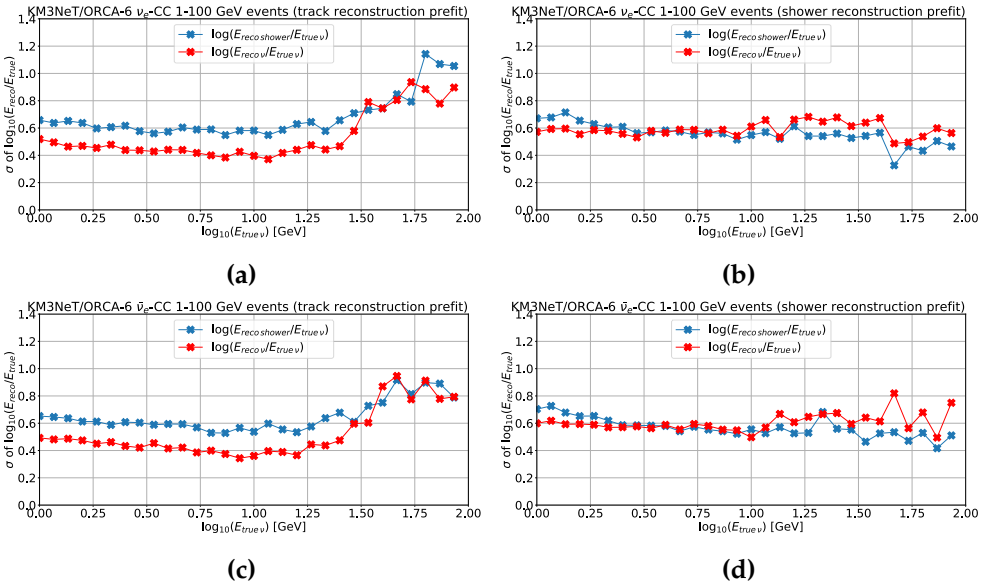


Figure 7.15.: The standard deviation σ of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ as a function of the true neutrino energy for charged-current ν_e and $\bar{\nu}_e$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure. The values for the reconstructed shower energy (blue) and reconstructed neutrino energy (red) are compared.

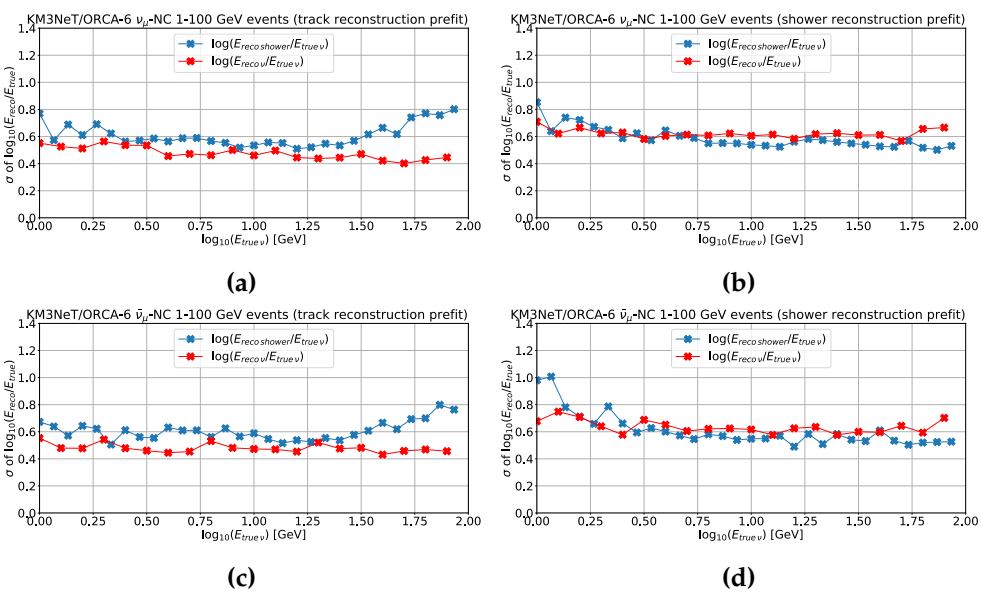
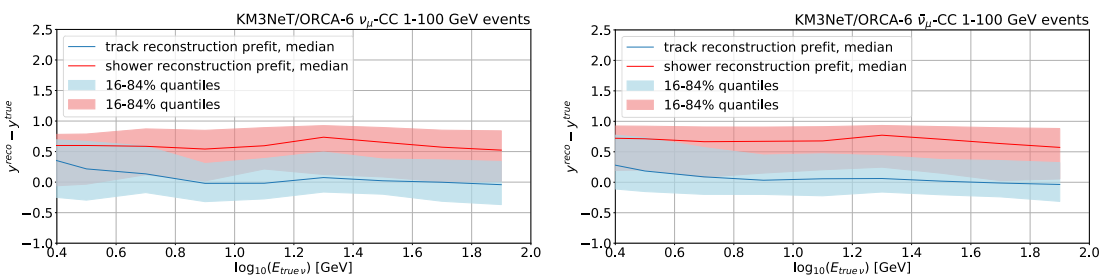


Figure 7.16.: The standard deviation σ of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ as a function of the true neutrino energy for neutral-current ν_μ and $\bar{\nu}_\mu$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure. The values for the reconstructed shower energy (blue) and reconstructed neutrino energy (red) are compared.

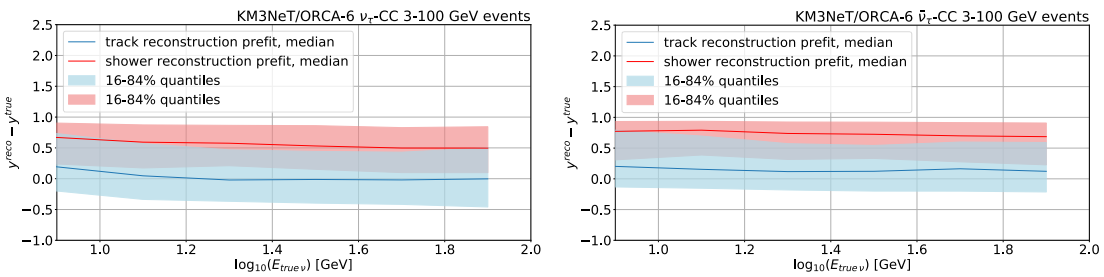
Bjorken-y Resolution

For track- and shower-like events, the difference between the true and reconstructed Bjorken-y as a function of the energy is presented in Figures 7.17 and 7.18 for neutrino and anti-neutrino events, for both prefits. The reconstructed Bjorken-y is defined using Equation (6.10), with the reconstructed track energy. The track reconstruction prefit gives a better estimate of the reconstructed Bjorken-y, with $y^{reco} - y^{true}$ closer to 0 following this prefit, and with a more narrow quantile (i.e. better resolution) as the energy increases. Using the shower prefit, the reconstructed shower energy and hence reconstructed y tends to be overestimated, resulting in positive values of $y^{reco} - y^{true}$. The distributions are slightly different between neutrino and anti-neutrino events, suggesting a potential to discriminate between the two.



(a) **(b)**

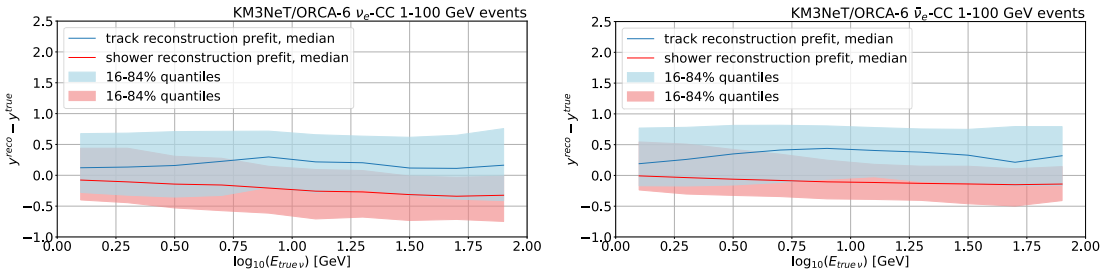
Figure 7.17.: The difference between the true and reconstructed Bjorken-y, as a function of the true neutrino energy, for charged-current ν_μ and $\bar{\nu}_\mu$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure. The median and quantiles when using the track reconstruction (blue) and shower reconstruction (red) to provide starting points to the fit are shown.



(a) **(b)**

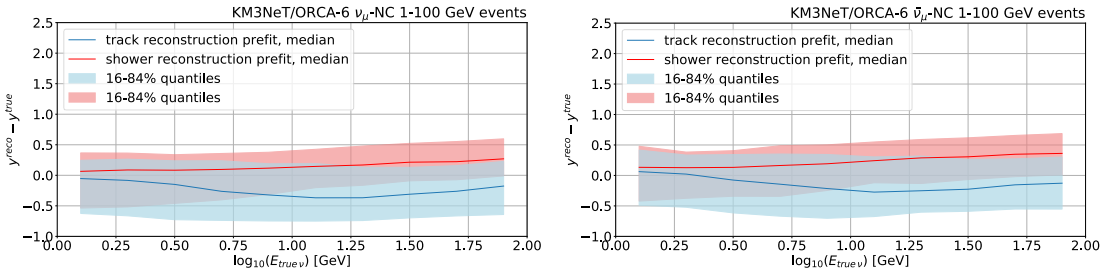
Figure 7.18.: The difference between the true and reconstructed Bjorken-y, as a function of the true neutrino energy, for charged-current ν_τ and $\bar{\nu}_\tau$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure. The median and quantiles when using the track reconstruction (blue) and shower reconstruction (red) to provide starting points to the fit are shown.

The value of $y^{reco} - y^{true}$ as a function of the energy is presented in Figures 7.19 and 7.20 for shower-like events. The reconstructed shower energy is used to define y^{reco}, instead of the track energy in Equation (6.10). Using the shower prefit results in a better resolution with increasing energy (as the quantiles band becomes more narrow). The track prefit results in a worse resolution. Again, there is a slight difference between neutrinos and anti-neutrinos.



(a) **(b)**

Figure 7.19.: The difference between the true and reconstructed Bjorken-y, as a function of the true neutrino energy, for charged-current ν_e and $\bar{\nu}_e$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure. The median and quantiles when using the track reconstruction (blue) and shower reconstruction (red) to proivde starting points to the fit are shown.



(a) **(b)**

Figure 7.20.: The difference between the true and reconstructed Bjorken-y, as a function of the true neutrino energy, for neutral-current ν_μ and $\bar{\nu}_\mu$ events simulated in the ORCA-6 detector, reconstructed with the track+shower reconstruction procedure. The median and quantiles when using the track reconstruction (blue) and shower reconstruction (red) to provide starting points to the fit are shown.

For ν_e-CC events, the track reconstruction prefit results in an overestimation of y^{true}, compared to the shower prefit wich results in an underestimation. Conversely, for ν_μ-NC the track prefit results in an underestimation of y^{true}, compared to an overestimated y^{true} using the shower prefit. In calculating y^{reco}, there is a non-zero track energy reconstructed by the track+shower reconstruction, for both event types. For the ν_e-CC events, one of the electromagnetic and hadronic components of the shower may be reconstructed as a track. For ν_μ-NC events, the track+shower reconstruction should ideally only reconstruct

a hadronic shower. This performance of the reconstructed Bjorken-y for the the two shower-like event classes warrants further investigation.

7.2.3. Summary of Reconstruction Performance

Some values of the presented median angular resolution and the standard deviation σ of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ are summarised in Table 7.3, for the different interaction channels and using both prefits. Median values of $y^{\text{reco}} - y^{\text{true}}$ are also given. Again, only the values for neutrinos are indicated, not for anti-neutrinos. These resolution values are shown for $E_\nu \sim 10$ GeV as an example value.

	$\Delta\theta^{\text{median}}_{\text{true }\nu,\text{ reco }\nu}$	$\sigma\log_{10}(E_{\text{reco }\nu}/E_{\text{true }\nu})$	$y^{\text{reco}} - y^{\text{true}}$
ν_μ-CC (track prefit)	18.8°	0.45	-0.02
ν_μ-CC (shower prefit)	28.6°	0.70	0.57
ν_e-CC (track prefit)	33.7°	0.31	0.25
ν_e-CC (shower prefit)	28.9°	0.54	-0.23
ν_τ-CC (track prefit)	37.1°	0.37	0.12
ν_τ-CC (shower prefit)	33.5°	0.59	0.62
ν_μ-NC (track prefit)	39.7°	0.38	-0.35
ν_μ-NC (shower prefit)	36.8°	0.60	0.14

Table 7.3.: A summary of the neutrino direction, neutrino energy, and Bjorken-y resolutions presented in this section for the track+shower reconstruction, at a true neutrino energy of $\sim$ 10 GeV. These resolution values are indicated for all interaction channels for the neutrino events, when using either the track or shower reconstruction as a prefit to the track+shower reconstruction.

Given the values noted in Table 7.3 and the performance of the track+shower reconstruction for the different neutrino and anti-neutrino interaction channels, some observations can be made.

The track+shower reconstruction algorithm, with the track prefit, is better at reconstructing the direction, energy and Bjorken-y of track- and shower-like (ν_μ-CC, ν_τ-CC) events, compared to using the shower prefit. For these events, using the reconstructed neutrino energy definition of Equation (6.3) gives an improved estimate of the true neutrino energy, compared to the reconstructed muon energy.

Using the shower reconstruction as a prefit results in better resolutions on the direction and Bjorken-y for shower-like (ν_e-CC, ν_μ-NC) events. When applying the track prefit on

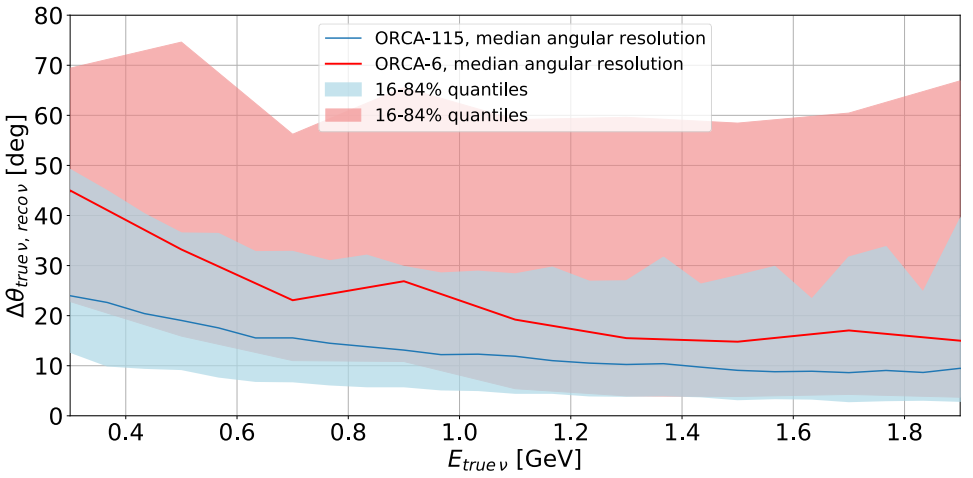
the shower-like events, using the reconstructed neutrino energy gives a better neutrino energy estimate compared to the shower energy. When using the shower prefit, the two energy estimates are very similar.

In estimating the true neutrino direction, for all of the above channels and for neutrinos and anti-neutrinos, the reconstructed neutrino direction does not give a much-improved estimate of the true neutrino direction, compared to using the direction of the track or shower components. For shower-like events, the two direction estimates are very similar in value. For track-like events, with the track prefit the muon direction estimate gives a starkly-improved angular resolution compared to the reconstructed neutrino estimate. This is also discussed in Chapter 6, where a similar conclusion is reached for muon neutrino events in the ORCA-115 line detector. When using the shower prefit for these track-like events, the two direction estimates are more comparable. An improved estimate of the neutrino energy is of major importance for an oscillation analysis (amongst others) with KM3NeT data. An improved energy estimate has the potential to improve the sensitivity of the experiment in measuring the neutrino oscillation parameters and determining the neutrino mass ordering.

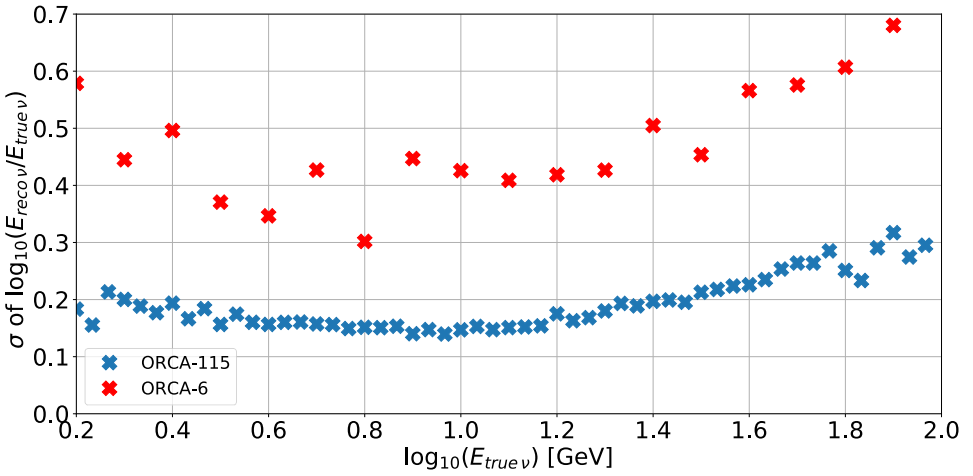
The resolutions presented above suggest a more optimal implementation of the starting points used in the track+shower reconstruction procedure. The shower prefit results in a better vertex estimate for the majority of the interaction channels, and also gives a better resolution on the shower direction. The track prefit results in a better estimate of the track direction and the track and shower energy. A combination of the prefits to provide starting points to the track+shower reconstruction algorithm may result in an improved performance: using the shower prefit for the vertex and shower direction parameters, and the track prefit for the track direction and and track and shower energies. It is also clear that using the reconstructed neutrino energy results in an improved estimate of the true neutrino energy, compared to the reconstructed track or shower energy alone. This observation holds for both track- and shower-like events, and shower-like events.

Comparison to the ORCA-115 Detector

The angular, energy, and Bjorken-y resolution of the track+shower reconstruction is compared between ORCA-115 and ORCA-6. This is done for ν_μ-CC events at energies of 1–100 GeV using the track reconstruction as a prefit. This neutrino interaction channel offers a benchmark for the performance in reconstructing events, since these events contains both a track and a shower component.



(a)



(b)

Figure 7.21.: The (a) angular resolution and (b) energy resolution achieved for the ORCA-115 (blue) and ORCA-6 (red) detectors, for ν_μ-CC events at energies of 1–100 GeV using the track reconstruction as a prefit.

The angular resolution achieved by the track+shower reconstruction is compared in Figure 7.21 (a) between the ORCA-6 and ORCA-115 detectors. The difference between the true and reconstructed neutrino direction is used, as a function of the neutrino energy. The angular resolution is better for the case of the detector with 115 detection units, with a much reduced quantiles band. At an energy of $E_\nu \sim$ 10 GeV, for example, the median angular resolution is $\sim 12^\circ$ and $\sim 23^\circ$ for the ORCA-115 and ORCA-6 detectors, respectively.

The energy resolution is shown in Figure 7.21 (b), where the standard deviation σ of $\log_{10}(E_{\text{reco}}/E_{\text{true}})$ is shown for the true and reconstructed neutrino energy, as a function of E_ν. Note that a different binning is used here between the two detectors, though

the trend is clear: the neutrino energy resolution is much improved for the case of the 115-line detector. At $E_\nu \sim$ 10 GeV, $\sigma \sim 0.15$ and 0.45 for the ORCA-115 and ORCA-6 detectors, respectively. The behaviour of the resolution as a function of the energy has been discussed previously in Chapters 5 and 6.

A comparison of the reconstructed Bjorken-y is made for the ORCA-115 and ORCA-6 detectors. In order to check events with a well-reconstructed Bjorken-y, the condition is added to the considered events that $E_{reco\,\nu} = 0.85\text{–}1.0 \times E_{true\,\nu}$, for the ν_μ-CC channel. This comparison is shown in Figure 7.22. From the 2D distribution of true and reconstructed y, in each bin of true y the mean value of the reconstructed y is taken, alongside the error on the mean value. For the ORCA-115 detector, more events lie closer to the diagonal line of $y^{true} = y^{reco}$ than in the ORCA-6 detector. Also for the ORCA-115 detector, events with higher values of y show a correlation between the true and reconstructed value. For these events, more energy goes into the hadronic shower. With improvements to the (energy) reconstruction of the hadronic shower, more events should be reconstructed with an even closer value to the true y value. For the ORCA-6 detector, there are much fewer events shown due to the limited number of statistics.

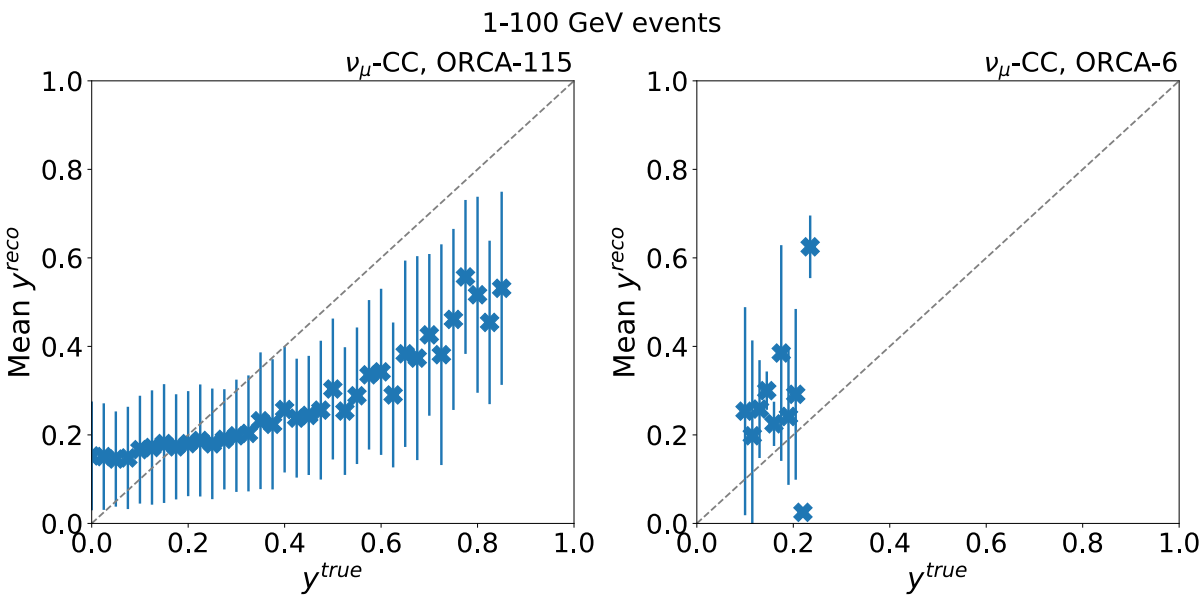


Figure 7.22.: The true and reconstructed y distributions are shown for ν_μ-CC events in the ORCA-115 detector (left) and the ORCA-6 detector (right). For every true y bin, the mean value of the reconstructed y is shown alongside the associated error.

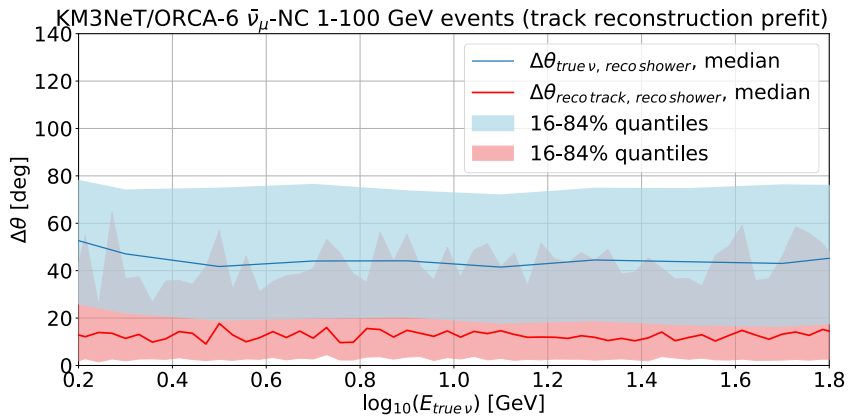
The better angular, energy, and Bjorken-y resolution in the ORCA-115 detector is likely due to the fact that with far more detection units and PMTs, there is more information available to determine the direction and energy of both the track and shower components of the events. A larger detector encloses the events more completely, and there is more

information available for reconstructing these events. The ORCA-6 detector has a much longer vertical reach than horizontal, making it difficult to detect the energy deposited by a track (or shower) traversing the entire detector (except diagonally). The ORCA-6 detector is also much smaller, making the reconstruction in this detector more susceptible to effects coming from the edge of the detector, where events are not contained. More optical modules clearly bring about a better ability to reconstruct the energies and thus the Bjorken-y parameter. The ORCA-6 detector is too small to do so, but with the addition of more detection units KM3NeT/ORCA becomes sensitive to the Bjorken-y with this track+shower reconstruction.

