

Master's Thesis in Physics

Detection of low-energy tau neutrinos with the ANTARES neutrino telescope - a feasibility study

Michail Chadolias
Supervisor: PD Dr. Thomas Eberl
Erlangen Centre for Astroparticle Physics

Submission Date: 17th of October 2024

Abstract

Neutrino physics has played a pivotal role in modern particle physics. From the postulation of its existence by Pauli in the 1930s and the introduction of neutrino oscillation phenomenology by Pontecovo in the 1950s, to the first experimental evidence of neutrino oscillations in the 1990s, the physics sector has made great strides and constantly grown, with two Nobel Prizes being awarded to related works since the turn of the century. The existence of neutrino oscillations showed that a neutrino of a given flavor has a non-zero probability of being detected after propagation as another flavor. As a result, the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix was developed to describe the mixing and oscillation of neutrinos between their three different flavors. Measurement of the elements of the PMNS matrix and constraining their errors have been an ongoing effort, with the ν_τ appearance measurement being one of them.

The ANTARES detector, a megaton-volume water Cherenkov neutrino telescope located in the Mediterranean Sea at a depth of 2.5 kilometers, operated successfully until its decommissioning in 2022. It consisted of several hundred photomultiplier tubes (PMTs) arranged in a three-dimensional grid. Primarily designed to detect high-energy neutrinos of astrophysical origin, it was also sensitive to all neutrino flavors in the atmospheric neutrino flux with an energy threshold of a few GeV. This work focuses on tau neutrino appearance, i.e. the existence of a tau neutrino atmospheric flux component resulting from neutrino flavor oscillations in the energy range of a few GeV up to 100 GeV. Using data from the full 15-year detector lifetime, an exploratory analysis was undertaken to examine the feasibility of detecting and characterizing a flux of tau neutrinos with the ANTARES detector. This is the first investigation to assess the potential sensitivity of the detector in detecting tau appearance. With a dedicated pipeline being developed to process all reconstructed data, the suitability of a reconstruction algorithm and initial cut selection criteria were identified. The calculation of the tau normalization parameter was achieved with *Swim*, the official tool for oscillation analysis by the Cubic Kilometre Neutrino Telescope (KM3NeT) collaboration. The sensitivities of the best-performing reconstruction algorithms have been determined when the Charged Current (CC) sector is investigated assuming Normal Mass Ordering (NMO). The results of Neural Network Fit (NNFit), the main reconstruction algorithm of this analysis, were compared with smeared versions of "perfectly" reconstructed events, revealing that at its current level, it exhibits a sensitivity of on par with smearing 90%. The potential gains of this analysis from a better performing reconstruction algorithm have also been reviewed.

Contents

1	Introduction to Neutrino Physics	1
1.1	Neutrino Properties	1
1.1.1	Neutrino Interactions	2
1.1.2	Tau Decay	3
1.2	Neutrino Sources	4
1.2.1	Atmospheric Neutrinos	5
1.3	Neutrino Oscillations Phenomenology	6
1.3.1	Neutrino oscillations in vacuum	8
1.3.2	Neutrino oscillations in matter	9
2	Tau Neutrino Appearance	11
2.1	Indirect detection of tau neutrino appearance	12
2.2	Tau neutrino cross-section	14
2.3	Status of current tau neutrino experiments	14
3	ANTARES detector	16
3.1	Science Goals	16
3.2	Detection principle	16
3.3	Detector layout	17
3.4	Detector design	18
3.5	Events in ANTARES	18
3.6	Background sources	19
4	Monte Carlo event processing	21
4.1	ANTARES MC production	21
4.2	"Antares Data Summary Tree"s (AntDSTs)	22
4.3	Processing pipeline	23
4.4	Reconstruction algorithms	30
4.4.1	AAFit	31
4.4.2	BBFit	34
4.4.3	NNFit	37
4.5	Expected number of events	42
4.5.1	Low-energy up-going	43
4.5.2	NNFit up-going Events	44
4.5.3	NNFit quality	45
5	Sensitivity of ANTARES to ν_τ appearance	46
5.1	Swim framework	46
5.1.1	User configurations	48
5.1.2	Analysis pipeline	52
5.2	Reconstruction algorithms	58
5.2.1	Low Energy up-going	59
5.2.2	NNFit up-going	59
5.2.3	NNFit quality	60
5.3	Smeared events	61
5.4	Reconstruction upgrade	62
6	Conclusion	65
6.1	Summary	65
6.2	Outlook	65
A	Oscillograms	71