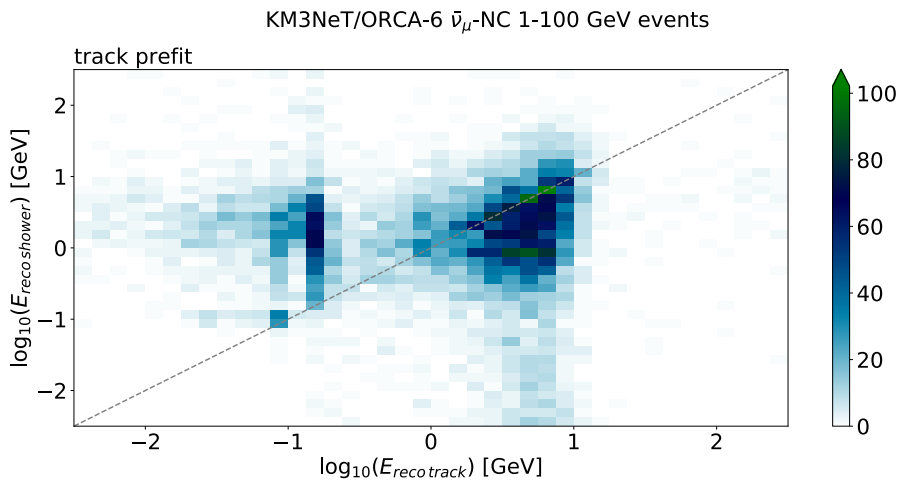
7.2.4. Limitations of the Reconstruction Performance

The reconstructed track and shower energies are shown in Appendix E for all the simulated neutrino channels, before any quality cuts are applied. As discussed already, the trend is also apparent that the prefit which is more suitable for the event signature results in a better reconstructed energy — that is, the track prefit yields more track and shower-like events reconstructed between 1–100 GeV, and comparatively for the shower prefit and shower-like events. Two distinctive features in the reconstructed energy distributions are many events reconstructed at $E_{reco} \sim$ 0.1 GeV and 0.003 GeV. A minimum threshold of 0.1 GeV is implemented in the track+shower minimisation strategy to fit the energies; some events remain at (or close to) this value after the fit procedure. The events with a reconstructed track or shower energy at $\sim$ 0.1 GeV tend to have true energies of a few GeV, and are reconstructed with values of $-\log\mathcal{L} <$ 1200. These low values of the likelihood may be related to not much hit information available to the fit. The large distribution of events at $\sim$0.003 GeV comes from the prefit reconstruction. A minimum threshold is also reached in the minimisation strategy for this reconstruction and some events are not fitted with an energy other than these values. Selection criteria, such as on $-\log\mathcal{L}$, can be used to exclude such events.

It is apparent from the reconstructed energy distributions and the above studies that a reconstructed track energy and direction are estimated for events in which there are no tracks. This is apparent in both ν_μ-NC and ν_e-CC events. In Figure 7.23 (a), the reconstructed track and shower directions are shown for $\bar{\nu}_\mu$-NC events, as an example set of events showing this issue. The angle between the true neutrino and reconstructed shower is shown in comparison to the angle between the reconstructed track and reconstructed shower direction. The track direction is reconstructed very close to the shower direction, much closer than the reconstructed shower differs from the neutrino direction, with a



(a)



(b)

Figure 7.23.: (a) The angular resolution of the true and reconstructed shower directions (blue) is shown alongside the angular resolution of the reconstructed track and shower directions (red). (b) The reconstructed shower energy is shown distributed as a function of the reconstructed track energy. Event interaction types of $\bar{\nu}_\mu$-NC are shown here to portray events which contain only a hadronic shower and no track component in the event interaction.

median angular deviation of $\sim 10^\circ$ across the entire neutrino energy range. The hits from the shower are used to estimate a track direction in the fit procedure. In Figure 7.23 (b), the reconstructed shower energy is shown compared to the reconstructed track energy for these events. Many events have a track energy reconstructed close to the shower energy, as indicated by the diagonal line. The track+shower reconstruction uses some of the hits induced by the hadronic shower to attempt to fit a track alongside the shower. The reconstructed track energy tends to be reconstructed with a value ≤ 10 GeV.

The non-zero reconstructed track energy for such hadronic shower events results in misleading values of the reconstructed Bjorken-y, as noted in Section 7.2.2. Improvements

to the reconstruction procedure should reduce the number of events mis-reconstructed as a track. In addition, the events which are reconstructed as a track for ν_e-CC should be investigated in a future study.

7.3. Discussion & Conclusion

The novel track+shower reconstruction algorithm has been applied to a sub-set of the KM3NeT/ORCA-6 data. The additional features provided by this reconstruction, such as the likelihood components of Equation (3.23), should offer additional input and discriminatory power to the particle identification step in analyses. Alongside the standard track or shower reconstruction methods, this reconstruction offers means of discriminating between track-like, shower-like, and atmospheric muon-like events. This can contribute to the particle identification stage of neutrino oscillation analyses.

A detailed overview of the performance of the reconstruction has been presented, categorised per neutrino interaction type and for both neutrino and anti-neutrino interactions. A summary of the angular resolution and energy resolution which can be expected for the track+shower reconstruction, applied to this data set, is given in Table 7.3. The results have been summarised, depending on the prefit used and the resolution differences have been outlined between neutrinos, anti-neutrinos and the prefit used per neutrino interaction channel.

The neutrino energy estimate is shown to be improved upon, compared to using the track or shower energy as a proxy for the neutrino energy. Improvements to the energy resolution are essential to improving the sensitivity of KM3NeT/ORCA to the neutrino oscillation parameters and the neutrino mass ordering. The Bjorken-y parameter is also directly estimated using the track+shower reconstruction. Although for the ORCA-6 detector the resulting resolution on y^{reco} is poor, with additional detection units the resolution will improve. An imperative study to improve the oscillation analyses with ORCA data will be using the (improved) reconstructed Bjorken-y parameter as an additional fit parameter to an oscillations analysis.

For future studies, the prefit starting points should be combined. One update that can be made to the reconstruction procedure is: using the shower prefit to provide the vertex, shower direction and shower energy starting points, and the track prefit for the track direction and track energy starting points.

With regards to the performance of the track+shower reconstruction, the reconstructed neutrino direction should not be used as the direction estimate of events. Reconstructing the direction of hadronic showers is quite a difficult task, as conveyed in this thesis. The reconstructed track or shower directions give a better estimate of the true neutrino direction compared to the reconstructed neutrino formulation of Equation (6.4). As mentioned previously, this formula neglects the four-momentum of the target, and the weights assigned to the shower and track directions are in need of further investigation.

The reconstructed neutrino energy should be used as the neutrino energy estimate in analyses with the ORCA data, as opposed to the reconstructed track or shower energies themselves. A reduced version of the track+shower fit should also be investigated, where the directions are not estimated but the vertex time and positions, and the track and shower energies are estimated. This would lead to a faster minimisation and provide a benchmark reconstruction which avoids difficult descriptions of the hadronic shower direction.

This reconstruction technique has been shown to improve the neutrino energy estimate and directly estimate a value of the Bjorken-y parameter for neutrino events. With further improvements to the reconstruction method, better estimates of these parameters are also expected. This in turn should lead to an improved sensitivity in measuring the neutrino mass ordering, which still eludes us. Hopefully, the development and use of this novel reconstruction algorithm, described in this thesis, is just the beginning of the story.

Appendix A.

Derivation of Hit Time Residuals

Consider a muon, originating from a point z_0, travelling in the z-direction. The muon travels a distance A, at which it emits a Cherenkov photon. The distance to the PMT from this point is denoted as D. The muon travels another distance, B, until it is at the point of closest approach to the PMT; this distance is denoted by R and this point is denoted by z. This geometry is shown in Figure 3.1.

The time for the light to reach the PMT will be

$$t_1 = t_0 + \frac{A}{c} + \frac{D}{c_{\text{water}}} = t_0 + \frac{A}{c} + \frac{D}{c/n} = t_0 + \frac{A}{c} + \frac{D \cdot n}{c}\,. \tag{A.1}$$

with n the refractive index of the medium, c_{water} the speed of light in water, and c the speed of light in a vacuum. With an emission angle of $\cos\theta = \frac{1}{n\cdot\beta}$, and $\beta \approx 1$, $cos\theta_C = \frac{1}{n}$.

Thus,

$$t_1 = t_0 + \frac{A}{c} + \frac{D}{\cos\theta_C \cdot c}\,. \tag{A.2}$$

The following relations can be made:

$$\sin\theta_C = \frac{R}{D}\,, \tag{A.3}$$

$$\cos\theta_C = \frac{B}{D}\,, \tag{A.4}$$

and

$$A = (z - B)\,. \tag{A.5}$$

Now, the arrival time becomes

$$t_1 = t_0 + \frac{z - D \cdot \cos\theta_C}{c} + \frac{R}{\sin\theta_C \cdot \cos\theta_C \cdot c} \tag{A.6}$$

$$= t_0 + \frac{z}{c} - \frac{R}{\sin\theta_C}\frac{\cos\theta_C}{c} + \frac{R}{\sin\theta_C \cdot \cos\theta_C \cdot c}\,. \tag{A.7}$$

Re-arranging:

$$t_1 = t_0 + \frac{z}{c} - \left(\frac{R \cdot \cos^2\theta_C + R}{\sin\theta_C \cdot \cos\theta_C \cdot c}\right) \tag{A.8}$$

$$= t_0 + \frac{z}{c} - \left(\frac{R(1 - \cos^2\theta_C)}{\sin\theta_C \cdot \cos\theta_C \cdot c}\right) \tag{A.9}$$

$$= t_0 + \frac{z}{c} - \left(\frac{R(\sin^2\theta_C)}{\sin\theta_C \cdot \cos\theta_C \cdot c}\right) \tag{A.10}$$

$$= t_0 + \frac{z}{c} - \left(\frac{R(\sin\theta_C)}{\cos\theta_C \cdot c}\right)\,. \tag{A.11}$$

This becomes

$$t_1 = t_0 + \frac{z}{c} - \left(\frac{R \cdot \tan\theta_C}{c}\right)\,. \tag{A.12}$$

For a muon, a reference expected arrival time to use for the time residual definition is where the muon emits light directly at the point where the light only travels to the PMT. In this case, the first two terms of Equation (A.12) are not necessary and

$$t_\mu^{\text{expected}} = \left(\frac{R \cdot \tan\theta_C}{c}\right)\,. \tag{A.13}$$

For the case of a particle shower, again the distance D between the light emission point and the PMT can be considered, rather than the road-width. From Equation (A.1), a reference expected arrival time for light from a shower is

$$t_{\text{shower}}^{\text{expected}} = \frac{D \cdot n}{c}\,. \tag{A.14}$$

Appendix B.

Derivation of Hit Time Relation

We define two relations:

$$t'_j = \frac{t_j c}{\tan(\theta_c)} - \frac{z_j}{\tan(\theta_c)}\,, \tag{B.1}$$

and

$$t'_0 = \frac{t_0 c}{\tan(\theta_c)}\,. \tag{B.2}$$

We can re-arrange Equation (B.1) and Equation (B.2) to give:

$$t_j = \frac{t'_j \tan(\theta_c)}{c} + \frac{z_j}{c}\,, \tag{B.3}$$

and

$$t_0 = \frac{t'_0 \tan(\theta_c)}{c}\,. \tag{B.4}$$

Substituting these into Equation (3.32) gives:

$$\frac{t'_j \tan(\theta_c)}{c} + \frac{z_j}{c} = \frac{t'_0 \tan(\theta_c)}{c} + \frac{z_j}{c} + \frac{\tan(\theta_c) R_j}{c}\,, \tag{B.5}$$

$$\implies t'_j = t'_0 + R_j \,. \tag{B.6}$$

With

$$t'_j - t'_0 = R_j \,, \tag{B.7}$$

and using Equation (3.33),

$$\implies t'_j - t'_0 = \sqrt{(x_j - x_0)^2 + (y_j - y_0)^2} \,, \tag{B.8}$$

$$\implies \left(t'_j - t'_0\right)^2 = (x_j - x_0)^2 + (y_j - y_0)^2 \,. \tag{B.9}$$

Thus

$$\begin{aligned} t'^2_j - 2t'_j t'_0 + t'^2_0 = x_j^2 - 2x_j x_0 + x_0^2 \\ + y_j^2 - 2y_j y_0 + y_0^2 \,. \end{aligned} \tag{B.10}$$

Following the same logic for a hit i on the PMT with expected arrival time t_i and distance of closest approach R_i gives

$$\begin{aligned} t'^2_i - 2t'_i t'_0 + t'^2_0 = x_i^2 - 2x_i x_0 + x_0^2 \\ + y_i^2 - 2y_i y_0 + y_0^2 \,. \end{aligned} \tag{B.11}$$

Subtracting Equation B.11 from Equation B.10 leaves us with

$$\begin{aligned} t'^2_j - t'^2_i - 2t'_j t'_0 + 2t'_i t'_0 = x_j^2 - x_i^2 - 2x_j x_0 + 2x_i x_0 \\ + y_j^2 - y_i^2 - 2y_j y_0 + 2y_i y_0 \,. \end{aligned} \tag{B.12}$$

Appendix C.

List of MUPAGE Parameters

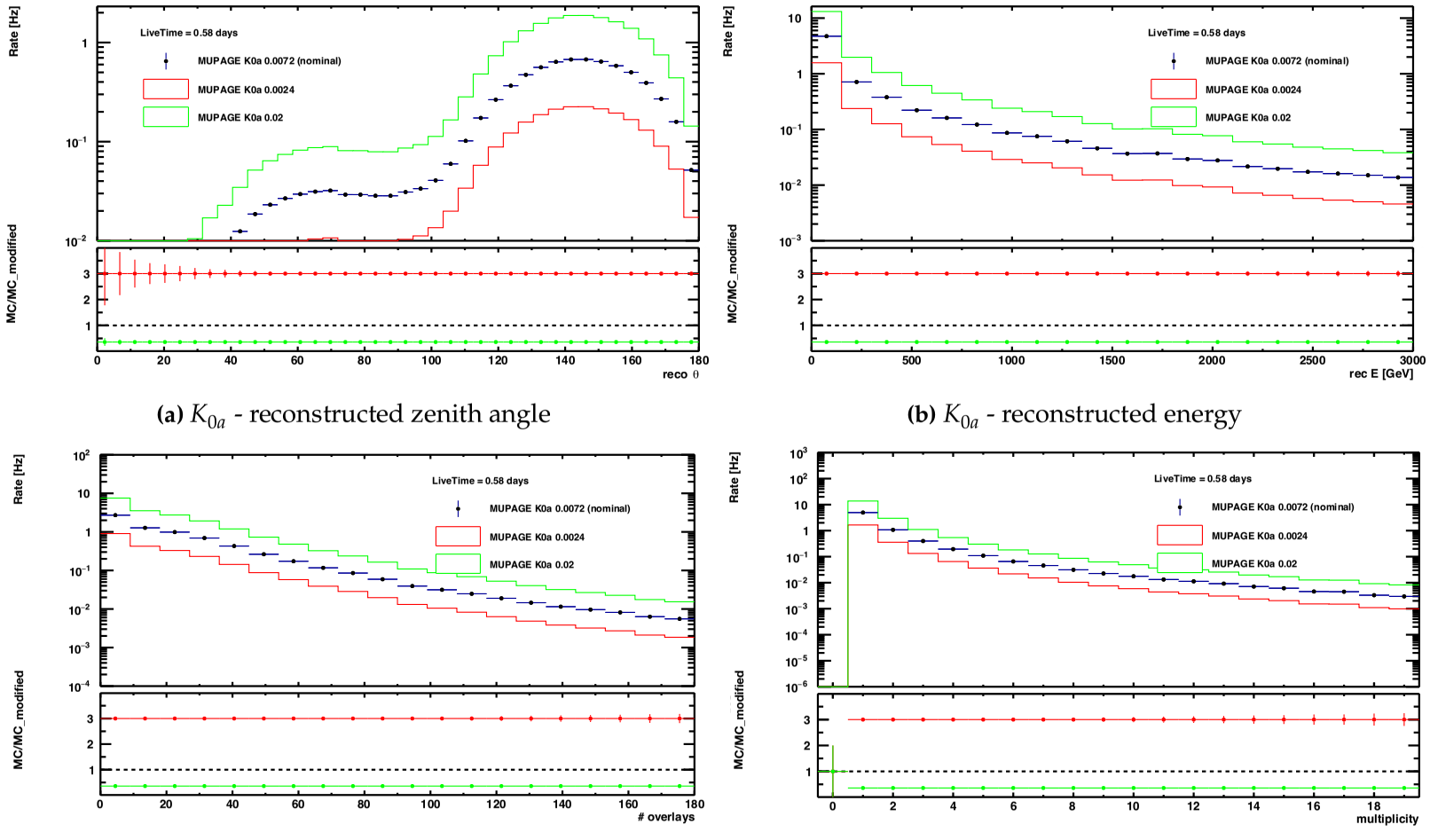
The **default** values for the parameters within the parametric equations are listed in Table C.1, in the syntax used within MUPAGE.

Parameter	Value	Parameter	Value
Flux		*Lateral Spread*	
K0a	0.0072	RHO0a	-1.786
K0b	-1.927	RHO0b	28.26
K1a	-0.581	RHO1	-1.06
K1b	0.034	THETA0	1.3
ni0a	-0.0771	Fpiccolo	10.4
ni0b	0.524	ALPHA0a	-0.448
ni0c	2.068	ALPHA0b	4.969
ni1a	0.03	ALPHA1a	0.0194
ni1b	0.470	ALPHA1b	0.276
Energy - Single Muons		*Energy - Multiple Muons*	
BETA	0.420	A0	0.0033
g0	-0.232	A1	0.0079
g1	3.961	B0a	0.0407
e0a	0.0304	B0b	0.0283
e0b	0.359	B1a	-0.312
e1a	-0.0077	B1b	6.124
e1b	0.659	Q0	0.0543
		Q1	-0.365
		C0a	-0.069
		C0b	0.488
		C1	-0.117
		D0a	-0.398
		D0b	3.955
		D1a	0.012
		D1b	-0.35

Table C.1.: Default values of the parameters used within parametric equations in MUPAGE, from Carminati, Margiotta, and Spurio (2008) [105].

Appendix D.

Impact of MUPAGE Parameters on Observables



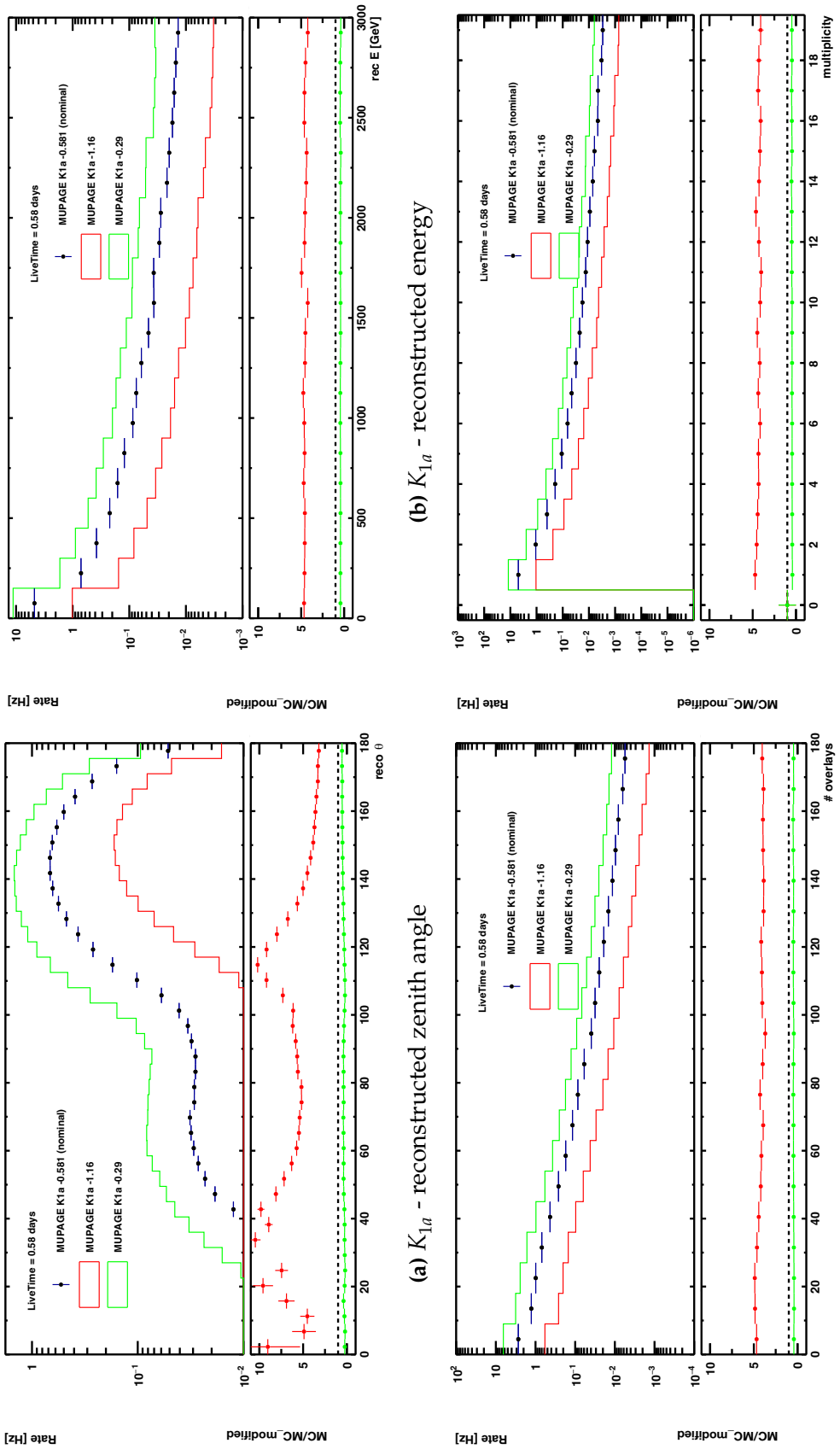
(a) K_{0a} - reconstructed zenith angle

(b) K_{0a} - reconstructed energy

(c) K_{0a} - overlays

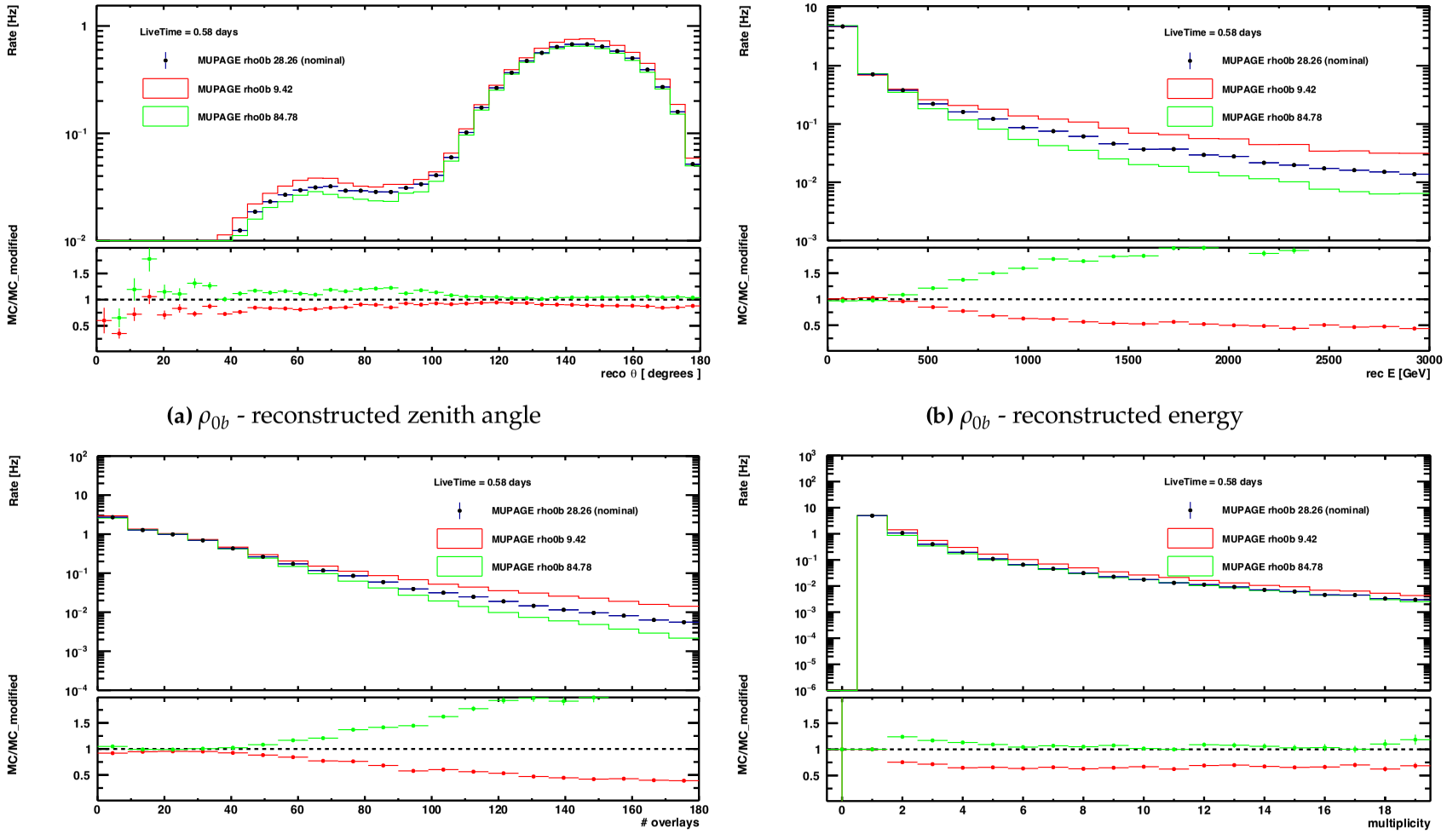
(d) K_{0a} - multiplicity

Figure D.1.: The effect of varying K_{0a} on the distributions of the (a) reconstructed zenith angle (b) reconstructed energy (c) number of overlays (d) multiplicity. Changing this parameter results in a scaling of the distributions, with no effect on the shape.



(a) K_{1a} - reconstructed zenith angle **(b)** K_{1a} - reconstructed energy **(c)** K_{1a} - overlays **(d)** K_{1a} - multiplicity

Figure D.2.: The effect of varying K_{1a} on the distributions of the (a) reconstructed zenith angle (b) reconstructed energy (c) number of overlays (d) multiplicity. This parameter affects the shape of these distributions, in particular that of the reconstructed zenith angle.



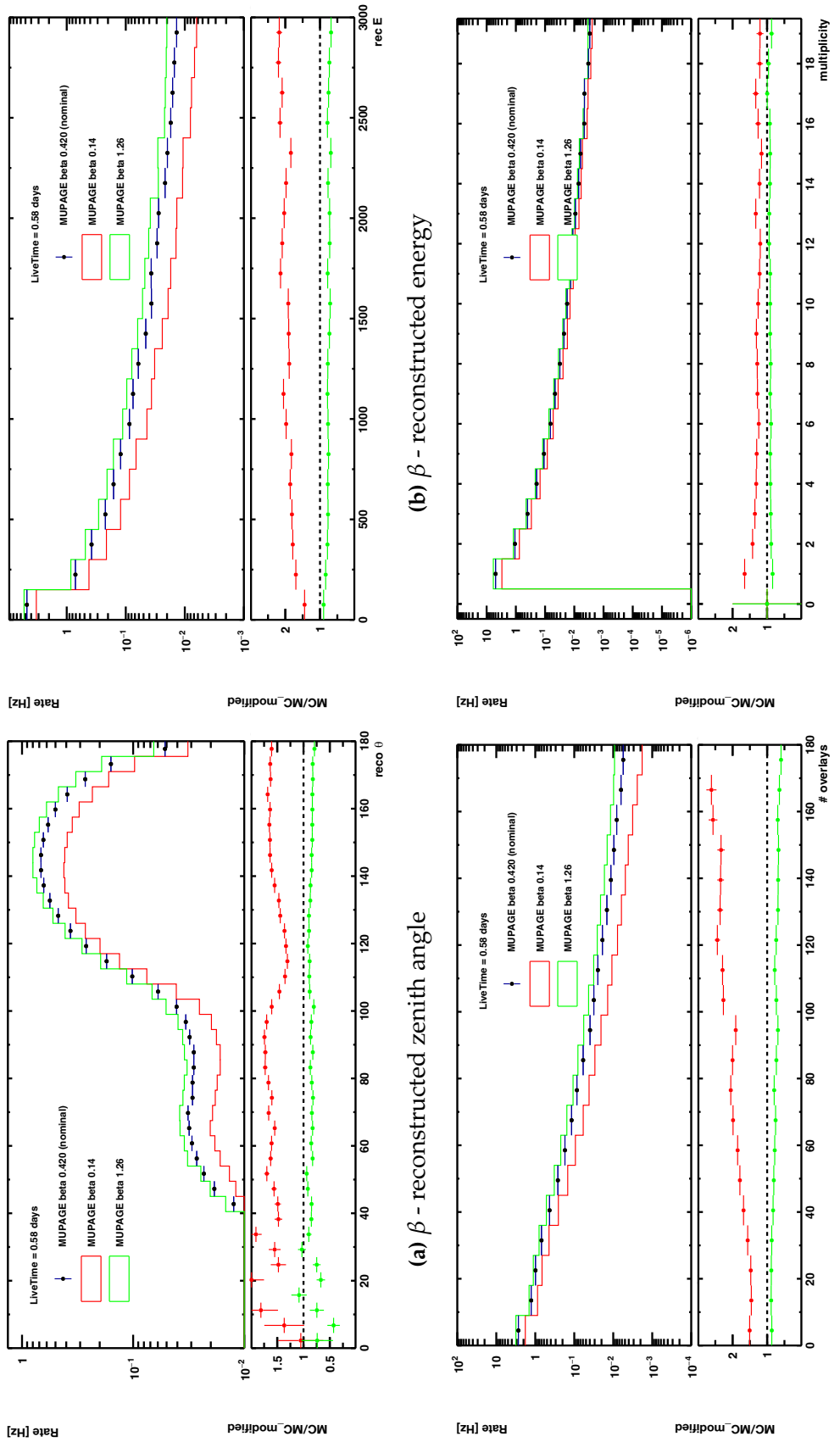
(a) ρ_{0b} - reconstructed zenith angle

(b) ρ_{0b} - reconstructed energy

(c) ρ_{0b} - overlays

(d) ρ_{0b} - multiplicity

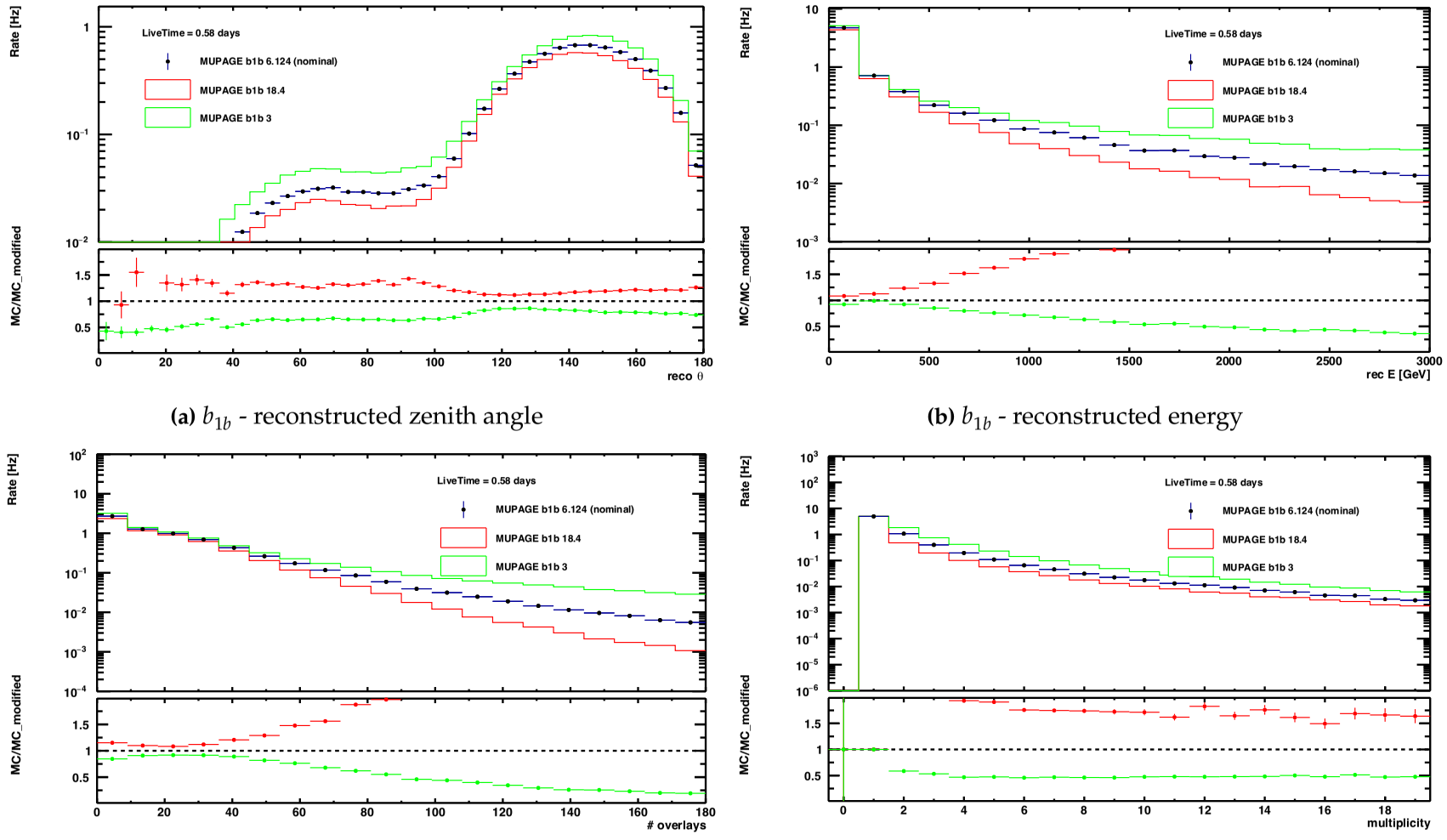
Figure D.3.: The effect of varying ρ_{0b} on the distributions of the (a) reconstructed zenith angle (b) reconstructed energy (c) number of overlays (d) multiplicity. While little effect is induced on the reconstructed zenith and multiplicity distribution, altering this parameter value affects the reconstructed energy and the number of overlays, decreasing or increasing the slope of these distributions.



(a) β - reconstructed zenith angle **(b)** β - reconstructed energy

(c) β - overlays **(d)** β - multiplicity

Figure D.4.: The effect of varying β on the distributions of the (a) reconstructed zenith angle (b) reconstructed energy (c) number of overlays (d) multiplicity. Altering this parameter value is shown to affect all the distributions, least so the multiplicity. Decreasing the value of β has the greatest effect on the distributions.



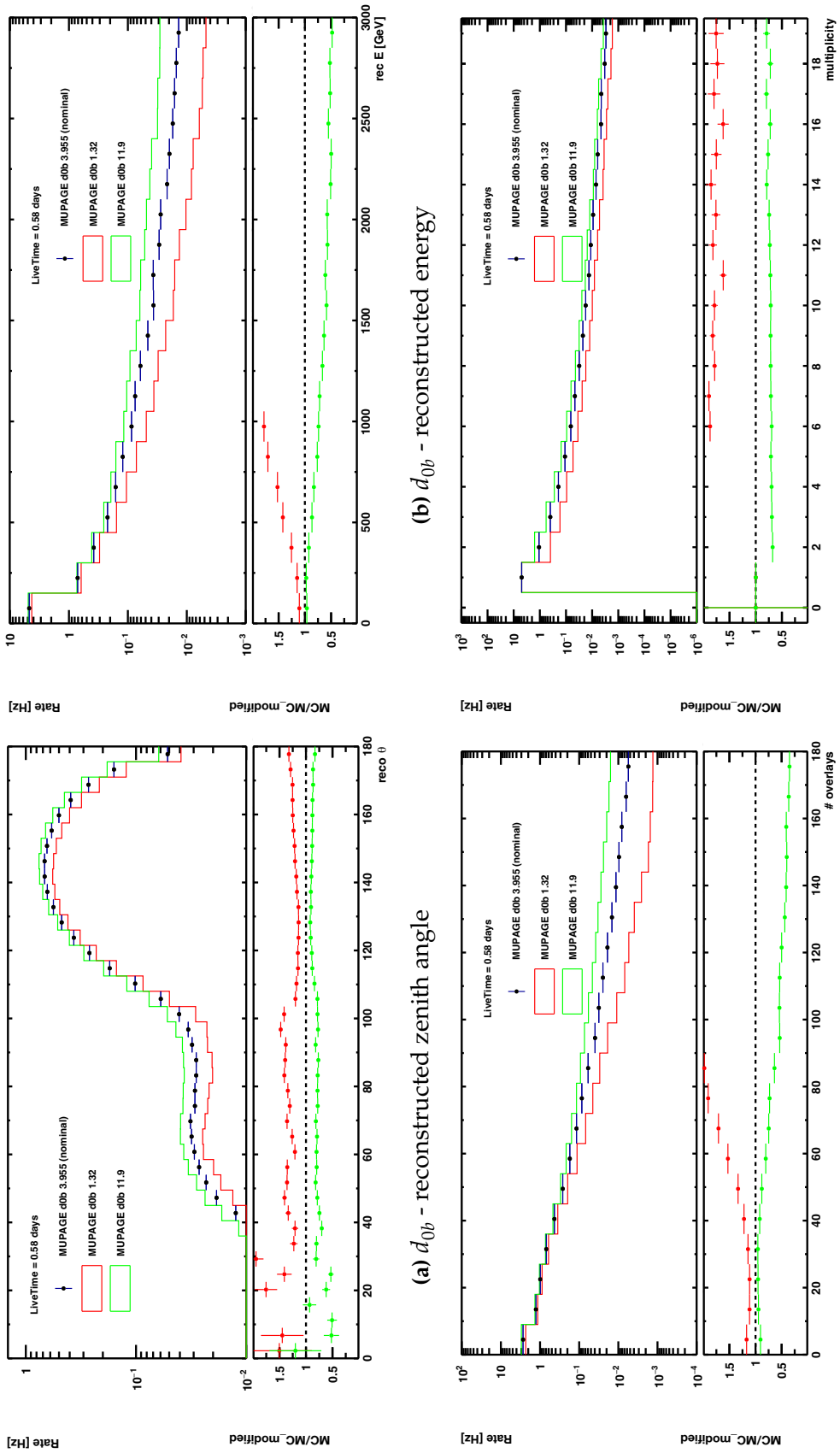
(a) b_{1b} - reconstructed zenith angle

(b) b_{1b} - reconstructed energy

(c) b_{1b} - overlays

(d) b_{1b} - multiplicity

Figure D.5.: The effect of varying b_{1b} on the distributions of the (a) reconstructed zenith angle (b) reconstructed energy (c) number of overlays (d) multiplicity. Varying this parameter value from the nominal affects all distributions, in particular changing the slope of the reconstructed energy and number of overlays.



(a) d_{0b} - reconstructed zenith angle **(b)** d_{0b} - reconstructed energy

(c) d_{0b} - overlays **(d)** d_{0b} - multiplicity

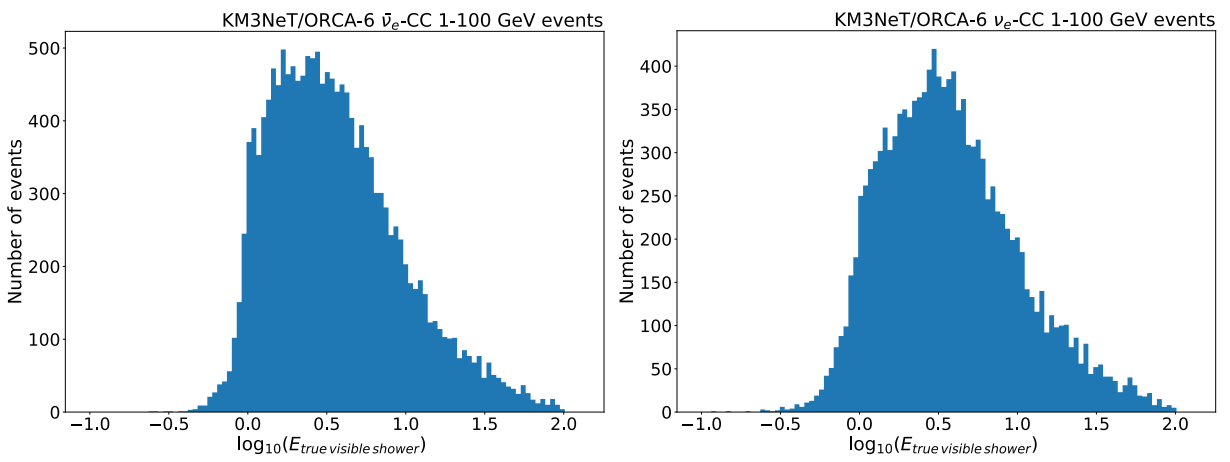
Figure D.6.: The effect of varying d_{0b} on the distributions of the (a) reconstructed zenith angle (b) reconstructed energy (c) number of overlays (d) multiplicity. Varying this parameter value from the nominal affects all distributions, with the zenith distribution the least affected.

Appendix E.

True & Reconstructed Distributions in the ORCA-6 detector

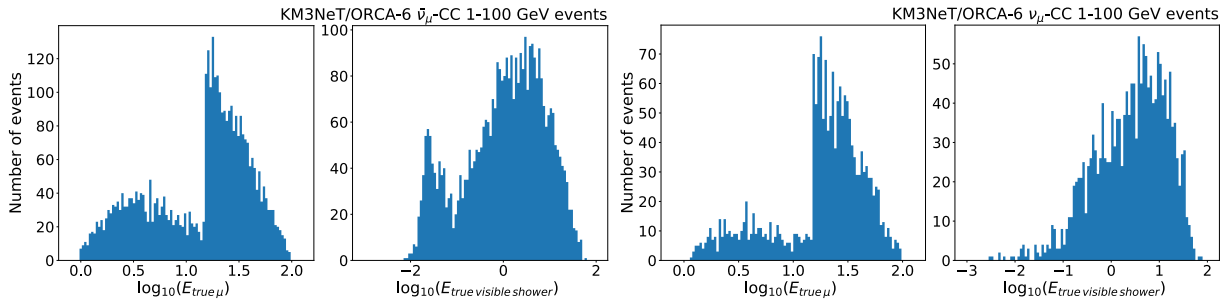
E.1. True Energy Distributions

The distributions of the true energies of the track and shower events are presented here, for the four interaction channels that are simulated for the ORCA-6 detector.



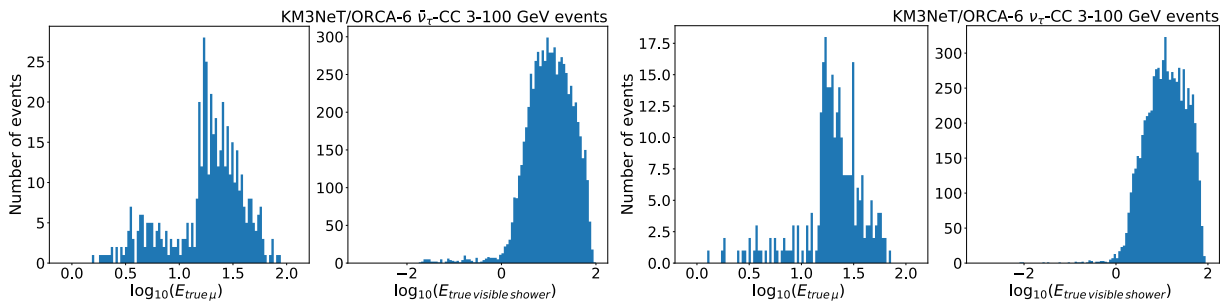
(a) **(b)**

Figure E.1.: The true energy of the visible showers for the ν_e-CC and $\bar{\nu}_e$-CC events used to evaluate the track+shower reconstruction performance.



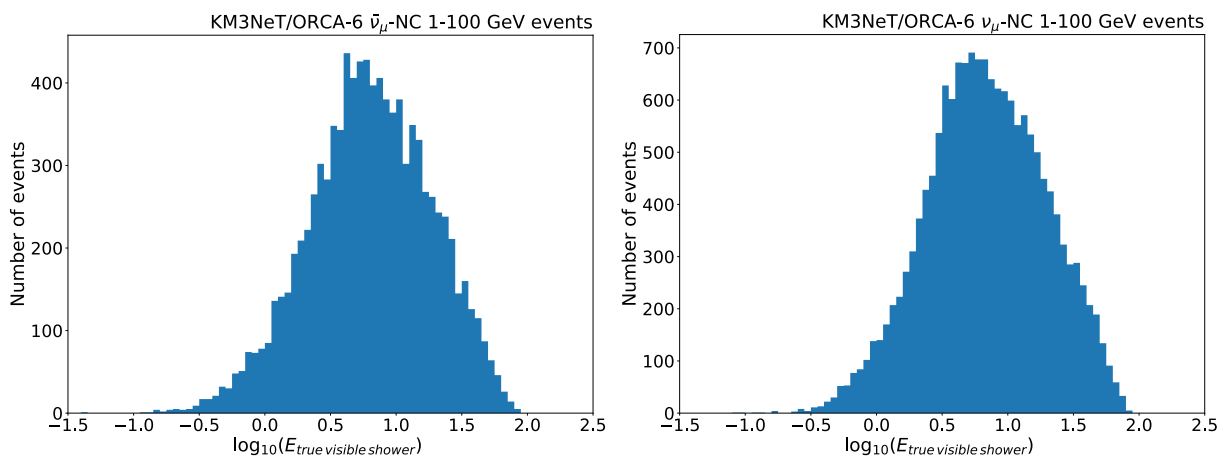
(a) **(b)**

Figure E.2.: The true energy of the muon tracks and visible showers for the ν_μ-CC and $\bar{\nu}_\mu$-CC events used to evaluate the track+shower reconstruction performance.



(a) **(b)**

Figure E.3.: The true energy of the muon tracks and visible showers for the ν_τ-CC and $\bar{\nu}_\tau$-CC events used to evaluate the track+shower reconstruction performance.



(a) **(b)**

Figure E.4.: The true energy of the visible showers for the ν_μ-NC and $\bar{\nu}_\mu$-NC events used to evaluate the track+shower reconstruction performance.

E.2. Reconstructed Energy Distributions

The reconstructed energy distributions of the track and shower components are presented here, when using both the track or shower reconstruction prefit, for the four neutrino interaction channels. No cuts have been applied for these distributions, in order to show the output of the track+shower reconstruction for all reconstructed events.

In the following figures, the shape of the energy distributions are very similar when the same prefit is used, i.e. the distributions appear similar in shape for most channels when the track or shower prefit is used.

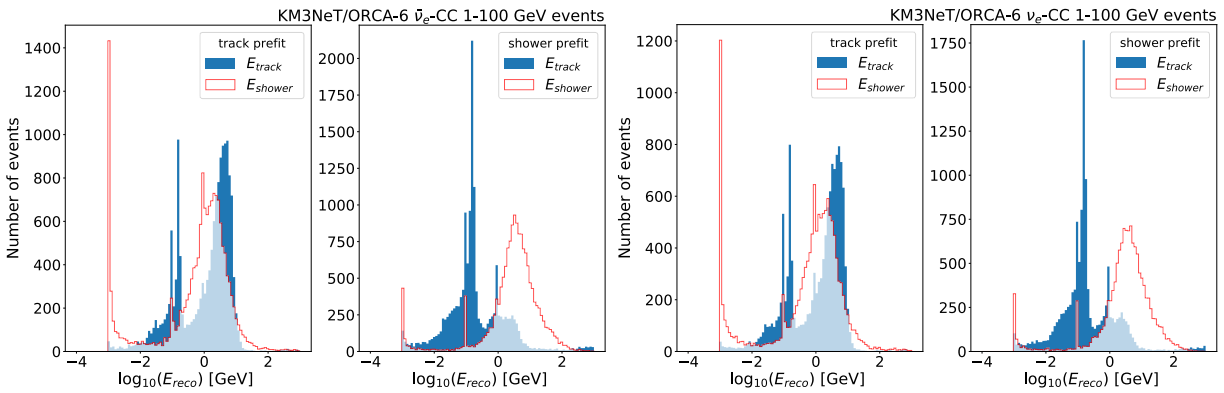


Figure E.5.: The reconstructed energy of the visible showers for the ν_e-CC and $\bar{\nu}_e$-CC events, as resulting from the track+shower reconstruction to ORCA-6 data.