B	ν_τ weighted distributions	73
C	NNFit file merging	74
D	JSON Files	78
E	Smearing Examples	82
	Bibliography	90

1 Introduction to Neutrino Physics

The Standard Model (SM) constitutes the most complete and established theory currently in the field of particle physics. The particles making the SM are classified as fermions or bosons. The Standard Model explains three of the four known fundamental forces in the universe: electromagnetic, weak, and strong interactions. In the 1960s, the electromagnetic and weak interactions were combined into a single description by Weinberg and Salam, relating the weak nuclear force and electromagnetism of the mechanism known as the Brout-Englert-Higgs mechanism. The mathematical validity of the model was ultimately confirmed, self-consistent (capable of being re-normalized) in the year 1971. The model devised by Glashow, Weinberg, and Salam, along with the theory of strong interactions, which is expressed using the same mathematical structure They reached their ultimate form in the mid-1970s and is now referred to as the Standard Model(SM) pertaining to the field of particle physics.

As aforementioned from the first categories between fermions and bosons. Fermions are particles that make up matter. They obey the Pauli exclusion principle [1], which states that no two fermions can occupy the same quantum state simultaneously. Fermions are divided into quarks and leptons.

- quarks: There exist six flavors of quarks which are known as up, down, charm, strange, top, and bottom. Quarks join together to create particles known as hadrons, including baryons like protons and neutrons, as well as mesons. Quarks are bound by the strong force, which is conveyed through gluons.
- fermions: There are six types of leptons: electron, muon, and tau particles, along with their respective neutrinos - electron neutrino, muon neutrino, and tau neutrino. Leptons are not affected by the strong forces, but the weak force and the electromagnetic force if the lepton involved is electrically charged.

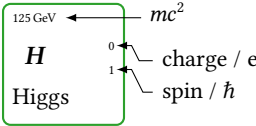
I		II		III				
2.16 MeV	$\frac{2}{3}$	1.27 GeV	$\frac{2}{3}$	172 GeV	$\frac{2}{3}$			
u	$\frac{1}{2}$	c	$\frac{1}{2}$	t	$\frac{1}{2}$			
up		charm		top				
4.67 MeV	$-\frac{1}{3}$	93 GeV	$-\frac{1}{3}$	4.18 GeV	$-\frac{1}{3}$			
d	$\frac{1}{2}$	s	$\frac{1}{2}$	b	$\frac{1}{2}$			
down		strange		bottom				
0.511 MeV	-1	105 MeV	-1	1.78 GeV	-1	0	0	
e	$\frac{1}{2}$	μ	$\frac{1}{2}$	τ	$\frac{1}{2}$	γ	g	
electron		muon		tauon		photon	gluon	
0	0	0	0	0	0	80.4 GeV	91.2 GeV	
ν_e	$\frac{1}{2}$	ν_μ	$\frac{1}{2}$	ν_τ	$\frac{1}{2}$	W^\pm	Z^0	
electron neutrino		muon neutrino		tau neutrino		W-boson	Z-boson	

Figure 1: The list of elementary particles according to the Standard Model. The colour code for the following particles corresponds to orange for quarks, blue for fermion, and green for bosons.

1.1 Neutrino Properties

The neutrino is classified in the list of particles of the SM as a lepton. An overview of the particles can be seen in Figure 1. Neutrinos do not have an electromagnetic charge, neither have color charge, which means they are not involved in the electromagnetic and strong interactions. Neutrinos can only interact with matter through the weak force which is mediated by the electrically neutral Z_0 boson and the electrically charged W_\pm gauge bosons. Given that neutrinos have mass, gravitational interaction can also be applied. However since gravitation is not part of the SM, extended theories beyond SM are needed.