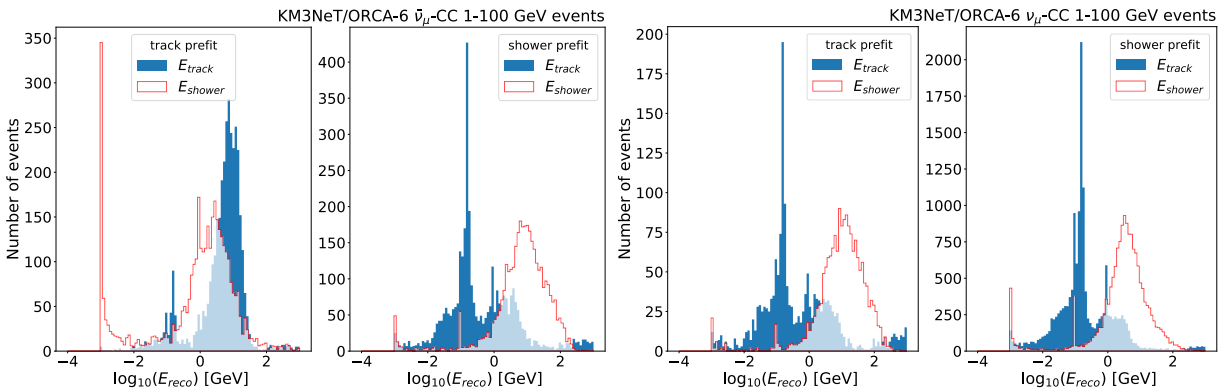
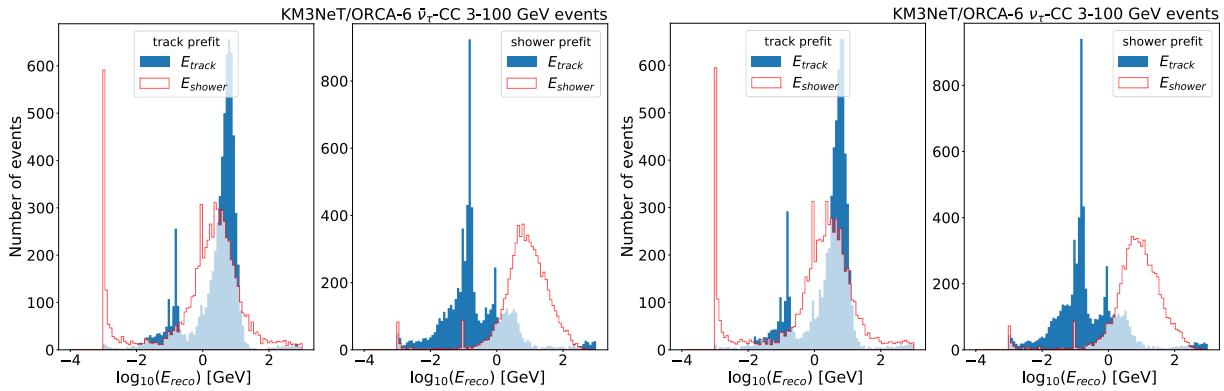
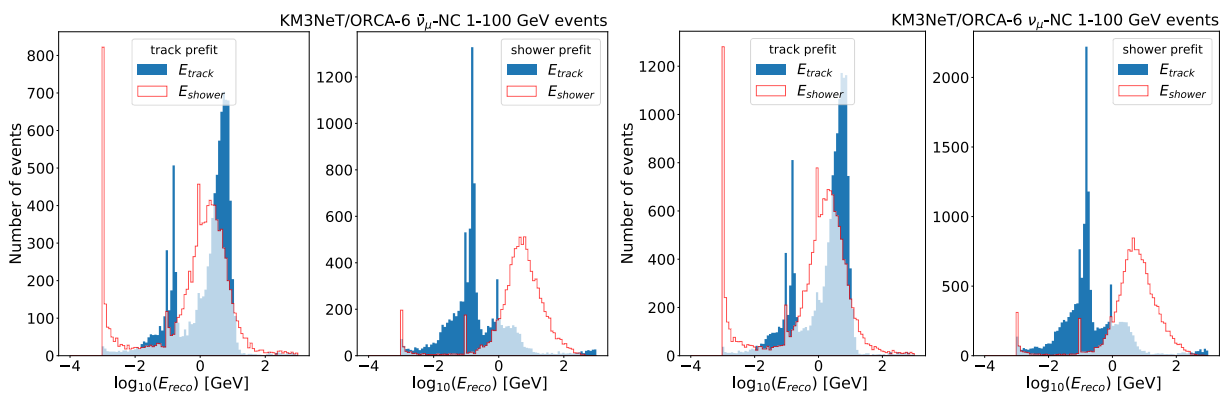


Figure E.6.: The reconstructed energy of the muon tracks and visible showers for the ν_μ-CC and $\bar{\nu}_\mu$-CC events, as resulting from the track+shower reconstruction to ORCA-6 data.



(a) **(b)**

Figure E.7.: The reconstructed energy of the muon tracks and visible showers for the ν_τ-CC and $\bar{\nu}_\tau$-CC events, as resulting from the track+shower reconstruction to ORCA-6 data.



(a) **(b)**

Figure E.8.: The reconstructed energy of the visible showers for the ν_μ-NC and $\bar{\nu}_\mu$-NC events, as resulting from the track+shower reconstruction to ORCA-6 data.

Bibliography

[1] D. V. Reames, "Solar energetic particles," *Lecture Notes in Physics* (2021) . http://dx.doi.org/10.1007/978-3-030-66402-2. (Cited on page 6)

[2] P. Tinyakov, "Course 10 - ultra-high energy cosmic rays," in *Particle Physics Beyond the Standard Model*, D. Kazakov, S. Lavignac, and J. Dalibard, eds., vol. 84 of *Les Houches*, pp. 537–572. Elsevier, 2006. https://www.sciencedirect.com/science/article/pii/S0924809906800302. (Cited on page 6)

[3] **Particle Data Group** Collaboration, R. L. Workman and Others, "Review of Particle Physics," *PTEP* **2022** (2022) 083C01. (Cited on page 7, 8, 18, 33, 39, 42, 65, 118, 170)

[4] P. Blasi, "The Origin of Galactic Cosmic Rays," *Astron. Astrophys. Rev.* **21** (2013) 70, arXiv:1311.7346 [astro-ph.HE]. (Cited on page 7)

[5] **KASCADE Grande** Collaboration, W. D. Apel *et al.*, "Kneelike structure in the spectrum of the heavy component of cosmic rays observed with KASCADE-Grande," *Phys. Rev. Lett.* **107** (2011) 171104, arXiv:1107.5885 [astro-ph.HE]. (Cited on page 7)

[6] K. van Dam, *HiSPARC A Scintillating Experiment*. PhD thesis, University of Twente, 2021. (Cited on page 7)

[7] V. Hess, "On the Observations of the Penetrating Radiation during Seven Balloon Flights," arXiv:1808.02927 [physics.hist-ph]. (Cited on page 9)

[8] P. Auger, P. Ehrenfest, R. Maze, J. Daudin, and R. A. Fréon, "Extensive cosmic-ray showers," *Rev. Mod. Phys.* **11** (Jul, 1939) 288–291. https://link.aps.org/doi/10.1103/RevModPhys.11.288. (Cited on page 9)

[9] **Pierre Auger** Collaboration, J. Abraham *et al.*, "Properties and performance of the prototype instrument for the Pierre Auger Observatory," *Nucl. Instrum. Meth. A* **523** (2004) 50–95. (Cited on page 9)

[10] **LHAASO** Collaboration, G. Di Sciascio, "The LHAASO experiment: from Gamma-Ray Astronomy to Cosmic Rays," *Nucl. Part. Phys. Proc.* **279-281** (2016) 166–173, arXiv:1602.07600 [astro-ph.HE]. (Cited on page 9)

[11] **Telescope Array** Collaboration, T. Abu-Zayyad *et al.*, "The surface detector array of the Telescope Array experiment," *Nucl. Instrum. Meth. A* **689** (2013) 87–97, arXiv:1201.4964 [astro-ph.IM]. (Cited on page 9)

[12] **AMS** Collaboration, A. Kounine, "AMS Experiment on the International Space Station," *32nd International Cosmic Ray Conference* **c** (2011) 5. (Cited on page 9)

[13] **Fermi-LAT** Collaboration, W. B. Atwood *et al.*, "The Large Area Telescope on the Fermi Gamma-ray Space Telescope Mission," *Astrophys. J.* **697** (2009) 1071–1102, arXiv:0902.1089 [astro-ph.IM]. (Cited on page 9)

[14] **HAWC** Collaboration, R. Alfaro *et al.*, "All-particle cosmic ray energy spectrum measured by the HAWC experiment from 10 to 500 TeV," *Phys. Rev. D* **96** no. 12, (2017) 122001, arXiv:1710.00890 [astro-ph.HE]. (Cited on page 9)

[15] E. Lorenz, "The MAGIC telescope project for gamma-ray astronomy in the 15-GeV to 300-GeV energy range," *Nucl. Phys. B Proc. Suppl.* **48** (1996) 494–496. (Cited on page 9)

[16] A. Kohnle, J. Mattes, G. Hermann, W. Hofmann, and M. Panter, "Photodetectors for HESS," *Nucl. Instrum. Meth. A* **442** (2000) 322–326. (Cited on page 9)

[17] **KM3NeT** Collaboration, S. Adrian-Martinez *et al.*, "Letter of intent for KM3NeT 2.0" *J. Phys. G* **43** no. 8, (2016) 084001, arXiv:1601.07459 [astro-ph.IM]. (Cited on page 9, 26, 35, 37, 42, 43, 45, 49, 53, 54, 55, 122, 150)

[18] **IceCube** Collaboration, M. G. Aartsen *et al.*, "The IceCube Neutrino Observatory: Instrumentation and Online Systems," *JINST* **12** no. 03, (2017) P03012, arXiv:1612.05093 [astro-ph.IM]. (Cited on page 9)

[19] W. Pauli, "Dear radioactive ladies and gentlemen," *Phys. Today* **31N9** (1978) 27. (Cited on page 10)

[20] E. Fermi, "Tentativo di una teoria dell'emissione dei raggi beta," *Ric. Sci.* **4** (1933) 491–495. (Cited on page 10, 16)

[21] H. Bethe and R. Peierls, "The 'neutrino'," *Nature* **133** (1934) 532. (Cited on page 11)

[22] R. Davis, Jr., "Attempt to detect the antineutrinos from a nuclear reactor by the Cl37(anti-nu,e-) A37 reaction," *Phys. Rev.* **97** (1955) 766–769. (Cited on page 11)

[23] C. L. Cowan, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire, "Detection of the free neutrino: A Confirmation," *Science* **124** (1956) 103–104. (Cited on page 11)

[24] F. Reines and C. L. Cowan, "The Neutrino," *Nature* **178** (1956) 446–449.

[25] M. Goldhaber, L. Grodzins, and A. W. Sunyar, "Helicity of Neutrinos," *Phys. Rev.* **109** (1958) 1015–1017. (Cited on page 11)

[26] B. Pontecorvo, "Electron and Muon Neutrinos," *Zh. Eksp. Teor. Fiz.* **37** (1959) 1751–1757. (Cited on page 11)

[27] G. Danby, J.-M. Gaillard, K. Goulianos, L. M. Lederman, N. Mistry, M. Schwartz, and J. Steinberger, "Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos," *Phys. Rev. Lett.* **9** (Jul, 1962) 36–44. `https://link.aps.org/doi/10.1103/PhysRevLett.9.36`. (Cited on page 11)

[28] M. L. Perl *et al.*, "Evidence for Anomalous Lepton Production in e+ - e- Annihilation," *Phys. Rev. Lett.* **35** (1975) 1489–1492. (Cited on page 11)

[29] **DONUT** Collaboration, K. Kodama *et al.*, "Observation of tau neutrino interactions," *Phys. Lett. B* **504** (2001) 218–224, `arXiv:hep-ex/0012035`. (Cited on page 11)

[30] H. Georgi and S. L. Glashow, "Unified weak and electromagnetic interactions without neutral currents," *Phys. Rev. Lett.* **28** (1972) 1494. (Cited on page 12)

[31] S. L. Glashow and S. Weinberg, "Natural Conservation Laws for Neutral Currents," *Phys. Rev. D* **15** (1977) 1958. (Cited on page 12)

[32] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc. C* **680519** (1968) 367–377. (Cited on page 12)

[33] S. Weinberg, "A model of leptons," *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266. `https://link.aps.org/doi/10.1103/PhysRevLett.19.1264`. (Cited on page 12)

[34] G. 't Hooft and M. J. G. Veltman, "Regularization and Renormalization of Gauge Fields," *Nucl. Phys. B* **44** (1972) 189–213. (Cited on page 12)

[35] **ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group** Collaboration, S. Schael *et al.*, "Precision electroweak measurements on the Z resonance," *Phys. Rept.* **427** (2006) 257–454, `arXiv:hep-ex/0509008`. (Cited on page 12, 13)

[36] P. Janot and S. Jadach, "Improved Bhabha cross section at LEP and the number of light neutrino species," *Phys. Lett. B* **803** (2020) 135319, arXiv:1912.02067 [hep-ph]. (Cited on page 12)

[37] J. A. Formaggio and G. P. Zeller, "From eV to EeV: Neutrino Cross Sections Across Energy Scales," *Rev. Mod. Phys.* **84** (2012) 1307–1341, arXiv:1305.7513 [hep-ex]. (Cited on page 14, 17, 18)

[38] A. Garcia, R. Gauld, A. Heijboer, and J. Rojo, "Complete predictions for high-energy neutrino propagation in matter," *JCAP* **09** (2020) 025, arXiv:2004.04756 [hep-ph]. (Cited on page 15, 18)

[39] M. Thomson, *Modern particle physics*. Cambridge University Press, New York, 2013. (Cited on page 16)

[40] J. G. H. de Groot *et al.*, "Inclusive Interactions of High-Energy Neutrinos and anti-neutrinos in Iron," *Z. Phys. C* **1** (1979) 143. (Cited on page 17)

[41] S. L. Glashow, "Resonant scattering of antineutrinos," *Phys. Rev.* **118** (Apr, 1960) 316–317. https://link.aps.org/doi/10.1103/PhysRev.118.316. (Cited on page 17)

[42] **IceCube** Collaboration, M. G. Aartsen *et al.*, "Detection of a particle shower at the Glashow resonance with IceCube," *Nature* **591** no. 7849, (2021) 220–224, arXiv:2110.15051 [hep-ex]. [Erratum: Nature 592, E11 (2021)]. (Cited on page 17)

[43] compiled by the Class for Physics of the Royal Swedish Academy of Sciences, "Scientific background on the nobel prize in physics 2015 - neutrino oscillations," *NobelPrize.org* (2015) 3–20. https://www.nobelprize.org/prizes/physics/2015/advanced-information/. (Cited on page 18, 25)

[44] P. F. De Salas, S. Gariazzo, O. Mena, C. A. Ternes, and M. Tórtola, "Neutrino Mass Ordering from Oscillations and Beyond: 2018 Status and Future Prospects," *Front. Astron. Space Sci.* **5** (2018) 36, arXiv:1806.11051 [hep-ph]. (Cited on page 22)

[45] **Planck** Collaboration, N. Aghanim *et al.*, "Planck 2018 results. VI. Cosmological parameters," *Astron. Astrophys.* **641** (2020) A6, arXiv:1807.06209 [astro-ph.CO]. [Erratum: Astron.Astrophys. 652, C4 (2021)]. (Cited on page 22, 27)

[46] J. Lesgourgues and S. Pastor, "Neutrino cosmology and Planck," *New J. Phys.* **16**

(2014) 065002, arXiv:1404.1740 [hep-ph]. (Cited on page 22)

[47] **Hyper-Kamiokande** Collaboration, K. Abe *et al.*, "Hyper-Kamiokande Design Report," arXiv:1805.04163 [physics.ins-det]. (Cited on page 22, 28)

[48] T. Ohlsson and H. Snellman, "Three flavor neutrino oscillations in matter," *J. Math. Phys.* **41** (2000) 2768–2788, arXiv:hep-ph/9910546. [Erratum: J.Math.Phys. 42, 2345 (2001)]. (Cited on page 25)

[49] T. Nosek, "Effects of Matter in Neutrino Oscillations and Determination of Neutrino Mass Hierarchy at Long-baseline Experiments," arXiv:1612.09132 [hep-ph]. (Cited on page 25)

[50] L. Wolfenstein, "Neutrino oscillations in matter," *Phys. Rev. D* **17** (May, 1978) 2369–2374. https://link.aps.org/doi/10.1103/PhysRevD.17.2369. (Cited on page 25)

[51] S. Choubey and P. Roy, "Probing the deviation from maximal mixing of atmospheric neutrinos," *Phys. Rev. D* **73** (2006) 013006, arXiv:hep-ph/0509197. (Cited on page 25)

[52] J. Hofestadt, *Measuring the neutrino mass hierarchy with the future KM3NeT/ORCA detector*. PhD thesis, Erlangen - Nuremberg U., 2017. (Cited on page 25)

[53] K. N. Vishnudath, S. Choubey, and S. Goswami, "New sensitivity goal for neutrinoless double beta decay experiments," *Phys. Rev. D* **99** no. 9, (2019) 095038, arXiv:1901.04313 [hep-ph]. (Cited on page 27)

[54] **KATRIN** Collaboration, M. Aker *et al.*, "Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN," *Phys. Rev. Lett.* **123** no. 22, (2019) 221802, arXiv:1909.06048 [hep-ex]. (Cited on page 27)

[55] **KATRIN** Collaboration, M. Aker *et al.*, "Direct neutrino-mass measurement with sub-electronvolt sensitivity," *Nature Phys.* **18** no. 2, (2022) 160–166, arXiv:2105.08533 [hep-ex]. (Cited on page 27)

[56] **Super-Kamiokande** Collaboration, Y. Fukuda *et al.*, "The Super-Kamiokande detector," *Nucl. Instrum. Meth. A* **501** (2003) 418–462. (Cited on page 28)

[57] **SNO** Collaboration, A. Bellerive, J. R. Klein, A. B. McDonald, A. J. Noble, and A. W. P. Poon, "The Sudbury Neutrino Observatory," *Nucl. Phys. B* **908** (2016) 30–51, arXiv:1602.02469 [nucl-ex]. (Cited on page 28)

[58] **Borexino** Collaboration, G. Alimonti *et al.*, "The Borexino detector at the

Laboratori Nazionali del Gran Sasso," *Nucl. Instrum. Meth. A* **600** (2009) 568–593, arXiv:0806.2400 [physics.ins-det]. (Cited on page 28)

[59] **Daya Bay** Collaboration, F. P. An *et al.*, "The Detector System of The Daya Bay Reactor Neutrino Experiment," *Nucl. Instrum. Meth. A* **811** (2016) 133–161, arXiv:1508.03943 [physics.ins-det]. (Cited on page 28)

[60] **Double Chooz** Collaboration, F. Ardellier *et al.*, "Double Chooz: A Search for the neutrino mixing angle theta(13)," arXiv:hep-ex/0606025. (Cited on page 28)

[61] **IceCube** Collaboration, R. Abbasi *et al.*, "The Design and Performance of IceCube DeepCore," *Astropart. Phys.* **35** (2012) 615–624, arXiv:1109.6096 [astro-ph.IM]. (Cited on page 28)

[62] **T2K** Collaboration, K. Abe *et al.*, "The T2K Experiment," *Nucl. Instrum. Meth. A* **659** (2011) 106–135, arXiv:1106.1238 [physics.ins-det]. (Cited on page 28)

[63] **NOvA** Collaboration, M. A. Acero *et al.*, "An Improved Measurement of Neutrino Oscillation Parameters by the NOvA Experiment," arXiv:2108.08219 [hep-ex]. (Cited on page 28)

[64] **MINOS** Collaboration, P. Adamson *et al.*, "A Study of Muon Neutrino Disappearance Using the Fermilab Main Injector Neutrino Beam," *Phys. Rev. D* **77** (2008) 072002, arXiv:0711.0769 [hep-ex]. (Cited on page 28)

[65] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou, "The fate of hints: updated global analysis of three-flavor neutrino oscillations," *JHEP* **09** (2020) 178, arXiv:2007.14792 [hep-ph]. (Cited on page 28, 29)

[66] M. A. Vermeulen, *A blessing in disguise : Characterisations of ProtoDUNE photon showers for neutrino measurements in DUNE*. PhD thesis, Nikhef, Amsterdam, 2021. (Cited on page 28)

[67] M. F. L'Annunziata, "9 - cherenkov counting," in *Handbook of Radioactivity Analysis (Second Edition)*, M. F. L'Annunziata, ed., ch. 9, pp. 719–797. Academic Press, second edition ed., 2003. https://www.sciencedirect.com/science/article/pii/B9780124366039500144. (Cited on page 32)

[68] P. A. Cerenkov, "Visible radiation produced by electrons moving in a medium with velocities exceeding that of light," *Phys. Rev.* **52** (1937) 378–379. (Cited on page 32)

[69] K. W. Melis, *Studying the universe from -3000m N.A.P.* PhD thesis, Universiteit van Amsterdam, 2021. (Cited on page 34, 41, 45, 52, 80, 125, 126)

[70] **KM3NeT** Collaboration, S. Aiello *et al.*, "The KM3NeT multi-PMT optical module," `arXiv:2203.10048 [astro-ph.IM]`. (Cited on page 35, 36, 47, 48, 51)

[71] **KM3NeT** Collaboration, S. Aiello *et al.*, "KM3NeT front-end and readout electronics system: hardware, firmware and software," *J. Astron. Telesc. Instrum. Syst.* **5** no. 4, (2019) 046001, `arXiv:1907.06453 [astro-ph.IM]`. (Cited on page 35)

[72] **KM3NeT** Collaboration, S. Aiello *et al.*, "Embedded Software of the KM3NeT Central Logic Board," `arXiv:2308.01032 [astro-ph.IM]`. (Cited on page 36)

[73] S. I. Klimushin, E. V. Bugaev, and I. A. Sokalski, "Precise parametrizations of muon energy losses in water," in *27th International Cosmic Ray Conference*. 6, 2001. `arXiv:hep-ph/0106010`. (Cited on page 41)

[74] J. Matthews, "A Heitler model of extensive air showers," *Astropart. Phys.* **22** (2005) 387–397. (Cited on page 43)

[75] **ANTARES** Collaboration, A. Albert *et al.*, "Long-term monitoring of the ANTARES optical module efficiencies using ^{40}K decays in sea water," *Eur. Phys. J. C* **78** no. 8, (2018) 669, `arXiv:1805.08675 [astro-ph.IM]`. (Cited on page 45)

[76] **KM3NeT** Collaboration, S. Aiello *et al.*, "Characterisation of the Hamamatsu photomultipliers for the KM3NeT Neutrino Telescope," *JINST* **13** no. 05, (2018) P05035. (Cited on page 45)

[77] **KM3NeT** Collaboration, M. Ageron *et al.*, "Dependence of atmospheric muon flux on seawater depth measured with the first KM3NeT detection units: The KM3NeT Collaboration," *Eur. Phys. J. C* **80** no. 2, (2020) 99, `arXiv:1906.02704 [physics.ins-det]`. (Cited on page 46, 47)

[78] **KM3NeT** Collaboration, P. Kalaczyński and R. Coniglione, "Comparison of the measured atmospheric muon flux with Monte Carlo simulations for the first KM3NeT detection units," *PoS* **ICRC2019** (2020) 927, `arXiv:2101.00907 [hep-ex]`. (Cited on page 47)

[79] R. Bruijn and T. Chiarusi, "KM3NeT Data Acquisition and and Trigger System," *PoS* **ICRC2019** (2019) 848. (Cited on page 48)

[80] C. Reiher, "The clique density theorem," *Annals of Mathematics* **184** no. 3, (Nov, 2016) 683–707. (Cited on page 50)

[81] **KM3NeT** Collaboration, M. Ageron *et al.*, "PMT gain calibration and monitoring based on highly compressed hit information in KM3NeT," *JINST* **16** no. 09, (2021) C09011, arXiv:2108.11467 [astro-ph.IM]. (Cited on page 51)