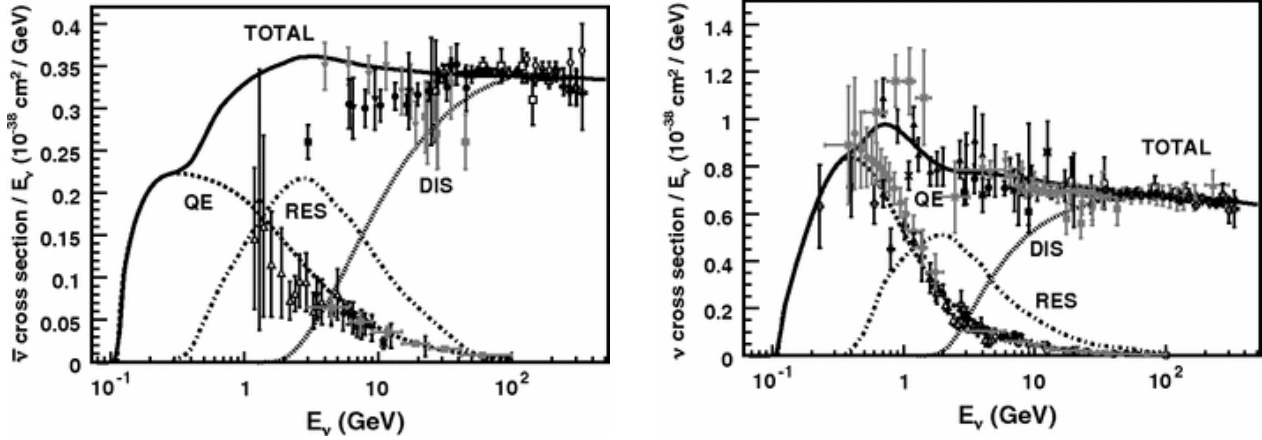


Figure 2: Neutrino cross-section for muon (anti)neutrinos. (Left) the antineutrino and (right) the neutrino cross-section as a function of energy up to 400 GeV [4].

Even though it was presumed to be massless since it wasn't required by SM. The discovery of neutrino oscillations indicated that neutrinos must have a mass since it is a prerequisite for their existence. The phenomenon of “neutrino oscillations” states that neutrinos can modify their flavor state during their propagation through space, although this will be explained further in a later section (see sec:1.3). For the determination of the neutrino mass, dedicated experiments are being carried out such as the KATRIN experiment with the latest result the mass of the electron neutrino is $m_{\nu_e} < 0.8$ meV [2, 3]. Even though the SM is considered to be the most accomplished theory in the field of particle physics, neutrino oscillations are the first experimental evidence that proves that there must be something further. Thus, neutrinos can be potential candidates for probing theories beyond the SM, that could answer fundamental questions.

Likewise, with the charged leptons, the neutrino appears in three different “flavor” generations, typically denoted as ν_e , ν_μ , and ν_τ . Furthermore, like charged leptons neutrinos have their antiparticle counterparts denoted, as $\bar{\nu}_e$, $\bar{\nu}_\mu$ and $\bar{\nu}_\tau$.

1.1.1 Neutrino Interactions

As already mentioned neutrinos interact through the weak force. Neutrinos can interact with nucleons, electrons, nuclei via both interaction processes, CC and Neutral Current (NC) interaction. The mediating bosons are W^\pm for CC and Z^0 for NC interactions.

Figure:2 shows the different contributions to the ν_e CC cross-section as a function of energy. For energies below the GeV scale, the dominant process is quasi-elastic (QE) interaction. In this process, the neutrino sees the ‘whole’ nucleon, and a corresponding charged particle is produced. In the case of NC, the neutrino that interacts remains in the final state. Conservation of charge and leptonic number allow for two possible scenarios depending on the target nucleon. If the target nucleon is neutron, this can interact with a neutrino producing the corresponding negatively charged lepton, and a positively charged proton. On the other hand, an antineutrino can interact with a proton producing a neutron and the corresponding charged anti-lepton. From energies above ~ 1 GeV, in the resonance production (RES) process, the nucleon produces a baryon resonance state, which then decays in a sea of nucleons and mesons. For energies above a few GeV, the neutrino can resolve the internal structure of the nucleon and interact directly with the individual constituents, breaking apart the nucleon and producing a cascade of hadrons, a process known as the deep inelastic scattering (DIS). For energies above ~ 10 GeV, DIS dominates and can be accomplished from both NC and CC.

A reminder before moving to the next section, a recap of the reactions is listed here

- Quasi-elastic scattering: $\nu_\mu + n \rightarrow \mu^- + p$
- Resonant inelastic scattering: $\nu_\mu + n \rightarrow \mu^- + \Delta^+ \rightarrow \mu^- + p + \pi^0$
- Deep-inelastic scattering: $\nu_\mu + N \rightarrow X + \mu^-$

$$\begin{array}{lcl}
\tau^- & \rightarrow & e^- \bar{\nu}_e \nu_\tau \quad (17.8\%) \\
\tau^- & \rightarrow & \mu^- \bar{\nu}_\mu \nu_\tau \quad (17.4\%) \\
\tau^- & \rightarrow & \pi^- \pi^0 \nu_\tau \quad (25.5\%) \\
\tau^- & \rightarrow & \pi^- \nu_\tau \quad (10.8\%) \\
\tau^- & \rightarrow & \pi^- \pi^0 \pi^0 \nu_\tau \quad (9.3\%) \\
\tau^- & \rightarrow & \pi^- \pi^- \pi^+ \nu_\tau \quad (9.0\%) \\
\tau^- & \rightarrow & \pi^- (> 2\pi) \nu_\tau \quad (9.0\%)
\end{array}
\left. \vphantom{\begin{array}{l} \tau^- \\ \tau^- \\ \tau^- \\ \tau^- \\ \tau^- \\ \tau^- \\ \tau^- \end{array}} \right\} \begin{array}{ll} \textbf{leptonic} & \textbf{(35.2\%)} \\ \textbf{hadronic} & \textbf{(64.8\%)} \end{array}$$

h!

Figure 4: τ branching ratios for each main decay. Taken from [7]

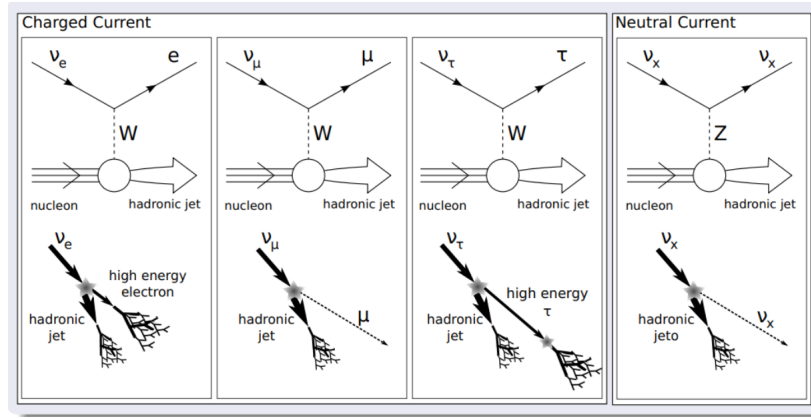


Figure 3: Illustrations of neutrino interactions for both NC and CC for deep-inelastic scattering. Taken from [5]

It should be noted, that in this work, given that the neutrino flux studied is in the range of tens of GeV the dominant interaction is deep-inelastic scattering (see Figure:3). A final reminder of the two interactions for the neutrino telescopes, for NC

$$\nu_x + N \rightarrow \nu_x + \text{hadronic jet} \quad x = e, \mu, \tau \quad (1)$$

and the CC

$$\nu_x + N \rightarrow x + \text{hadronic jet} \quad x = e, \mu, \tau \quad (2)$$

1.1.2 Tau Decay

From the Particle Data Group (PDG) the tau lepton has a lifetime of $\tau_\tau = 2.9 \times 10^{-13}\text{s}$ [6]. In the GeV ranges, where the transition probabilities to tau neutrinos are the highest, the lepton only traverses less than a millimeter before it decays. The final states from such a decay and their branching ratios are well-defined and documented also in the PDG. The main decay modes with their branching ratios are gathered in Figure:4

In practice for a Cherenkov detector, the spatial reconstruction capabilities of these experiments are not sufficient enough to distinguish individual particles in the final state and therefore a distinction between the interaction vertex and the tau decay vertex is not possible. Another important detail is that since the vertices cannot be separated, the hadronic and the leptonic shower will overlap, thus trying to isolate the two components is also not feasible. Therefore the topologies that can be produced by the tau are both shower-like and track-like. The track-like case appears in 17.4% of the events when the tau lepton decays into a muon. A muon will appear inside the detector as a track and in principle, it couldn't be identified if it was from a ν_μ or a ν_τ . Thus the majority of the events will have a shower-like signature inside the detector with 64.5% coming

from the hadronic decays and another 17.8% from taus decaying into electrons. Another thing to consider is that, since neutrinos are produced in the tau-lepton decays, and these typically escape from the detector without interacting, a fraction of the tau lepton's energy will be invisible, thus lowering the reconstructed energy of the original neutrino.