[82] **KM3NeT** Collaboration, R. Coniglione, A. Creusot, I. Di Palma, D. Guderian, J. Hofestädt, G. Riccobene, and A. Sánchez Losa, "KM3NeT Time Calibration," *PoS* **ICRC2019** (2021) 868. (Cited on page 52)

[83] M. Bouwhuis, "Time synchronization and time calibration in KM3NeT," *PoS* **ICRC2015** (2016) 1170. (Cited on page 52)

[84] **KM3NeT** Collaboration, S. Aiello *et al.*, "Nanobeacon: A time calibration device for the KM3NeT neutrino telescope," arXiv:2111.00223 [astro-ph.IM]. (Cited on page 52)

[85] L. Nauta, *Quantum Effects on a Planetary Scale - The First Neutrino Oscillation Measurement with KM3NeT/ORCA*. PhD thesis, Nikhef, Amsterdam, 2022. (Cited on page 52, 57, 59, 138)

[86] **IceCube** Collaboration, M. G. Aartsen *et al.*, "Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert," *Science* **361** no. 6398, (2018) 147–151, arXiv:1807.08794 [astro-ph.HE]. (Cited on page 54)

[87] **IceCube** Collaboration, M. G. Aartsen *et al.*, "Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector," *Science* **342** (2013) 1242856, arXiv:1311.5238 [astro-ph.HE]. (Cited on page 54)

[88] **KM3NeT** Collaboration, S. Aiello *et al.*, "Determining the Neutrino Mass Ordering and Oscillation Parameters with KM3NeT/ORCA," arXiv:2103.09885 [hep-ex]. (Cited on page 56, 119, 122)

[89] **ANTARES** Collaboration, M. Ageron *et al.*, "ANTARES: the first undersea neutrino telescope," *Nucl. Instrum. Meth. A* **656** (2011) 11–38, arXiv:1104.1607 [astro-ph.IM]. (Cited on page 57)

[90] R. Muller, A. Heijboer, and T. van Eeden, "Search for cosmic neutrino point sources and extended sources with 6-21 lines of KM3NeT/ARCA," *PoS* **ICRC2023** (2023) 1018. (Cited on page 58)

[91] V. Carretero, "Measuring atmospheric neutrino oscillation with KM3NeT/ORCA6," *PoS* **ICRC2023** (2023) 996. (Cited on page 57, 59, 138, 159, 160)

[92] W. J. Metzger, *Statistical methods in data analysis*. Nijmegen Univ. Fys. Lab.,

Nijmegen, 2002. https://cds.cern.ch/record/482882. (Cited on page 62, 79, 80)

[93] M. de Jong and E. van Campenhout, "The probability density function of the arrival time of Čerenkov light," *arXiv e-prints* (May, 2023) arXiv:2305.19626, arXiv:2305.19626 [astro-ph.IM]. (Cited on page 63, 64, 65, 67, 68)

[94] M. de Jong, "Multi-dimensional interpolations in C++," *arXiv e-prints* (July, 2019) arXiv:1907.02597, arXiv:1907.02597 [cs.MS]. (Cited on page 66)

[95] **KM3NeT** Collaboration, S. Aiello *et al.*, "gSeaGen: The KM3NeT GENIE-based code for neutrino telescopes," *Comput. Phys. Commun.* **256** (2020) 107477, arXiv:2003.14040 [astro-ph.IM]. (Cited on page 70)

[96] C. Andreopoulos *et al.*, "The GENIE Neutrino Monte Carlo Generator," *Nucl. Instrum. Meth. A* **614** (2010) 87–104, arXiv:0905.2517 [hep-ph]. (Cited on page 70)

[97] **KM3NeT** Collaboration, A. G. Soto, C. Distefano, and P. Kalaczyński, "Upgrading gSeaGen: from MeV to PeV neutrinos," *JINST* **16** no. 09, (2021) C09008, arXiv:2107.13880 [hep-ex]. (Cited on page 70)

[98] **KM3NeT** Collaboration, J. Schumann and B. Jung, "GiBUU based neutrino interaction simulations in KM3NeT," *JINST* **16** no. 09, (2021) C09022, arXiv:2107.13947 [hep-ex]. (Cited on page 71)

[99] O. Buss, T. Gaitanos, K. Gallmeister, H. van Hees, M. Kaskulov, O. Lalakulich, A. B. Larionov, T. Leitner, J. Weil, and U. Mosel, "Transport-theoretical Description of Nuclear Reactions," *Phys. Rept.* **512** (2012) 1–124, arXiv:1106.1344 [hep-ph]. (Cited on page 71)

[100] D. Bailey, *Monte Carlo tools and analysis methods for understanding the ANTARES experiment and predicting its sensitivity to dark matter*. PhD thesis, Oxford U., 2002. (Cited on page 71)

[101] P. Antonioli, C. Ghetti, E. V. Korolkova, V. A. Kudryavtsev, and G. Sartorelli, "A Three-dimensional code for muon propagation through the rock: Music," *Astropart. Phys.* **7** (1997) 357–368, arXiv:hep-ph/9705408. (Cited on page 71)

[102] J. H. Koehne, K. Frantzen, M. Schmitz, T. Fuchs, W. Rhode, D. Chirkin, and J. Becker Tjus, "PROPOSAL: A tool for propagation of charged leptons," *Comput. Phys. Commun.* **184** (2013) 2070–2090. (Cited on page 71)

[103] M. Chrzaszcz, T. Przedzinski, Z. Was, and J. Zaremba, "TAUOLA of τ lepton

decays—framework for hadronic currents, matrix elements and anomalous decays," *Comput. Phys. Commun.* **232** (2018) 220–236, arXiv:1609.04617 [hep-ph]. (Cited on page 71)

[104] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, and T. Thouw, "CORSIKA: A Monte Carlo code to simulate extensive air showers,". (Cited on page 71)

[105] G. Carminati, A. Margiotta, and M. Spurio, "Atmospheric MUons from PArametric formulas: A Fast GEnerator for neutrino telescopes (MUPAGE)," *Comput. Phys. Commun.* **179** (2008) 915–923, arXiv:0802.0562 [physics.ins-det]. (Cited on page 72, 76, 194)

[106] Y. Becherini, A. Margiotta, M. Sioli, and M. Spurio, "A Parameterisation of single and multiple muons in the deep water or ice," *Astropart. Phys.* **25** (2006) 1–13, arXiv:hep-ph/0507228. (Cited on page 72, 73, 74, 75, 76)

[107] C. Forti, H. Bilokon, B. d'Ettore Piazzoli, T. K. Gaisser, L. Satta, and T. Stanev, "Simulation of atmospheric cascades and deep underground muons," *Phys. Rev. D* **42** (1990) 3668–3689. (Cited on page 72)

[108] **MACRO** Collaboration, M. Ambrosio *et al.*, "Measurement of the residual energy of muons in the Gran Sasso underground laboratories," *Astropart. Phys.* **19** (2003) 313–328, arXiv:hep-ex/0207043. (Cited on page 72)

[109] A. G. Tsirigotis, A. Leisos, and S. E. Tzamarias, "HOU Reconstruction & Simulation (HOURS): A complete simulation and reconstruction package for very large volume underwater neutrino telescopes," *Nucl. Instrum. Meth. A* **626-627** (2011) S185–S187. (Cited on page 77)

[110] **GEANT4** Collaboration, S. Agostinelli *et al.*, "GEANT4–a simulation toolkit," *Nucl. Instrum. Meth. A* **506** (2003) 250–303. (Cited on page 77)

[111] **KM3NeT** Collaboration, K. Melis, A. Heijboer, and M. de Jong, "KM3NeT/ARCA Event Reconstruction Algorithms," *PoS* **ICRC2017** (2018) 950. (Cited on page 85, 88)

[112] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of Applied Mathematics* **2** no. 2, (1944) 164–168. http://www.jstor.org/stable/43633451. (Cited on page 86)

[113] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics* **11** no. 2,

(1963) 431–441. http://www.jstor.org/stable/2098941. (Cited on page 86)

[114] A. Domi, *Shower Reconstruction and Sterile Neutrino Analysis with KM3NeT/ORCA and ANTARES*. PhD thesis, Aix-Marseille U., Genoa U., 2019. (Cited on page 88)

[115] **KM3NeT** Collaboration, T. van Eeden, J. Seneca, and A. Heijboer, "High-energy reconstruction for single and double cascades using the KM3NeT detector," *PoS* **ICRC2021** (2021) 1089. (Cited on page 88)

[116] T. P. Li and Y. Q. Ma, "Analysis methods for results in gamma-ray astronomy," *Astrophys. J.* **272** (1983) 317–324. (Cited on page 97)

[117] B. Ó Fearraigh, "Tuning parametric models of the atmospheric muon flux in MUPAGE to data from the KM3NeT detector," *PoS* **ICRC2021** (2021) 1176. (Cited on page 98)

[118] **AMANDA** Collaboration, J. Ahrens *et al.*, "Muon track reconstruction and data selection techniques in AMANDA," *Nucl. Instrum. Meth. A* **524** (2004) 169–194, arXiv:astro-ph/0407044. (Cited on page 121)

[119] **IceCube** Collaboration, M. G. Aartsen *et al.*, "Energy Reconstruction Methods in the IceCube Neutrino Telescope," *JINST* **9** (2014) P03009, arXiv:1311.4767 [physics.ins-det]. (Cited on page 121)

[120] M. Dentler, "Investigation of the One-Particle Approximation in the ANTARES simulation package KM3," bachelor thesis, U. Erlangen-Nuremberg (main), 2012. (Cited on page 124)

[121] A. Rohatgi, "Webplotdigitizer: Version 4.6," 2022. https://automeris.io/WebPlotDigitizer. (Cited on page 125)

[122] C. Kopper, *Performance Studies for the KM3NeT Neutrino Telescope*. PhD thesis, Erlangen - Nuremberg U., 2010. (Cited on page 126)

[123] S. Adrián-Martínez *et al.*, "Intrinsic limits on resolutions in muon- and electron-neutrino charged-current events in the KM3NeT/ORCA detector," *JHEP* **05** (2017) 008, arXiv:1612.05621 [physics.ins-det]. (Cited on page 127, 128, 141)

[124] F. James and M. Roos, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations," *Comput. Phys. Commun.* **10** (1975) 343–367. (Cited on page 136)

[125] Y. Coadou, "Boosted decision trees," arXiv:2206.09645 [physics.data-an].

(Cited on page 168)

Summary

'Since the dawn of time, mankind has looked to the stars and wondered–'

We do not need to resort to such highfalutin speech to introduce someone to science. 'Science', by the way, should not be such a scary word as it is to some people. In this summary chapter, I will try to present the work I carried out during my PhD in as least a scary way as possible.

The title of this thesis is **'Following the Light; Novel event reconstruction techniques for neutrino oscillation analyses in KM3NeT/ORCA'**. This PhD thesis - put simply - is all about light, and in particular the kind of light caused by neutrinos.

Everything in the world that we can interact with, the tables we sit at, the air we breath, the planet we live on, is made up of atoms. Let us start not being scared by the word 'atom'. An atom is the basic building block of matter. The cells in your body are building blocks of you and everyone you know. What makes up every cell, the parts of the cells, are atoms. Atoms are they themselves made up of smaller particles, called protons, neutrons and electrons. The different combinations of the number of protons, neutrons and electrons that are put together make up different atoms. In this way we get atoms of oxygen, gold, silver and so on: elements we are familiar with. Now, electrons are fundamental particles; they cannot be broken down into anything smaller. Electrons have some properties that define them, such as their mass and electric charge. The flow (movement) of electrons is a flow of electric charge that gives us the electricity we are familiar with in everyday life.

Neutrinos are also fundamental particles, much, much lighter than electrons. They do not 'stick' to protons the way electrons do to form atoms. They are neutral particles, meaning they have no electric charge associated with them. Neutrinos are created in interactions between particles, particle decays, and in the decay of atoms. Protons, coming from outer space, hit the Earth's atmosphere, which create neutrinos, raining down from the sky and passing through you, me, and all other atoms on Earth. Neutrinos are also created in radioactive decays, where an atom changes into another kind of atom;

this is caused by neutrons turning into protons and electrons, and emitting a neutrino. In fact, the neutrino was first thought of because the laws of physics of these decays made little sense without the existence of neutrinos. Because of their creation in decays and nuclear fusion (where protons can combine to form new atoms), every star in the sky produces neutrinos. So, neutrinos are created in the atmosphere and also in outer space. Our Sun sends 100 billion neutrinos through an area the size of your thumbnail every second.

Okay, what I just suggested is very scary. We are apparently subject to an onslaught, a barrage of neutrinos every second. Should we be living in fear because of this? The answer is no, because of neutrinos being extremely light and having no electric charge. This means they pass through atoms usually without doing anything, and just continue travelling through the Universe until the very, very rare chance that they do interact with a particle they encounter. This however, makes them nearly impossible to detect and study. Why would we want to study them, you must be wondering? Well, we want to learn everything about the Universe we live in, of course. Neutrinos come in 3 types, which we call 'flavours'. Neutrinos have been proven to 'oscillate', or change from one flavour to the other at any given moment in time and space. That means when we detect a neutrino (finally) interacting with an atom, there is a certain probability associated with what kind of neutrino we will see at that moment in time. This understandably crazy concept is not something we see in day-to-day life, tables do not spontaneously turn into chairs, or water to beer, alas. We naturally want to study and understand this concept of 'neutrino oscillations'. Furthermore, the fact that neutrinos oscillate mean they have mass. This was for a very long time not believed to be the case, we used to think of neutrinos as massless, the same way light has no mass. The three flavours of neutrinos have three corresponding masses, but to make things more complicated it is not true that neutrino of flavour A has a mass of X. Instead, because the quantum mechanical (i.e. the particle) world does not behave the same way as the world we can see and touch, at any given moment a neutrino exists as a combination of all three masses and all three flavours at once. When neutrinos interact with atoms we can figure out what flavour they were (because of the kind of particles they create - more on this later), and because we understand the mathematics of their oscillations, we can build a picture of what combinations of masses they had at that moment of interaction. At the moment, we know that one of the mass 'states' is heavier than another. The third, we have no idea if it is the heaviest or the lightest of the three. Finding this out, known as the 'neutrino mass ordering' problem, is an urgent and fundamentally important goal in the field of particle physics research. It is a major goal of the KM3NeT experiment, along with

understanding and precisely measuring neutrino oscillations. Knowing the neutrino mass ordering has important implications for how our Universe works and was formed.

How can we study these complex, invisible particles, which I just said do not interact very often? Well, they do interact, you just need a very large volume in which neutrinos will interact, and study that volume. Since the probability of neutrinos interacting with an atom is so very small, increasing the number of atoms you have, i.e. the volume, increases the probability of them interacting and you detecting them! Now what kind of volume should we study? Well, we can never 'see' a neutrino, but when they have enough energy they will interact with an atom and create other particles that we can detect. We cannot see these new particles with the naked eye, but when the neutrino has enough energy it results in the creation of particles which emit light as they travel. Light is something we can definitely see and detect.

About 70% of the Earth's surface is covered in water, seas, oceans, and polar ice caps. We understand well how light travels through water. This makes water a very abundant and natural volume in which we could study neutrino interactions. The Cubic Kilometre Neutrino Telescope, or 'KM3NeT' for short, is an experiment being built right now to do exactly that. In the Mediterranean Sea, two locations off the coast of France and Italy are being dedicated to enclose a huge volume of water with digital sensors, which can detect and record light as a signal. This involves filling glass spheres with light sensors, and arranging the spheres along vertical lines connected to the sea floor. KM3NeT records the data (the light signals) and sends them to computers on shore to be analysed. An illustration of the KM3NeT experiment is shown in Figure S.1. The two different locations are used to measure different kinds of neutrinos, one site detects low energy neutrinos created in the atmosphere, like I described above, and the other site detects high energy neutrinos created in outer space: in stars, in the centres of galaxies, in turbulent and unimaginably high-energy processes we are still trying to understand. The detector dedicated to measuring neutrino oscillations and determining the neutrino mass ordering is called KM3NeT/ORCA. Upon its completion, the ORCA detector will consist of 115 of these vertical lines. The work carried out in this thesis is dedicated to both the envisaged 115-line detector, and to the 6-line ORCA detector, which operated from January 2020 – November 2021 before the addition of more lines.

So, KM3NeT records light from neutrino interactions, we send them to computers, and then analyse the data to figure out these scientific mysteries - easy. Alas, not so easy. Since neutrinos interact infrequently, we need to collect many years of data to record enough neutrinos that can give us a precise measurement of their oscillations, and also enough neutrinos to allow us to distinguish between the case of the third mass state

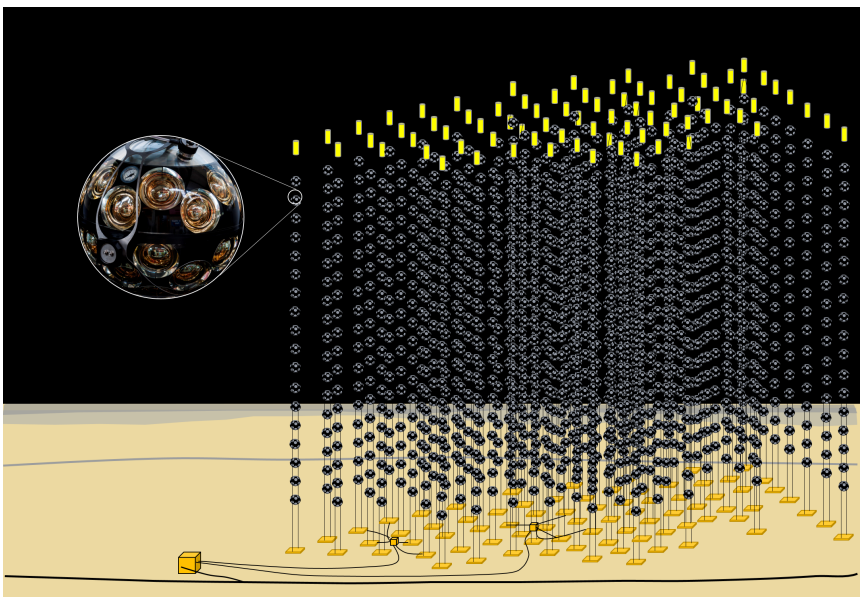


Figure S.1.: The glass sphere, filled with golden digital light sensors, is shown. These glass spheres are in turn arranged along vertical lines and connected to a network on the Mediterranean Sea floor. This infrastructure forms the KM3NeT experiment.

being the lightest, or it being the heaviest — the difference between the two cases changes the probability of neutrino oscillations and what we measure, but the difference is subtle. Furthermore, even in the dark depths of the Mediterranean Sea, other light sources exist besides the particles created in neutrino interactions. Microscopic sea creatures emit light, and radioactive decays in the sea water also can be detected as light signals in KM3NeT. Neutrinos are not the only particle that make it to the depths either. Muons (particles a bit heavier than electrons) are also created in the atmosphere, and can also make it below sea level. These muons also emit light, and to make matters worse muons are also a particle created in neutrino interactions! These atmospheric muons form a main 'background' signal to the products of neutrino interactions we want to detect. Luckily, atmospheric muons mainly come from straight above the detector, while neutrinos can pass through the Earth, and create a muon in the KM3NeT detectors that goes upwards. This difference in direction gives us a handle on how to exclude atmospheric muons from analyses.

The study of atmospheric muons and the particles that created them is in itself a whole research area. The scientific community would also like to understand these interactions as well as possible. However, these interactions in the atmosphere are very complex, they involve the creation and decay of many particles, and are not so easily simulated. It takes quite intensive computer simulations to model these processes. One thing I contributed to KM3NeT as part of my research is improving one of the atmospheric muon

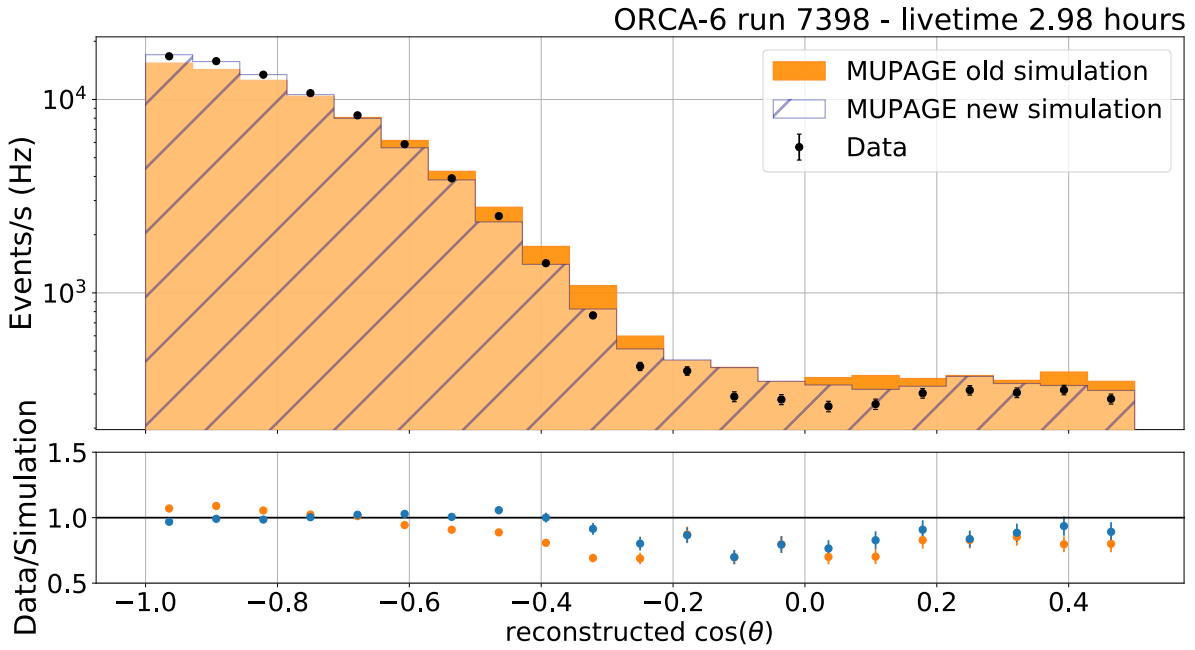


Figure S.2.: A distribution of the cosine of the direction angle θ of the atmospheric muons, as represented by data points (black), the old MUPAGE simulation (orange), and the new MUPAGE simulation (blue). The top figure shows the number of events per second for a particular data run, the bottom figure shows the data divided by the two simulations, where data/simulation = 1 is the most desirable outcome.

simulations so it describes the data better. This simulation software, called 'MUPAGE', is less complex and intensive than others, as it describes using some formulae, the energy and direction of atmospheric muons measured at different depths at or below sea level. MUPAGE however, does not perfectly perfectly describe the atmospheric muon data we collect in KM3NeT, meaning we do not simulate or understand our background well enough. (Take note that although for neutrino analyses these muons are background, for the study of atmospheric muon development and the particles that created them, these muons are in fact signal! That is a story for another day). I performed a 'tuning' of the formulae within MUPAGE to better describe the data we take in KM3NeT/ORCA. A result of this tuning is shown in Figure S.2. When we look at the angle of the arriving atmospheric muons, you can see the data points, the old simulation, and the new one I have tuned. The new simulation much better describes the data. This new simulation has in fact been used to simulate the background atmospheric muons in the latest published research using KM3NeT data.

Back to neutrinos. When they interact in the detectors, their flavour (and how they interact) determines the kind of particles they create. We call some of these particle signatures, like those created by muons, 'tracks'. They travel through the detector in a straight line, emitting light we can detect. Another signature is called a 'shower',

caused by electrons or heavier particles which result in a burst of light, usually more localised than a track signature. Conventionally, because they are easy to separate, we 'reconstruct' neutrino interactions as tracks or shower, not both. To reconstruct in this context means determining their energy and direction. Treating these signatures separately is a bit naive, though. In fact, almost every neutrino interaction will result in a shower. For neutrino interactions which result in a track signature, there will also be a shower, and reconstructing the energy and direction of **both** of these would result in a much better determination of the direction and energy of the neutrino that created them. What's more, such a track *and* shower reconstruction directly gives you an estimate of how much energy has gone into the muon and into the shower (something we call 'Bjorken-y'). Knowing this quantity means we could describe neutrino interactions far more accurately. I spent the vast majority of my PhD research developing such a track and shower reconstruction, as is described within this book.