1.2 Neutrino Sources

Neutrinos are the most abundant particles in the Universe, after photons. Their energy spans a wide range of orders of magnitude (see Figure:5). Their flux is likewise widely varied from 10^{20} to $10^{-24} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \text{ MeV}^{-1}$.

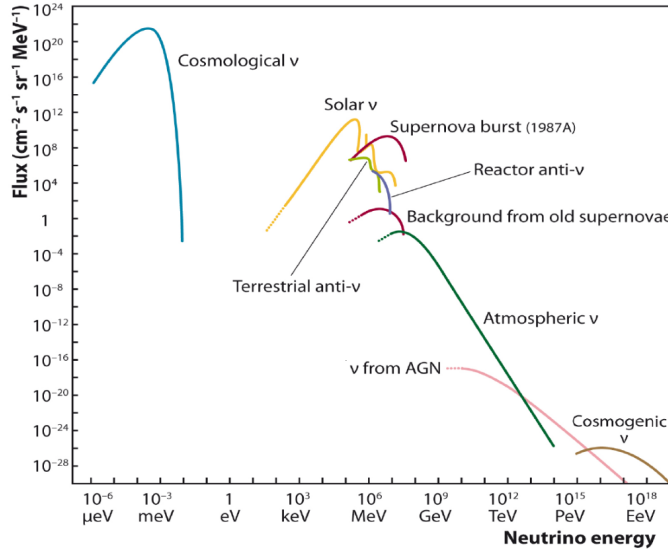


Figure 5: Measured and expected fluxes of natural and reactor neutrinos. Taken from [8].

Neutrinos can be produced by a variety of sources both natural and man-made. Starting from the lowest energy ones to the highest, the **cosmological neutrinos** (ν) are remnants from the freeze-out period after the Big Bang[9]. **Solar** ν are products from the various chains taking place in the Sun as part of nuclear fusion reactions, the main chain being the proton-proton chain that starts with the production of deuterium $p + p \rightarrow d + e^+ + \nu_e$ [10]. **Terrestrial neutrinos** originated from radioactive decays in Earth's core and are sources of $\bar{\nu}_e$. **Reactor anti-neutrinos** are due to radioactive decay inside a nuclear reactor and are mainly $\bar{\nu}_e$.

Furthermore, **supernova** explosions (SN) result in a massive emission of neutrinos. Depending on the distance of the source, neutrino events can be detected below the energy threshold of the neutrino telescope even if a sufficient rate is achieved.

From GeV to PeV **atmospheric neutrinos** dominate this region, where Cherenkov detectors are the most sensitive. At GeV energies, neutrinos are produced in the Earth's atmosphere after the interaction of cosmic rays with the atmosphere. This region is also going to be the focus of this study since neutrino oscillation studies are most sensitive in this region. After the TeV-scale, the flux from **astrophysical sources** (or ν from AGN) overcomes the atmospheric one.

On the highest point of the spectrum, **cosmogenic neutrinos** are generated in the interactions of Ultra-High-Energy (UHE) cosmic rays with photons (γ) of the Cosmic Microwave Background (CMB) like the following $N + \gamma \rightarrow N + p^+$ [11]. Neutrinos are produced from the decay of pions, which will be explained more in-depth in the next section. The highest energies of neutrinos can be achieved from protons of the GZK cut-off [12]. Although no events have been detected yet, future radio detectors aim to reach this energy region.

1.2.1 Atmospheric Neutrinos

As stated beforehand, primary cosmic rays interact with nuclei from the atmosphere to produce atmospheric neutrinos. The Earth's upper atmosphere is constantly bombarded by high energy Cosmic Rays (CR) from outer space. They are composed mainly of protons (95%) with the remaining being helium 4% with the remaining being heavier nuclei up to iron (1%). The collisions between the initial cosmic rays and the atoms in the atmosphere produce secondary cosmic rays, such as hadrons and their decay products. These secondary cosmic rays have an energy distribution peaking in the GeV range and gradually extending to higher energies following a power law pattern [13].