Figure S.3 shows the neutrino energy resolution (how well we can determine the neutrino energy) as a function of the true (simulated) neutrino energy in the fully-complete KM3NeT/ORCA detector. Using the new, reconstructed neutrino energy from the track and shower reconstruction results in a much better estimate than using the track energy by itself — the lower the resolution in value, the better we can determine the energy.

Some neutrinos with a very close value of true and reconstructed Bjorken-y (or y for short) in the KM3NeT/ORCA detector are shown in Figure S.4. There are many neutrinos which lie close to the diagonal line, indicating the neutrinos for which the true and reconstructed Bjorken-y are the same.

This new track+shower reconstruction is capable of reconstructing the neutrino energy with an improved resolution, and the Bjorken-y parameter for the first time. This track+shower reconstruction can also be used to differentiate between neutrino events with tracks, and those without, as discussed within this book. The improved energy resolution and the Bjorken-y parameter are essential ingredients to analyses with the KM3NeT/ORCA data, and will lead to more precise measurements of the oscillation parameters and a more precise determination of the yet-unknown neutrino mass ordering.

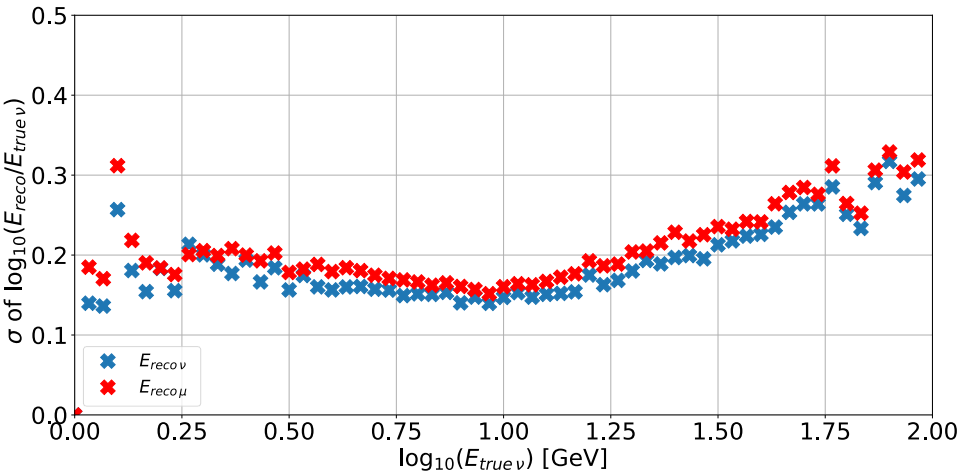


Figure S.3.: The energy resolution as a function of the true neutrino energy for the KM3NeT/ORCA detector, using the track+shower reconstruction technique. The energy resolution of the reconstructed neutrino energy (blue points) is better (lower in value) than the reconstructed track energy (red points), meaning the reconstructed neutrino energy gives a better estimate of the true neutrino energy. This is valid across the entire energy range relevant for KM3NeT/ORCA.

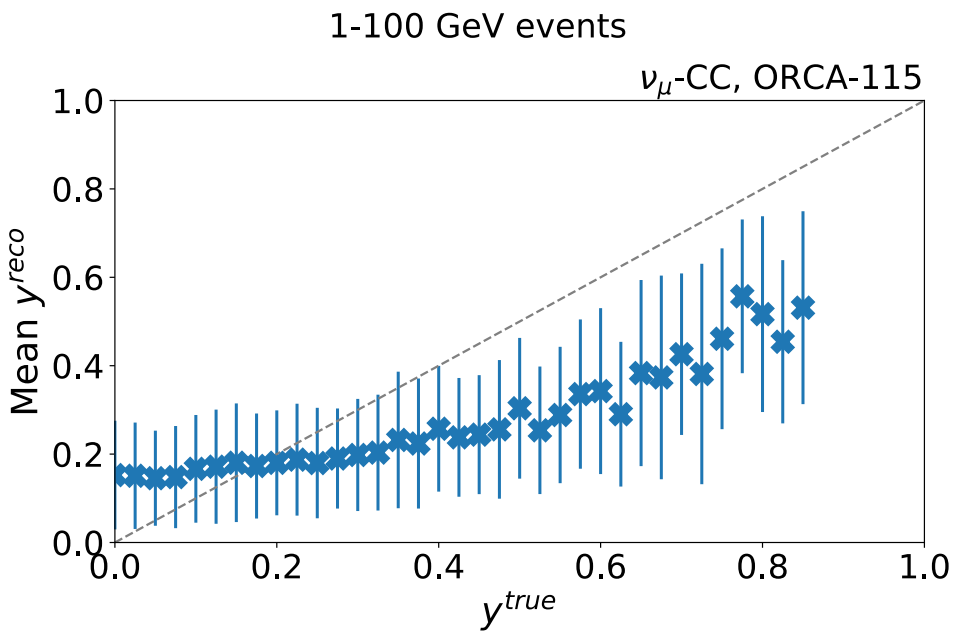


Figure S.4.: The energy transferred from the neutrino to the shower it creates (called 'Bjorken-y') is determined by the track+shower reconstruction. For some neutrinos simulated in the KM3NeT/ORCA detector, the true and reconstructed values of Bjorken-y are very close, indicated by the diagonal line.

Samenvatting - Nederlands

'Sinds het begin der tijden keek de mensheid naar de sterren en verwonderde zich–'

We hoeven niet zo hoogdravend te doen om iemand kennis te laten maken met wetenschap. 'Wetenschap' zou trouwens niet zo'n eng woord moeten zijn als het voor sommige mensen is. In deze samenvatting zal ik proberen om het werk dat ik tijdens mijn promotie traject heb gedaan zo min mogelijk eng uit te leggen.

De titel van dit proefschrift is **'Following the Light; Novel event reconstruction techniques for neutrino oscillation analyses in KM3NeT/ORCA'**. Dit proefschrift gaat - eenvoudig gezegd - over licht, en in het bijzonder over het soort licht dat door neutrino's wordt veroorzaakt.

Alles waarmee we in contact kunnen komen in de wereld, de tafels waar aan we zitten, de lucht die we in ademen, de planet waarop we leven, is opgebouwd uit atomen. Een atoom is is de bouwsteen van de materie. De cellen in je lichaam zijn bouwstenen van jou en iedereen die je kent. Waar elke cel uit bestaat de onderdelen van de cellen, zijn atomen. Atomen zijn op hun beurt opgebouwd uit kleinere deeltjes, protonen, neutronen en elektronen genaamd. De verschillende combinaties van het aantal protonen, neutronen, en elektronen die samengevoegd worden, vormen verschillende atomen. Op deze manier krijgen we atomen van zuurstof, goud, zilver enzovoort: elementen die we kennen. Nu zijn elektronen fundamentele deeltjes; ze kunnen niet worden opgesplitst in iets kleiners. Elektronen hebben een aantal eigenschappen die hen definiëren, zoals hun massa en elektrische lading. De verplaatsing van elektronen is het bewegen van elektrische lading, die en geeft elektriciteit zoals we die in het dagelijks leven kennen.

Neutrino's zijn ook fundamentele deeltjes. En ze zijn veel en veel kleiner dan elektronen. Ze 'kleven' niet aan protonen zoals elektronen dat doen om atomen te vormen. Het zijn neutrale deeltjes, wat betekent dat ze geen elektrische lading hebben. Neutrino's ontstaan door interacties tussen deeltjes en het verval van atomen. Protonen, afkomstig uit de ruimte, botsen op de atmosfeer van de aarde, waardoor neutrino's ontstaan die uit de lucht naar beneden regenen en door jou, mij en alle andere atomen op aarde

heen gaan. Neutrino's ontstaan ook in radioactief verval, waarbij een atoom verandert in een ander soort atoom; dit wordt veroorzaakt doordat neutronen in protonen en elektronen veranderen en een neutrino uitzenden. In feite werd het neutrino voor het eerst bedacht omdat de natuurkundige wetten van deze vervalprocessen niet klopten zonder het bestaan van neutrino's. Door de creatie van neutrino's in verval en kernfusie (waarbij protonen kunnen combineren om nieuwe atomen te vormen) produceert boven dien elke ster aan de hemel neutrino's. Neutrino's ontstaan dus in de atmosfeer én in de ruimte. Onze zon stuurt elke seconde 100 miljard neutrino's door een gebied zo groot als je duimnagel.

Oké, wat ik net suggereerde klinkt heel eng. We zijn blijkbaar onderworpen aan een aanval, een vloedgolf van neutrino's elke seconde. Moeten we daarom in angst leven? Het antwoord is nee, omdat neutrino's extreem licht zijn en geen elektrische lading hebben. Dit betekent dat ze meestal zonder iets te doen atomen passeren en gewoon door het heelal blijven reizen tot de extreem zeldzame kans dat ze een interactie aangaan met een ander deeltje dat ze tegenkomen. Dit maakt het echter bijna onmogelijk om ze te detecteren en te bestuderen. Waarom zouden we ze willen bestuderen, zul je afvragen? Nou, we willen natuurlijk alles leren over het heelal. Neutrino's zijn er in 3 soorten, die we 'smaken' noemen. Het is bewezen dat neutrino's 'oscilleren', of op elk moment in tijd ruimte van de ene smaak naar de andere veranderen. Dat betekent dat wanneer we een neutrino detecteren, dat (eindelijk) een interactie heeft met een atoom, er een bepaalde kans is verbonden aan welke neutrino smaak we op dat moment zullen zien. Dit concept is niet iets wat we in het dagelijks leven zien. Tafels veranderen niet spontaan in stoelen, of, jammer genoeg, water niet in bier. Natuurlijk willen we het concept van 'neutrino oscillaties' bestuderen en begrijpen. Daarnaast betekent het feit dat neutrino's oscilleren dat ze massa hebben. Dit werd lange tijd niet voor waar aangenomen. We dachten dat neutrino's massaloos waren, net als licht dat geen massa heeft. De drie smaken neutrino's hebben drie corresponderende massa's, maar om het nog ingewikkelder te maken is het niet zo dat neutrino van smaak A een massa heeft van X. Omdat de kwantummechanische wereld (d.w.z. de deeltjeswereld) zich niet op dezelfde manier gedraagt als de wereld die we kunnen zien en aanraken, bestaat een neutrino op elk moment uit een combinatie van alle drie de massa's en alle drie de smaken tegelijkertijd. Wanneer neutrino's een wisselwerking hebben met atomen, kunnen we achterhalen welke smaak ze hadden vanwege het soort deeltjes dat ze creëren - hierover later meer. Omdat we de wiskunde van de oscillaties begrijpen, kunnen we ons een beeld vormen van de combinaties van massa's die ze op het moment van interactie hadden. Op dit moment weten we dat één van de drie zwaarder is dan de andere. Van de derde weten we niet of die de zwaarste of de lichtste van de drie is. Dit staat bekend als het 'neutrino

mass ordering' probleem, en is van fundamenteel belang binnen de deeltjesfysica. Het probleem is een belangrijk doel van het KM3NeT experiment, samen met het begrijpen en nauwkeurig meten van neutrino oscillaties. Kennis over de ordening van neutrino massa's heeft belangrijke implicaties voor ons begrip van het heelal.

Hoe kunnen we deze complexe, onzichtbare deeltjes bestuderen, waarvan ik net zei dat ze niet vaak interageren? Nou, ze interageren wel, je hebt alleen een heel groot volume nodig om interacties te bestuderen. Omdat de kans op interactie tussen neutrino's en een atoom zo klein is, moet je het aantal atomen dramatisch vergroten. Als het volume vergroot, vergroot de kans dat ze op elkaar reageren en dat jij dit detecteert! Wat voor volume moeten we nu bestuderen? Nou, we kunnen een neutrino nooit 'zien', maar als ze genoeg energie hebben, zullen ze een wisselwerking aangaan met een atoom en andere deeltjes creëren die we kunnen detecteren. We kunnen deze nieuwe deeltjes niet met het blote oog zien, maar als het neutrino genoeg energie heeft, ontstaan er deeltjes die lucht uitzenden terwijl ze reizen. En licht kunnen we wel degelijk zien en detecteren.

Ongeveer 70% van het aardoppervlak is bedekt met water, zeeën, oceanen en poolkappen. We begrijpen goed hoe licht door water reist. Dit maakt water tot een zeer overvloedig en natuurlijk volume waarin we neutrino interacties zouden kunnen bestuderen. De 'Cubic Kilometre Neutrino Telescope', kortweg 'KM3NeT', is een experiment dat op dit moment wordt gebouwd om precies dat te doen. Op twee locaties voor de kust van Frankrijk en Italieë in de Middellandse Zee wordt een enorme hoeveelheid water omsloten met digitale sensoren die licht als signaal kunnen detecteren en registreren. Hiervoor worden glazen bollen gevuld met lichtsensoren en langs verticale lijnen op de zeebodem geplaatst. De KM3NeT detector registreert de gegevens (de lichtsignalen) en stuurt deze naar computers aan wal voor analyse. Een illustratie van het KM3NeT experiment wordt getoond in Figuur N.1. De twee verschillende locaties worden gebruikt om verschillende soorten neutrino's te meten, één locatie detecteert laagenergetische neutrino's die in de atmosfeer ontstaan, zoals ik hierboven. De andere locatie detecteert hoogenergetische neutrino's in de ruimte ontstaan: in sterren, in de centra van sterrenstelsels, in turbulente en onvoorstelbaar hoogenergetische processen die we proberen te begrijpen. De detector die is gewijd aan het meten van neutrino oscillaties en het bepalen van de ordening van neutrino massa's heet KM3NeT/ORCA. Na voltooiing zal de ORCA-detector uit 115 van deze verticale lijnen bestaan. Het werk in dit proefschrift is gewijd aan de ORCA detetctor met 6 lijnen, die van januari 2020 to november 2021 in werking was voordat er meer lijnen werden toegevoegd.

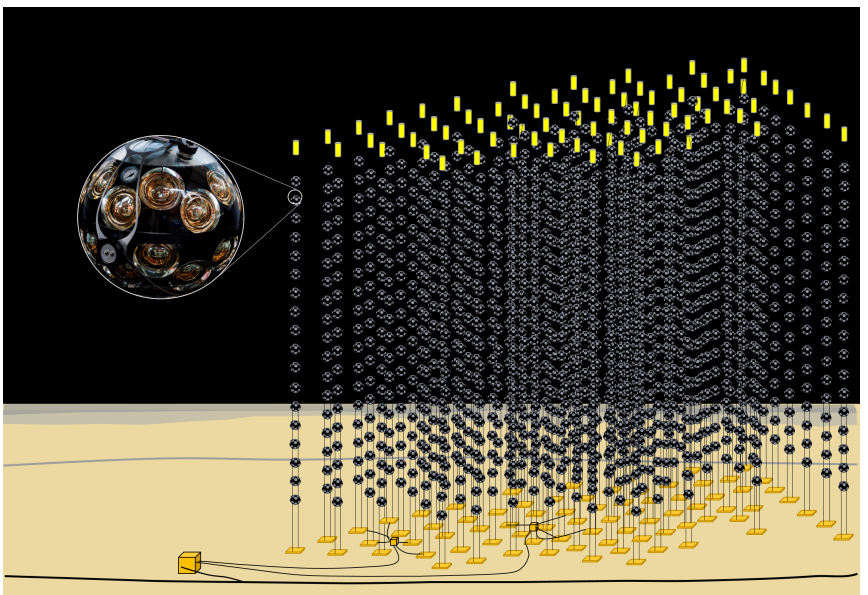


Figure N.1.: De glazen bol, gevuld met gouden digitale lichtsensoren, wordt getoond. Deze glazen bollen zijn op hun beurt langs verticale lijn en bevestigd verbonden met een netwerk op de bodem van de Middellandse Zee.

Dus KM3NeT registreert het licht van neutrino interacties, dat sturen we naar computers en analyseren de gegevens om deze wetenschappelijke mysteries op te lossen: eenvoudig. Jammer genoeg, niet zo eenvoudig. Omdat neutrino's niet vaak interageren, moeten we vele jaren aan gegevens verzamelen om genoeg neutrino's te registreren. Dit zal een precieze meting van hun oscillaties kunnen geven, en ook genoeg neutrino's om ons in staat te stellen onderscheid te maken tussen het geval dat de derde massatoestand de lichtste is en het geval dat deze de zwaarste is. Het verschil tussen de twee gevallen verandert de effecten van neutrino oscillaties, maar het verschil is subtiel. Daarnaast bestaan er zelfs in de donkere diepten van de Middellandse Zee andere lichtbronnen dan de deeltjes die ontstaan bij neutrino interacties. Microscopische zeediertjes zenden licht uit en radioactief verval van zeezout kan ook worden gedicteerd als lichtsignalen in KM3NeT. Neutrino's zijn ook niet de enige deeltjes die de diepte bereiken. Muonen (deeltjes die iets zwaarder zijn dan elektronen) worden ook in de atmosfeer gemaakt en kunnen ook onder zeeniveau komen. Deze muonen zenden ook licht uit en tot overmaat van ramp zijn muonen ook een deeltje dat ontstaat bij neutrino interacties! Deze atmosferische muonen vormen een belangrijk 'achtergrond'-signaal voor de producten van neutrino interacties die we willen detecteren. Gelukkig komen atmosferische voornamelijk recht boven de detector vandaan, terwijl neutrino's door de aarde heen kunnen gaan en in KM3NeT detectoren een muon creëren dat omhoog gaat. Dit verschil in richting geeft ons houvast om atmosferische muonen uit te sluiten van analyses.

Het bestuderen van atmosferische muonen en de deeltjes die deze muonen creëren is op zich al een heel onderzoeksgebied. De wetenschappelijke gemeenschap zou ook graag deze interacties zo goed mogelijk begrijpen. Deze interacties in de atmosfeer zijn echter zeer complex: ze omvatten de creatie en het verval van vele deeltjes en zijn niet zo gemakkelijk te simuleren. Er zijn behoorlijk intensieve computersimulaties nodig om deze processen te modelleren. Als onderdeel van mijn onderzoek heb ik onder andere bijgedragen aan KM3NeT experiment door een van de atmosferische muonsimulaties te verbeteren, zodat deze de metingen beter beschrijven. Deze simulatiesoftware, 'MUPAGE' genaamd, is minder complex en intensief dan andere, omdat het met behulp van een aantal formules de energie en richting beschrijft van atmosferische muonen die op verschillende diepten op of onder zeeniveau worden gemeten. MUPAGE beschrijft echter niet perfect de atmosferische muon data die we verzamelen in KM3NeT. Dit betekent dat we onze achtergrond niet goed genoeg simuleren of begrijpen. (Merk op dat hoewel deze muonen voor neutrino analyses achtergrond zijn, de muonen voor de studie van atmosferische muon ontwikkeling en de deeltjes die deze muonen creëren in feite signaal zijn! Dat is een verhaal voor een ander keer). Ik heb een 'afstemming' van de formules binnen MUPAGE uitgevoerd om de data die we in KM3NeT/ORCA gebruiken beter te beschrijven. Een resultaat van deze afstemming is te zien in Figuur N.2. Als we kijken naar de hoek van de aankomende atmosferische muonen, zie je de data punten, de oude simulatie en de nieuwe simulatie die ik heb afgestemd. De nieuwe simulatie beschrijft de gegevens veel beter. Deze nieuwe simulatie is in feite gebruikt om de atmosferische muonen te simuleren in het laatst gepubliceerde onderzoek met KM3NeT gegevens.

Terug naar neutrino's. Wanneer ze reageren (in de detectoren), bepaalt hun smaak (en het type reactie) het soort deeltjes dat ze creëren. Sommige van deze deeltjes, zoals muonen, noemen we 'sporen'. Ze reizen in een rechte lijn door de detector en zenden licht uit dat we kunnen detecteren. Een andere 'signatuur' heet een deeltje 'lawine', veroorzaakt door elektronen of zwaardere deeltjes die resulteren in een uitbarsting van licht, meestal meer gelokaliseerd dan een spoor signatuur. Het is conventie om neutrino interacties als sporen of lawines te reconstrueren, niet allebei, omdat ze gemakkelijk te scheiden zijn. In deze context betekent reconstrueren het bepalen van hun energie en richting. Deze eigenschappen apart behandelen is echter een beetje naïef. In feite zal bijna elke neutrino interactie resulteren in een deeltje lawine. Voor neutrino interacties die resulteren in een spoor signatuur zal er ook een lawine zijn, en het reconstrueren van de energie en richting van beide zou resulteren in een veel beter bepaling van de richting en energie van het neutrino dat de reactie. Bovendien kun je met zo'n spoor- *en* lawine reconstructie direct schatten hoeveel energie in het muon en in de lawine is gegaan (dit

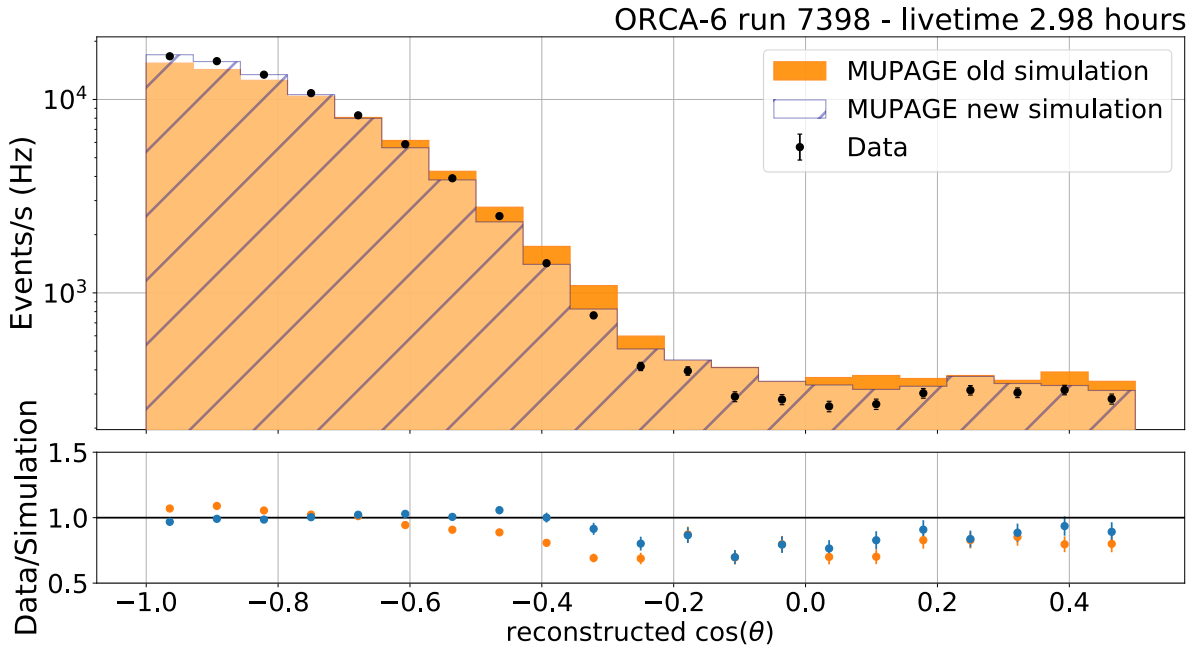


Figure N.2.: Een verdeling van de cosinus van de richtingshoek θ van de atmosferische muonen, zoals weergegeven door data punten (zwart), de oude MUPAGE simulatie (oranje) en de nieuwe MUPAGE simulatie (blauw). De bovenste figuur toont het aantal events per seconde voor en bepaalde datarun, de onderste figuur toont de data gedeeld door de twee simulaties, waarbij data/simulation = 1 de meest wenselijke uitkomst is.

noemen we 'Bjorken-y'). Als we deze Bjorken-y kennen, kunnen we neutrino interacties veel nauwkeuriger beschrijven. Het overgrote deel van mijn promotieonderzoek heb ik besteed aan het ontwikkelen van zo'n spoor- en lawine reconstructie, zoals beschreven in dit boek.

Figure N.3 toont de neutrino energieresolutie (hoe goed we de neutrino energie kunnen bepalen) als functie van de ware (gesimuleerde) neutrino energie in de volledige KM3NeT/ORCA detector. Het gebruik van de nieuwe, gereconstrueerde neutrino energie uit de spoor- en lawine reconstructie resulteert in een veel betere schatting dan het gebruik van de spoor energie op zichzelf - hoe lager de resolutie waarde, hoe beter we de energie kunnen bepalen.

Enkele neutrino's met een zeer dicht bij elkaar liggende waarde van ware en gereconstrueerde Bjorken-y in de KM3NeT/ORCA-detector getoond in Figure N.4 worden. Er zijn veel neutrino's die langs of dicht bij de diagonale lijn liggen, wat aangeeft dat voor deze neutrino's de ware en gereconstrueerde Bjorken-y hetzelfde zijn.

Deze spoor- en lawine reconstructie is in staat om de neutrino energie met een verbeterde resolutie te bepalen, en voor het eerst de Bjorken-y-parameter te reconstrueren.

Deze spoor- en lawine reconstructie kan ook worden gebruikt om onderscheid te maken tussen neutrino's met en zonder sporen, zoals in dit boek beschreven. De beter energieresolutie en de Bjorken-y-parameter zijn essentiële ingrediënten voor analyses met de KM3NeT/ORCA gegevens, en zullen leiden tot nauwkeurigere metingen van de oscillatie parameters en een nauwkeurigere bepaling van de nog onbekende neutrino mass ordening.

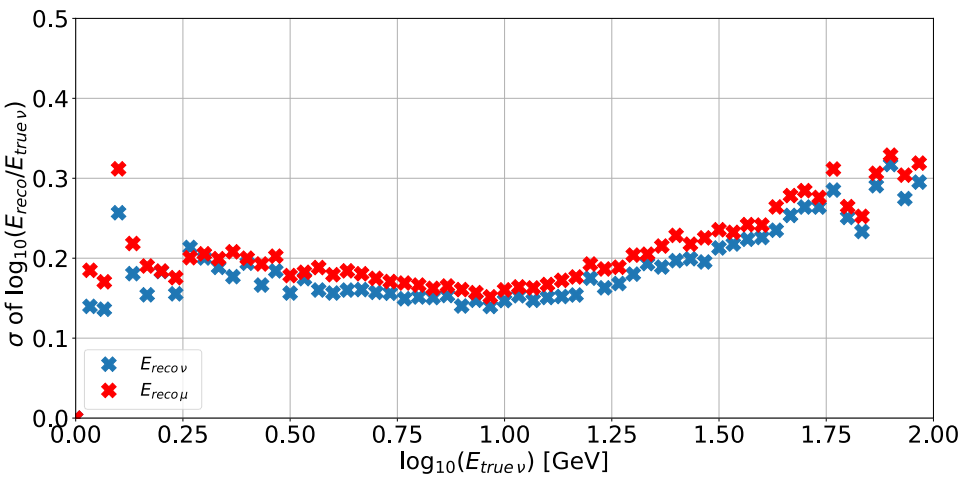


Figure N.3.: De energieresolutie als functie van de ware neutrino energie voor de KM3NeT/ORCA detector, met behulp van de spoor- en lawine reconstructietechniek. De energieresolutie van de gereconstrueerde neutrino energie (blauwe punten) is beter (lager in waarde) dan de gereconstrueerde spoor energie (rode punten), wat betekent dat de gereconstrueerde neutrino energie een betere schatting geeft van de ware neutrino energie. Dit geldt voor het hele energiebereik dat relevant is voor KM3NeT/ORCA.

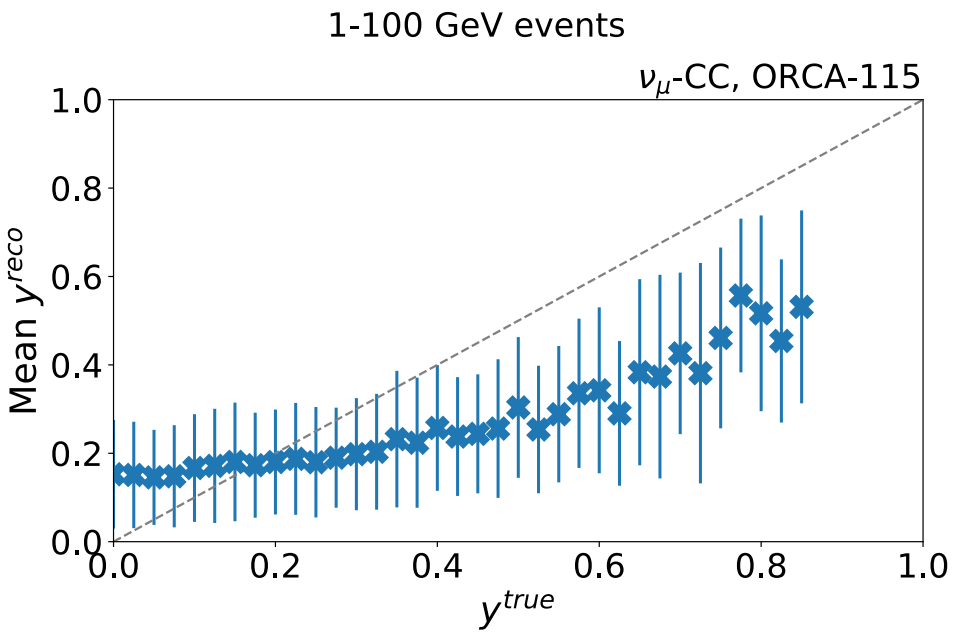


Figure N.4.: De energie die wordt overgedragen van het neutrino naar de deeltje lawine die hij creëert (genaamd 'Bjorken-y') wordt bepaald door de spoor- en lawine reconstructie. Voor sommige neutrino's die in de KM3NeT/ORCA detector zijn gesimuleerd, liggen de werkelijke en gereconstrueerde waarden van Bjorken-y zeer dicht bij elkaar, aangegeven door de diagonale lijn.

Achoimre - Gaeilge

‘Ó thús ama, tá iontas déanta ag an gcine daonna de na réaltaí sa spéir–”

Níl gá ar bith le hóráid bhreá chun eolaíocht a mhíniú do dhuine. Ní gá go mbeadh daoine neirbhíseach nó faiteach faoin focal ‘eolaíocht’. Sa chaibidil seo, déanfaidh mé achoimre ar an obair a rinne mé i rith mo thráchtais chomh héasca agus soiléir is féidir liom.

An teideal atá ar an tráchtas seo ná **‘Following the Light; Novel event reconstruction techniques for neutrino oscillation analyses in KM3NeT/ORCA’**. Baineann an tráchtas seo le solas, go háirithe le solas a bhaineann le neoidríonónná.

Tá gach uile rud ar domhan déanta de adaimh - an t-aer, an chathaoir ar a shuímid, ár bplainéad. Adamh is ea bloc tógála ábhar. Blocanna tógála díot féin is ea na cealla i do chorp. Tá gach cuid d’ár gcealla déanta de adaimh. Tá na hadaimh déanta de phíosaí, nó cáithníní, níos lú, ar a dtugtar prótóin, neodróin agus leictreoin. Nuair a chuirtear teaglaim éagsúla uimhreacha de na prótóin, neodróin agus leictreoin le chéile, is é sin is cúis le na hadaimh éagsúla. Is mar seo a chruthaítear adaimh de ocsaigín, ór, airgead agus rl: dúile a aithnímid. Is cáithníní bunúsacha iad leictreoin nach féidir a dhéanamh níos lú. Tá cáilíochtaí áirithe ag leictreoin a dhéanann cur síos orthu, mar shampla a mais agus a lucht leictreachais. Gluaiseacht leictreoin is ea sruth lucht leictreachais a thugann ár gcuid leictreachais dúinn.

Cáithníní bunúsacha freisin is ea neoidríonónná, atá i bhfad níos éadroime ná leictreoin. Ní ‘ghreamaíonn’ said le prótóin mar a dhéanann leictreoin chun adaimh a dhéanamh. Níl aon lucht leictreachais ag baint le neoidríonónná. Cruthaítear neoidríonónná san idirghníomhaíocht idir cháithníní agus mheath adaimh. Buaileann prótóin, a thagann ó spás, atmasféar an Domhain agus cruthaítear neoidríonónná a thagann anuas ón spéir agus a ghabhann trí dhaoine agus trí gach adamh eile ar Domhan. Cruthaítear neoidríonónná chomh maith i meath radaighníomhach, nuair a athraíonn adamh go cineál eile adamh; tarlaíonn sé seo nuair a athraíonn neodróin go prótóin agus leictreoin agus astaítear, ní chuirtear amach, neoidríonó. I ndáiríre, tháining eolaithe ar

thuairim an neoidríonó ar dtús mar ní raibh ciall ná brí le rialacha fisice na meath seo gan neoidríonónná a bheith i láthair. De bharr a gcruthú i meath agus i gcomhleá núi-cléach (nuair a thagann prótóin le chéile chun adaimh nua a dhéanamh), gineann gach réalt sa spéir neoidríonónná. Mar sin, cruthaítear neoidríonónná san atmasféar agus i spás. Seolann ár nGrian 100 billiún neoidríonónná trí achar méid do ionga ordóige gach soicind!

Is cosúil go mbímid faoi ionsaí ag na neoidríonónná seo gach soicind, ach ní gá go gcuirfeadh sé seo faitíos orainn, mar tá neoidríonónná an-éadrom agus níl lucht leictreachais iontu. Gabhann siad trí adaimh gan tionchar ar bith, agus leanann siad orthu tríd An Chruinne, agus is fíor-annamh a idirghníomhaíonn siad le cáithnín lena mbuaileann siad. Tá sé beagnach dodhéanta, mar sin, iad a bhraith agus staidéar a dhéanamh orthu. Cén fáth staidéar a dhéanamh orthu? Bhuel, tugann siad níos mó eolais dúinn faoin Chruinne ina mhairimid. Tá 3 shaghas neoidríonó ann, ar a nglaotar 'blasanna'. Tá sé cruthaithe go ndéanann said 'ascalú', nó athrú, ó bhlas amháin go blas eile uair ar bith in am agus spás. Mar sin, nuair a bhraithimid neoidríonó (ar deireadh) ag idirghníomhú le hadamh, tá dóchúlacht i gceist leis an saghas neoidríonó a léirítear ag an am áirithe sin. Ní fheictear an coincheap aisteach seo inár saol laethúil, ní athraíonn bord go cathaoir, nó uisce go beoir, faraor. Ní mór staidéar a dhéanamh ar an 'ascalú neoidríonó' seo agus é a thuiscint. Ciallaíonn an t-ascalú a dhéanann siad go bhfuil mais acu. Ceapadh ar feadh i bhfad nach raibh mais acu, faoi mar nach bhfuil mais ag solas. Tá an trí bhlas neoidríonónná comhfhreagrach le trí mhais, ach ní fíor go bhfuil mais X ag neoidríonó le blas A. In ionad sin, toisc nach bhfuil an iompar céanna ag an domhan candam meicniúil (domhan cáithnín) is atá sa domhan againne, maireann neoidríonó mar chomhcheangal den trí bhlas agus an trí mhais le chéile. Nuair a bhíonn idirghníomhú idir neoidríonónná agus adaimh, is féidir an blas a bhí acu a dhéanamh amach mar gheall ar na cáithníní a chruthaíonn siad, agus toisc go dtuigimid mata a n-ascalaithe, is féidir linn comhcheangal na mais a dhéanamh amach ag am an idirghníomhú. Ag an am seo, tuigimid go bhfuil ceann amháin de 'staid' na mais níos troime ná an ceann eile. Níl tuairim dá laghad againn an é an tríú mais an ceann is éadroime nó is troime díobh. Tugtar 'the neutrino mass ordering' nó 'ord mais neoidríonó' ar an gceist seo agus tá sé práinneach go réiteofar é seo i dtaighde fisic cháithníní. Is aidhm an-mhór é sin den turgnamh KM3NeT, chomh maith le hascalú neoidríonónná a thuiscint agus a thomhas go cruinn. Tá impleachtaí suntasacha ag réiteach na ceiste faoi 'ord mais neoidríonó' ar conas mar a cruthaíodh agus a n-oibríonn ár gCruinne.

Conas a dhéantar staidéar ar na cáithníní casta, dofheicthe seo, nach ndéanann mórán idirghníomhú? Ar ndóigh, má mhéadaítear an méid adaimh atá againn - an toirt - ansin

méadaítear an seans go ndéanfaidh siad idirghníomhú, agus is féidir iad a aimsiú! Anois, cén toirt a bheadh i gceist? Tá neoidríonónná dofheicthe dúinn, ach má bhíonn go leor fuinnimh acu déanfaidh siad idirghníomhú le hadamh agus cruthófar cáithníní eile gur féidir linn a aimsiú. Beidh na cáithníní nua seo dofheicthe don tsúil dhaonna, ach má bhíonn go leor fuinnimh ag an neoidríonó, cruthófar cáithníní a astaíonn solas nuair a ghluaiseann siad. Gan amhras, tá solas sofheicthe dúinn.

Tá breis is 70% de dhromchla an domhain clúdaithe le huisce, farraigí, aigéin agus oighearchaidhpeanna polacha. Tuigimid conas a ghluaiseann solas trí uisce, mar sin is toirt nádúrtha agus flúirseach é uisce inar féidir staidéar a dhéanamh ar idirghníomhú neoidríonónná. Is é sin díreach an turgnamh atá á thógáil i láthair na huaire chun an staidéar sin a dhéanamh, sin an 'Cubic Kilometre Neutrino Telescope', nó 'KM3NeT'. Tá dhá shuíomh amach ó chósta na Fraince agus na hIodáile, sa Mheánmhuir, roghnaithe ina bhfuil braiteoirí digiteacha curtha i dtoirt ollmhór uisce, agus braitheann na braiteoirí sin solas mar chomhartha, agus déanann siad taifeadadh air. Líontar sféir ghloine le braiteoirí solais, agus leagtar iad i línte ceartingearacha atá ceangailte le grinneall na farraige. Tógann KM3NeT na sonraí taifeadta (na comharthaí solais) agus seoltar go ríomhairí ar ais ar tír. Tá léaráid den turgnamh KM3NeT i bhFigiúr G.1. Úsáidtear an dá shuíomh chun cineálacha éagsúla neoidríonónná a thomhas, braitheann suíomh amháin neoidríonónná le fuinneamh íseal a cruthaíodh san atmaisféar, mar atá luaite cheana, agus braitheann an suíomh eile cinn le fuinneamh níos airde a cruthaíodh sa spás: sna réalta, i lár réaltra, i bpróisis shuaite ina bhfuil fuinneamh ard doshamhlaithe nach dtuigeann eolaithe fós. Tugtar KM3NeT/ORCA ar an mbrathadóir atá dírithe ar ascalú neoidríonónná a thomhas agus an t-ord mais neoidríonó a dhéanamh amach. Beidh 115 de na línte ceartingearacha seo sa bhrathadóir ORCA nuair a bheidh sé réidh go hiomlán. Tá an obair sa tráchtas seo dírithe ar am mbrathadóir ORCA 6-líne, a d'oibrigh ó Eanáir 2020 go Samhain 2021 sular cuireadh línte breise leis.

Mar sin, déanann KM3NeT taifeadadh ar sholas ó idirghníomhú neoidríonónná, cuirimid na sonraí go ríomhairí gus déanaimid anailís ar na mistéir eolaíochta seo: éasca. Faraor, níl sé chomh héasca sin. Is annamh a dhéanann neoidríonónná idirghníomhú, mar sin ní mór blianta de shonraí a bhailiú chun go leor neoidríonónná a thaifeadadh a thabharfaidh tomhas cruinn dúinn ar a n-ascalaithe. Freisin, beidh go leor uainn chun a fháil amach an é an tríú staid mais an ceann is éadroime nó is troime — tá an difríocht caolchúiseach ach athraíonn sé dóchúlacht ascalú neoidríonó agus conas iad a thomhas. Lena chois sin, fiú amháin i ndoimhneacht dhorcha na Meánmhara, tá foinsí solais eile ann seachas na cáithníní ó idirgníomhú neoidríonó. Astaíonn ainmhithe mara micreascópacha solas, agus bíonn comharthaí solais ó mheath radaighníomhach sa tsáile

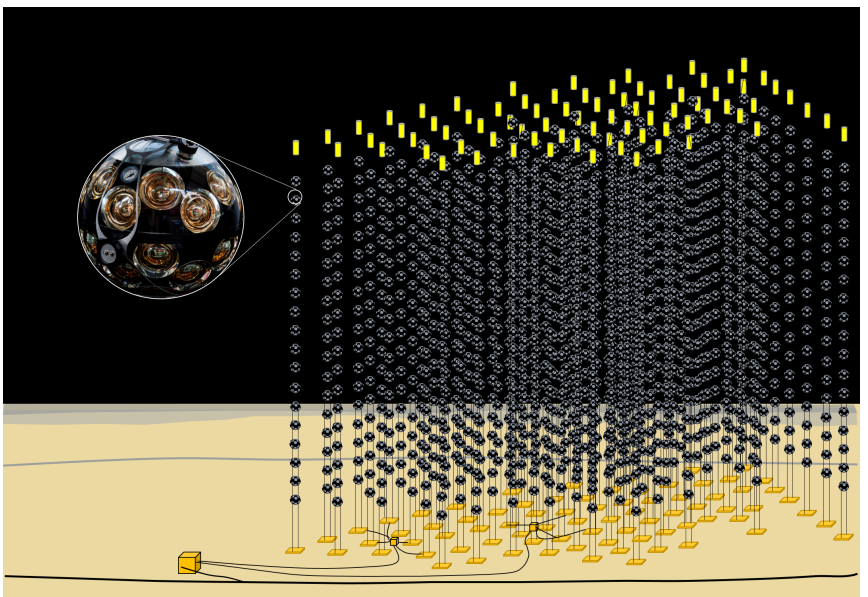


Figure G.1.: An sféar gloine, lán le braiteoirí solais digiteacha óir. Leagtar na sféir ghloine seo i línte ceartingearacha atá ceangailte le gréasán ar ghrinneall na Meánmhara.

le braith i KM3NeT. Sroicheann cáithníní eile, seachas neoidríonónná, doimhneachtaí na farraige freisin, mar shampla múóin (cáithníní atá níos troime ná leictreoin) atá cruthaithe san atmaisféar. Astaíonn siad sin solas, agus níos measa fós, is cáithnín é an múón a chruthaítear in idirghníomhaíochtaí neoidríonónná! Is 'cúlra' iad na múóin ón atmaisféar do na cáithníní atá cruthaithe in idirghníomhaíochtaí neoidríonónná gur mian linn a bhraith. Tá an t-ádh linn go dtagann na múóin atmaisféaracha seo ó díreach taobh thuas den bhrathadóir, cé gur féidir le neoidríonónná a mbealach a dhéanamh tríd an Domhan, agus múón a chruthú sna brathadóirí KM3NeT a théann in airde. Tugann an difríocht seo sa treo ina ghabhann na múóin treoir dúinn ar conas na múóin ón atmaisféar a fhágáil ar lár ón anailís.

Is taighde ann féin é an staidéar seo ar mhúóin ón atmaisféar agus ar na cáithníní a chruthaigh iad. Tá suim mhór ag an bpobal eolaíochta sna hidirghníomhaíochtaí seo agus tá siad ag iarraidh iad a thuiscint chomh maith agus is féidir. Tá na hidirghníomhaíochtaí seo san atmaisféar an-chasta , áfach, mar tá cruthú agus meath a lán cáithníní i gceist, agus is deacair ionsamhlú a dhéanamh orthu. Tá gá le hionsamhlú ríomhaireachta an-dian chun na próisis seo a mhúnlú. Mar chuid de mo thaighde féin, ionchur amháin a rinne mé do KM3NeT ná chuir mé feabhas ar ionsamhlú múón atmaisféarach amháin chun bheith in ann cur síos níos cruinne a dhéanamh ar na sonraí. Tugtar 'MUPAGE' ar an mbogearraí ionsamhlaithe seo, nach bhfuil chomh casta ná chomh dian le cinn eile, mar úsáideann sé roinnt foirmlí chun cur síos ar fuinneamh agus treo na múóin atmaisféaracha a thomhastar ag doimhneachtaí éagsúla faoi leibhéal na farraige. Ní

dhéanann MUPAGE cur síos iomlán cruinn ar na sonraí ó múóin atmaisféaracha a bhailítear i KM3NeT, áfach; ní ionsamhlaímid ná ní thuigimid an cúlra seo go hiomlán. (Nóta: is cúlra iad na múóin seo do anailís neoidríonó, ach maidir le staidéar ar fhorbairt múóin atmaisféaracha agus na cáithníní a chruthaigh iad, is comhartha iad na múóin seo, i ndáiríre. Ach sin scéal eile). Rinne mé 'tiúnáil' ar na foirmlí i MUPAGE chun cur síos níos fearr a dhéanamh ar na sonraí a thógtar i KM3NeT/ORCA. Léiríonn Figiúr G.2 toradh an tiúnála sin. Má dhírimid ar uillinn na múóin atmaisféaracha a thagann isteach, feictear na pointí sonraí, an sean ionsamhlú agus an t-ionsamhlú nua atá tiúnálta agam. Déanann an t-ionsamhlú nua cur síos i bhfad níos fearr ar na sonraí. Ar ndóigh, tá úsáid á bhaint as anois chun na múóin atmaisféaracha a ionshamhlú sa taighde foilsithe is déanaí ag úsáid sonraí KM3NeT.

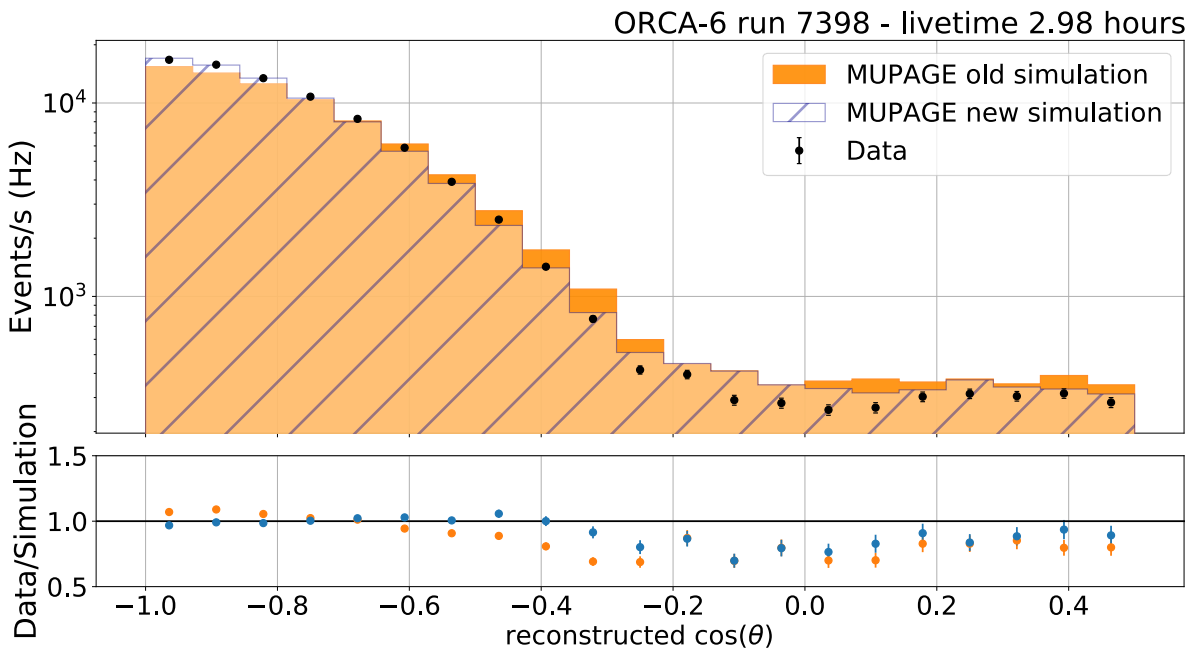


Figure G.2.: Dáileadh an chomhshínis den treo uillinn θ de na múóin atmaisféaracha, léirithe ag pointí sonraí (dubh), an sean MUPAGE ionsamhlú (oráiste), agus an t-ionsamhlú nua MUPAGE (gorm). Léiríonn an figiúr ag barr méid na n-eachtraí gach soicind do chúrsa áirithe sonraí, léiríonn an figiúr ag bun na sonraí roinnte ag an dá ionsamhlú, nuair is data/simulation = 1 an toradh is inmhianaithe.

Ar ais go neoidríonónná anois. Nuair a idirghníomhaíonn siad sna brathadóirí, braitheann an saghas cáithníní a chruthaíonn siad ar a mblas (agus ar an slí a dhéanann siad idirghníomhú). Tugtar 'traiceanna' ar chuid de na saghsanna próifílí cáithníní seo, mar shampla na próifílí atá cruthaithe ag múóin. Gluaiseann na cáithníní seo tríd an bhrathadóir i líne dhíreach, ag astú solais gur féidir linn a bhraith. Saghas próifíl eile is ea 'cith', a thagann ó leictreoin nó ó cháithníní níos troime, a astaíonn pléasc solais a bhíonn níos lárnaí ná an próifíl solais a thagann as traiceanna. De ghnáth, déantar 'atógáil'

ar idirghníomhú neoidríonó mar traic nó mar chith, mar tá sé éasca iad a dheighilt óna chéile. Chun atógáil mar seo a dhéanamh, ní mór an fuinneamh agus an treo san idirghníomhú a chinntiú. Ní dhéanann sé ciall iomlán iad a dheighilt óna chéile, áfach, mar bíonn toradh cith ar beagnach gach idirghníomhú neoidríonó. Má bhíonn toradh traic ar idirghníomhú neoidríonónná, beidh toradh cith ann chomh maith, agus trí treo agus fuinneamh an **dá** cheann a atógáil, bheadh cinntiú i bhfad níos fearr ar treo agus fuinneamh an neoidríonó a chruthaigh iad. Freisin, tugann atógáil traic *agus* cith tomhas díreach ar an méid fuinnimh atá imithe isteach sa mhúón agus sa chith (ar a dtugtar 'Bjorken-y'). Leis an eolas seo, táimid in ann cur síos i bhfad níos cruinne a dhéanamh ar idirghníomhú neoidríonónná. Chaith mé formhór de mo thaighde tráchtais ag forbairt atógáil traic agus cith dá leithéid, mar a phléigh mé sa leabhar seo.

Léiríonn Figiúr G.3 réiteach fuinnimh neoidríonó (chomh maith agus is féidir an fuinneamh neoidríonó a chinntiú) ag feidhmiú mar chuid den fíor fuinneamh ionsamhlaithe sa bhrathadóir KM3NeT/ORCA nuair atá sé críochnaithe. Bíonn tomhas i bhfad níos fearr ag an fuinneamh neoidríonó athógtha nua seo ón atógáil 'traic agus cith' le chéile ná ón fuinneamh traic leis féin — dá laghad é luach an réitigh, is ea is fearr gur féidir an fuinneamh a chinntiú.

Léiríonn Figiúr G.4 neoidríonónná le luach an-chóngarach don Bjorken-y fíor agus athógtha (nó y) sa bhrathadóir KM3NeT/ORCA. Tá a lán neoidríonónná ina luí gar don líne trasnánach, a léiríonn go bhfuil an Bjorken-y fíor agus athóghta mar an gcéanna.

Tá an atógáil traic+cith nua seo in ann an fuinneamh neoidríonó a chinntiú le réiteach níos fearr, agus an Bjorken-y a chinntiú don chéad uair. Is féidir an atógáil traic+cith seo a úsáid freisin chun idirdhealú a dhéanamh idir idirghníomhú neoidríonó le traiceanna, agus iad siúd nach bhfuil traiceanna ann, mar a phléitear sa leabhar seo. Is comhábhair riachtanacha iad an réiteach feabhsaithe fuinnimh agus an Bjorken-y chun anailís a dhéanamh le sonraí KM3NeT/ORCA, agus beidh tomhais níos beachta ar na paraiméadair ascalaithe agus cinneadh níos cruinne ar an ord mais neoidríonó, nach bhfuil a fhios againn go fóill cén ord atá air.

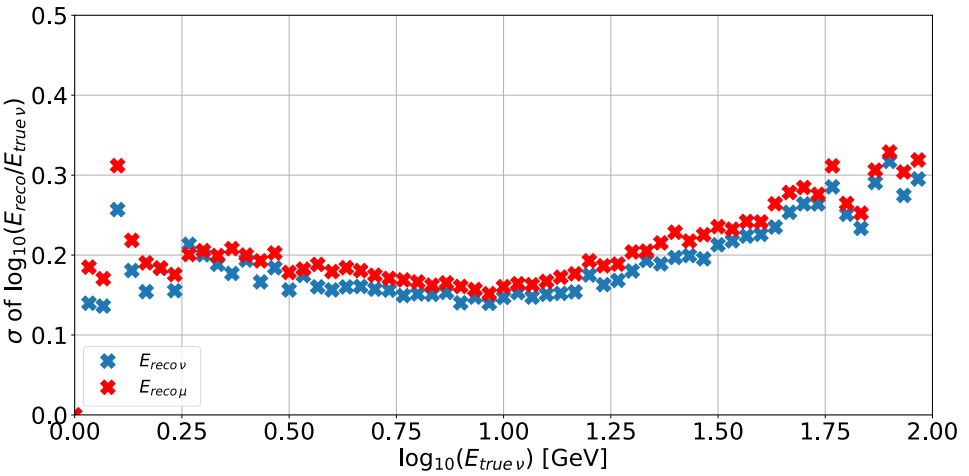


Figure G.3.: An réiteach fuinnimh mar fheidhm den fuinneamh neoidríonó fíor sa bhrathadóir KM3NeT/ORCA, ag úsáid an modh athógála traic+cith. Tá réiteach fuinnimh níos fearr (luach níos ísle) ag an fuinneamh neoidríonó athógtha (pointí gorma) ná ag an fuinneamh traic athógtha (pointí dearga). Mar sin tugann an fuinneamh neoidríonó athógtha tomhas níos fearr den fíor fuinneamh neoidríonó. Tá an feabhas seo le feiceáil don réimse iomlán fuinnimh a bhaineann le KM3NeT/ORCA.

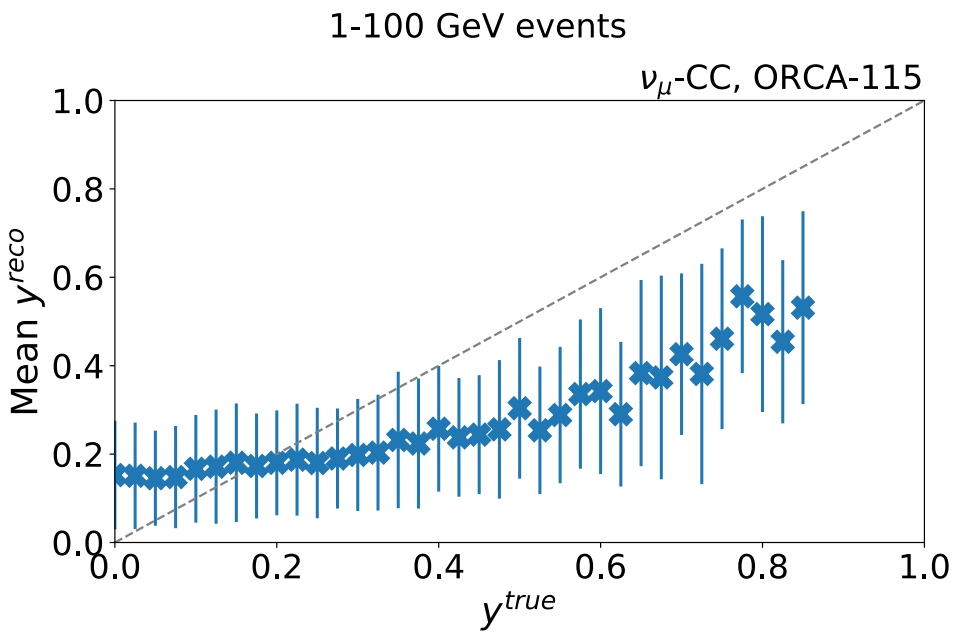


Figure G.4.: Cinntíonn an atógáil traic+cith an fuinneamh a aistrítear ón neoidríonó go dtí an cith a chruthaíonn sé (ar a dtugtar 'Bjorken-y'). Tá an Bjorken-y fíor agus athógtha an-chóngarach - léirithe ag an líne trasnánach - do chuid de na neoidríonónná ionsamhlaithe sa bhrathadóir KM3NeT/ORCA.

Acknowledgements

"Part of the journey is the end." - Tony Stark.

Writing acknowledgements for this thesis is not something that feels real. I'm not sure if it'll feel real before the thesis is sent for publication, or even afterwards. I want to take some time, and many pages, to thank the people who were a part of this PhD experience over the last number of years. A PhD thesis is about scientific research and results, yes, but it is also a life experience. I want to thank the people who helped me get through all of the PhD years and contributed to this book.

First and foremost, I thank my supervisors Ronald Bruijn and Paul de Jong. Thank you for the endless reading of these chapters, for the input and feedback to shape this thesis into a far more elegant and professional version than appeared in the first drafts, and for getting me to the finish line. Ronald, this reconstruction was your vision, and I followed through with that vision as best as I could. Your open-door policy of always being approachable and answering any questions I had when wandering into your office was very much appreciated. Thank you also for the time you devoted to debugging software with me, especially looking at this minimiser for literal hours on end with me. I enjoyed the time together at collaboration meetings, and finding someone who knows how to appreciate Terry's chocolate oranges properly. The way in which you are inspired by the physics we can do with KM3NeT, and talking about this with you, always encouraged me. Your supervision over the years has shaped me into a better researcher. Paul, thank you for leading the Nikhef-KM3NeT group for the majority of my time in Nikhef, for feedback and comments on my research, for reading these chapters, and importantly, for always being available for a meeting on where to go with this research when it was needed. During one C3 meeting you said to me "I believe you can do it!", and the encouragement that gave stayed with me.

Of course, I have to thank the members of the Nikhef-KM3NeT group who spent these years with me. We socialised, discussed, shared lunch and coffee breaks, and did our research together. Although I worked on the topics in this thesis very much by myself, there were always people willing to help and answer questions when asked.

I'll begin with my fellow PhD students. Lodewijk, man, I appreciated so much having you in this group. I realised this more so when you finished. Thanks for all the fun times together, for talking about work (and importantly, not work), discussing code and physics with me, for the never-ending memes, for big glasses of Coke in Mexico, for movie nights, for putting up and joining in with all my pop culture references. Bouke, even though we worked on different topics, I felt we followed a similar research path and had similar experiences during our PhD. I owe you countless thanks for all the software help over the last few years, for always being willing to help with problems, for all the fun times outside of work, and for discussing the Jpp PDF coordinate system with me once every second month when I would get confused again. Jordan, I have known you almost since my first day in the Netherlands. It would have been a far more boring experience without you. Thank you for the countless fun times together, for trips after collaboration meetings, experiences in Ireland and Mexico with one another, discussions about physics and PhD woes, jokes and discussions about everyone and everything, and for listening to my stream of conscience every day. I'm afraid we are now more friends than colleagues, and hope we remain friends for a long time.

Rasa, thanks for the fun times together. I appreciated having you in the group very much, including our discussions of the way we do things in the group, in KM3NeT, and in Nikhef as a whole. Thijs, my long-lost twin! Thanks for all the fun we had, movies we watched, help with physics you gave me, collaboration on reconstruction together. May the Lord of the Rings jokes continue for a long time. Clara, you saw me during my absolutely worst period when sharing an office with me in my last year. You still spoke to me, considering this, so I hope I haven't scared you as a friend. Thanks for all our *lovely* and *fantastic* conversations together about work and life. Karel, thank you for all our nice times together at the beginning of my PhD, for being the go-to person for a naive new PhD student, for all the help and discussions about reconstruction work we had. Milo, you were part of the wider neutrino group, and I am so glad for it. Thanks for all the fun times, texting each other when some word was woefully mispronounced, for appreciating pints of Guinness and texting me on days when one was needed. With you and Casandra I feel like I have friends for life. Long may we continue playing the tin whistle! You helped this Irishman unlock new levels of Irishness. Vincent, good luck with your PhD, I hope seeing me in the office at the end of mine didn't scare you off. Thanks also to Vikas and Marjolein for the nice conversations we had.

I also want to thank some postdocs, past and present. Rodrigo, without you half of this thesis would not be here. Thanks for all the help and support when you were in Nikhef, working on MUPAGE with me as a de-facto supervisor, for all the fun times

together, including drinking beers and talking about music (U2 included). Valentin, together with Lodewijk we shared the best damn office! You were such an important person in the group, not only for all the work you did, help you gave students, but also for someone to take a coffee with and discuss work or life. I still think of the evening I cycled through quite a storm for croziflette in your apartment. You inspire people to talk about physics and I hope to work together in the future. Bruno, man, thanks for all the fun times at the beginning of my PhD! Not only for the help with physics/programming, but going to see Dylan Moran together and proving that one can be an excellent researcher and also not treat things too seriously. Alba, thank you for answering the hundreds of emails and messages about shower reconstruction. Thanks also for the fun times at collaboration meetings, and for 'spacca sta piastra' high-fives! Alfonso, thanks also for the fun times at collaboration and working group meetings, and for the help and explanation of simulation software. Thank you Suzanne, Jhilik and Paco for the nice times as part of the group.

I also thank the Nikhef-KM3NeT staff members for their help with this research, in the form of group meeting discussions and for the social outings together. Thanks go to Maarten, Dorothea, Mieke, Aart, Daan and Ernst-Jan for this. Aart, I appreciated someone to share the Fr. Ted references with in the Greek collaboration meeting. Thank you Dorothea in turn for the social activities organised for the group, and for being so understanding in my difficulties at the end of this PhD. Mieke, during the COVID period I felt really welcomed by the dinners you invited me for, like I had a second home in Amsterdam during those difficult times. I remember the very flattering gift of green peas for St. Patrick's Day I received through the mail. Thanks for helping me feel like I had people to rely on in Amsterdam. Maarten, I cannot thank you enough for the time you spent answering the millions of emails I have sent starting with 'I have a question'. Your input to this research, in the form of conversations about reconstruction, comments during presentations, and coming up with ideas during coffee breaks helped fill this thesis. I want to emphasise how much I appreciated having you in the group, for being a person willing to have coffee breaks and discuss things for hours, for non-work related discussions and stories and for always treating me like a colleague, not a student. I hope you appreciated all of English-language things we had fun discussing over the years.

There are too many people to thank individually in the wider KM3NeT collaboration, but I have many people to mention for the professional help and nice social times we shared during the years building and developing this underwater neutrino telescope. Vladimir, Kay, Tamás, Jannik, Luigi, Rosa Anna, João, Jürgen, Jerzy, Johannes, Massimiliano, Marta, Zineb, Piotr, Andrey, Vasilis, Bruny, Victor, Alfonso (Lazo), Oleg...the

list goes on. I give a special shout-out to Luigi for answering so many questions on data processing and productions, and in turn to Jürgen and João for the enthusiasm and discussions they showed on this track+shower reconstruction. Zineb, thanks for the wonderful trip to Marseille, I hope to visit you again soon. Vladimir, I look forward to working with you in the near-future!

Returning to Nikhef, there is one group I have to thank for being there for me during all the years. I have to try now not to fill 20 pages. I don't know how it happened, but the LHCb group in Nikhef became a second family to me. Thank you to all the wonderful people in this group for accepting this random Irish guy into your midst. Some of you, I love very much, and if you don't know that I can confirm for you. Receiving four texts from different people when Mr. Brightside played at an LHCb BBQ made me feel very much loved.

Maarten, I cannot emphasise enough how much I loved having you around during these years. Thanks for the countless jokes, the even-more-countless gin and tonics, all of the concerts and appreciation of rock 'n roll. We will forever share the nearly-a-native-speaker moment, and driving through the Scottish highlands together. Aleksandra, we are married in the eyes of the Dutch healthcare system. How can I find someone more perfect for me than someone who asks to watch Marvel at 3 am? Thank you for the countless evenings we shared, watching movies, talking about life, battling people off chairs in Brecht, inviting me for the coffee breaks, carrying my tray for me when I had one useful hand, always backing me up in my opinion of people, for being a co-founder of the Mads Mikkelsen fan club. I'm sorry about the times I ate chocolate and offered you none of it, for hitting you on the nose, for rushing past you on your birthday. Unfortunately I have now gotten myself into the situation where I have become your Irish tea dealer, but it's a sacrifice I'm willing to make. In return, we limit the number of mushroom-based meals we share in the future. I'm still working on the list of things you requested during housewives viewings. Most importantly, I'm here for whenever you text me that you are crying in your kitchen because Only Lovers Left Alive is being shown in the cinema. Daniel (Hynds), well, you did your very best to kill me with alcohol poisoning during my first few years! Thanks for the many evenings of (Fr Jack voice) drink!, curries, whiskey, quaich, therapy, photo and stick lessons, for the trip to Scotland, for keeping in contact even after you entered the big leagues. I am sorry my playing the tin whistle entered your head during the trip as a fever dream for you. Not having you around to meet in Nikhef and suggest an after-work drink may have increased my productivity but made me far less happy.

Michele, man, I miss you since you left. The endless stupid jokes we had kept me going during the PhD. You gave me some classic Michele-isms (Caltech, tropical lectures, wait!, big blocks of cheese..) which always made me laugh. Thanks for appreciating beautiful pints with me. I'm glad we have kept in contact from opposite ends of the globe. Jacco, you take the high road, I take the low roadanyways, here's Wonderwall. Thanks to you and Linde for housing a homeless Irish person (it was apparently a common theme in Utrecht), for helping me out when I broke my wrist, for being there on trips abroad and the fun times together. Daniel (Magdalinski), one of the biggest regrets of my PhD is that I didn't get to spend all of it in Nikhef at the same time as you. With that said, I'm so happy we got to spend the last year together, from trips to places in the Netherlands, coffee breaks, going to concerts together, and all the jokes we shared. I'm glad to have had someone to bombard with so much Irish culture and niche jokes and for you to be fully on board with it (I return the favour for all Swedish things). There are not so many people I have met where after just a few weeks of knowing each other, we can sit and talk for hours on end. Thanks in turn for housing and clothing me on evenings where the Dutch weather took it out on us. I hope we remain friends long after I leave.

Lex, though we shared the BND school experience together in Spa, I feel like we became better friends in the last year of our PhDs, and I appreciate it a lot. Thanks for the bouldering sessions (and chalk), the memes, sharing end-of-PhD agonies, and of course our greatest achievement is the formidable playlist we made. I'm sure we'll keep in touch long after our PhDs! Lera, thanks for all the fun times and jokes we shared, for inviting me to Ukrainian music concerts, and for voice messages when you meet crazy English people. Mis chicas Silvia and Cristina, what am I going to do without you around? Silvia, the passion, energy, jokes and character you bring with you wherever you go is amazing, I love you very much. Thanks for all of our long chats and walks together, PhD complaints, jokes we shared, for inspiring me with the thesis title! You're still in my phone contacts as 'the one and only', though I can't remember which of us wrote that. Cristina, the love of my life, I told you I'd write that in my thesis (so, Hannes had plenty of time not to be mad about it). Thank you for the wonderful housewives evenings you hosted, they were a perfect antidote to finishing the PhD. You are so much craic, I regret so much I didn't acknowledge it earlier. Thanks for all the chats, jokes and gossip we shared, I'm glad we got so close. Miriam, we had a Stitch gif to share with the other person for every occasion. Thanks for being my person to dream with about the chandelier lifestyle, while telling each other we deserved Domineaux evenings. Thanks for the jokes, the many chats about life, love, work. We'll always have the movies we watched together in the cinema (good and very bad), and shared that moment of hearing

the most joyful laugh ever. I'll never use the deep-Irish-voice with you again, that was a dark moment. Thanks also for giving me the ultimate tortilla recipe.

Jordy, thank you for being a constant presence and friend during the majority of my PhD years, especially since the first BND school. Thanks for being a great neighbour, such a reliable person and loyal friend. Thanks for the trips we shared during these years, the cinema visits, for being the Iron (/Butter) Bank and unfortunately for documenting my many embarrassing moments by photograph. Efrén, when I returned from holidays to the Nikhef canteen you said "Welcome back to hell!". Maybe, but you made it much less so. You were such a great addition to the Bfys group. On arrival you looked fresh and happy, by the end you were unfortunately broken like many of us. You should know that it takes someone special to me to get me to use voice messages. Sese, thanks also for being a great friend and neighbour, for the movie evenings, for absolutely rocking pub quizzes with me with your insane general knowledge. Ganrong, you were such an amazing addition to the Bfys group for a year. You made me laugh so much, you are so funny, and left us with some golden quotes. I hope we cross paths again and we can watch Ireland play rugby together.

Roman, my dear friend, thank you for the countless pish and countless biertjes we shared. You overshadowed my supervisors by proofreading this thesis at record speed, in about 3 weeks. Thanks for all the trips abroad, the Schloss tour of Germany, the classical music concerts, and for being the ultimate person to complain to about something being too 'Protestant'. Igor, at the first BND school, I sat down beside you and we talked about history and politics, and have been friends since. I have a treasured memory of us at the dinner table in Lunteren, brains melted after the day of talks. Thanks for the many jokes, conversations about history, politics and PhD woes. Krista, thank you for holding my hand every time we went ice-skating, for the nice day trip to Schemer, for the karaoke parties and for always giving me a hug in Nikhef and telling me you had a dream about me. Hilbrand, I had so many nice times with you in the early year of the PhD. Thanks for the movie nights, the BND school experience we shared, and for some jokes that had me in tears. Alice, thanks for the nice times together, we still owe each other a Kings of Leon concert. Carlos, you were actually one of the first people to welcome and bring me into the Bfys group. Thanks for approaching me in the early days of my PhD, connecting me to the LHCb group through table football conversations and the invite to your house party. Roel, you gave me a lot of Python tips and tricks during my early years when we were supervising students together on the GPU-KM3NeT trigger project, and also had some nice conversations together. At risk of going over too many pages, thanks also to Maurício, Elena, Tim, Laurent, MD, Davide, Carolina, Maurice, Emmy, Suzanne,

Andrii, Sean, Robbert, Kazu for all the nice conversations and welcoming me as part of the group.

To the seniors of the LHCb group I also say thanks, for not batting a single eyelid at this random outsider joining the group, and even invading and sitting in the corridor. For this I thank Marcel, Mara, Gerhard, Niels, Antonio, Wouter. Marcel and Mara, you welcomed me so much into the group, and even invited me for talks! I was very honoured and flattered, it gave me a very warm feeling to be embraced by the staff as well as the students of LHCb. Marcel, I had no idea when we first spoke together at the borrel that you were leader of the group, you were so down to Earth. I don't know how you hired so many people that worked and socialised so well together. Thanks for always welcoming me for drinks with the group. Gerhard, thank you for being good craic and a great person to have coffee break conversations with, you were one of the people who kept me going in my last year because of this. I'll always remember the conversation about the tiny Danish actor.

There are other people in Nikhef I want to thank: Federica, for our domestic time together, sushi and pizza dinners; Alice/Alis, for our many gossip sessions, love advice, BND school fun, and sharing All These Things That I've Done with me; Rahul for the nice times together and sharing PhD woes; Marko for the fun times and a very random evening from pizza to beers to guitar songs in your flat; (Spanish) Andrea, for being so much fun to go out with and for the energy you bring. I extend thanks to Mat, (Swiss) Andrea, Alessio, Brian, Ashley, Marc de Beurs, Maria Bader, Navrit, Kevin, Peter Gaemers, Kasper for the nice times, and complaining about the bad times during the Nikhef PhD experience. I want to thank the true backbone of the institute who keep it up-and-running, the secretariaat. I also want to thank members of the PDP group for the help, especially in the last few months to get my reconstruction running on the grid. For this I thank Jeff, Mary, Dennis and David for their time. Thanks also to Pieter for support at the end of my PhD.

I could not have done this thesis without the support of friends outside of Nikhef. Paddy, you shared the PhD experience in the Netherlands with me, and together with Nicolas we embraced the Swamp life. Thanks to both of you for the cycle trip we did to the Veluwe, for the meet-ups in Delft. Pa, thanks for the trip to Athens, and sorry for giving you strike 12 by day one. Aidan (Kelly), you listened to my thoughts and opinions and woes every day. Thanks for remaining my friend during all the turbulent times. Oskar, thanks for being my go-to for English grammar questions. Jess, thanks for the wonderful cover you designed, we have come far from the early days of our PhDs sending money to each other to stay afloat. To the whole TCD physics crowd, thanks

to Cian, Kevin, Aidan, Paddy, Oskar, Laurence, Bevan, Binions, Brendan, Simon, Kate, Conal, Jack, Jess, Abhinav, Cully, for always keeping me humble and being such an amazing crowd of messers. The Gang Dies Abroad trips during my PhD years kept me going. In turn, thanks to the Schnacken crowd of Niall, John, Steve, Seán, Hugh and Hurl for the friendship and pure messing throughout the years. Closer to home, thanks to Aidan (McGrath), Caoimhe, Seán, Jessica, Ryan, Finbar for the support and friendship during these years.

Thanks go to my family, immediate and extended, for all the support during the years. I would like to thank my aunts and uncles for always being so curious about my research and inspiring me further, and to my grandmother for her constant love, to my brother and sisters for growing up with me during this PhD journey. A final thanks go to my parents, I could not have done this without your love and support, and for giving me the freedom to pursue my career so far. *Go raibh míle maith aguibh as ucht bhur ngrá agus bhfur gcabhair*.