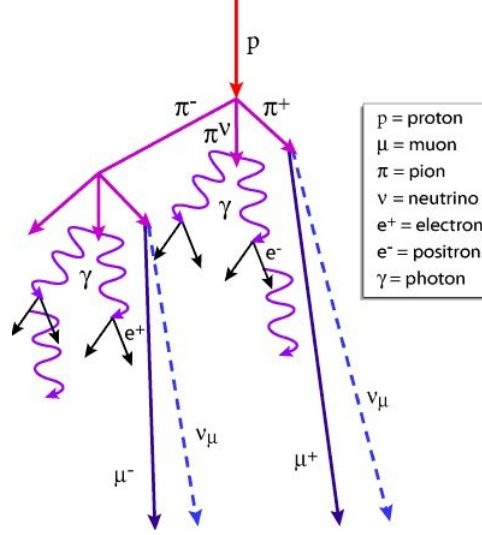


Figure 6: Air shower schematic. Taken from [14]

From the secondary particles, pions (π^\pm) and kaons (K^\pm) decays can subsequently produce neutrinos. The decays of charged kaons:

$$K^+ \rightarrow \mu^+ + \nu_\mu \quad (3)$$

$$K^+ \rightarrow \pi^+ + \pi^0, \pi^+ + \pi^+ + \pi^-, \pi^+ + \pi^0 + \pi^0 \quad (4)$$

$$K^+ \rightarrow \pi^0 + e^+ + \nu_e, \pi^0 + \mu^+ + \nu_\mu \quad (5)$$

The neutral kaons lastly can decay according to their modes either short or long (K_S and K_L) respectively. For the short-kaon:

$$K_S^0 \rightarrow \pi^0 + \pi^0 \quad (6)$$

$$K_S^0 \rightarrow \pi^+ + \pi^- \quad (7)$$

Whereas the long-kaon:

$$K_L^0 \rightarrow \pi^\pm + e^\mp + \bar{\nu}_e \quad (8)$$

$$K_L^0 \rightarrow \pi^\pm + \mu^\mp + \bar{\nu}_\mu \quad (9)$$

$$K_L^0 \rightarrow 3\pi^0 \quad (10)$$

$$K_L^0 \rightarrow \pi^+ + \pi^- + \pi^0 \quad (11)$$

The interactions for pions can be shown below as:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \quad (12)$$

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu \quad (13)$$

Lastly, the flux of muons produced by all these interactions mentioned above depending on their energy per particle can decay before reaching sea level as:

$$\mu^+ \rightarrow \bar{\nu}_\mu + \nu_e + e^+ \quad (14)$$

From equations 14 and 12, in the final state of events it is expected that the ratio of muon neutrinos to electron neutrinos is close to two[15].

$$(\bar{\nu}_\mu + \nu_\mu)/(\nu_e + \bar{\nu}_e) \simeq 2 \quad (15)$$

From the aforementioned reaction chain, the atmospheric neutrino flux should only consist of ν_e and ν_μ , thus the existence of ν_τ in this energy range of GeV can only be due to neutrino oscillations (see more regarding oscillation in sec:1.3). Up to this point the aforementioned mechanisms contribute to the "conventional" neutrino flux, however in much higher energies another contribution called "prompt" has been postulated [13]. The "prompt" flux is derived from the decay of charmed mesons (such as D^0 , D^+ , B^0 , B^+ , etc), which can only be produced in energies above 1 TeV [16]. According to models, the "prompt" component overtakes the "conventional" at energies above 100 TeV. Experimental evidence for the "prompt" component is challenging, since at these energies the astrophysical flux becomes the dominant one. It is also theorized, that through the decays of the charmed mesons, taus can be produced and after its decay tau neutrinos too.

This ratio, Equation:15, holds only at low energies as one moves toward higher energies and at the nadir/zenith, the ratio rises. This is a result of an increasing proportion of muons reaching the ground instead of decaying in the atmosphere. They only produce low-energy neutrinos there after being halted. Large-volume neutrino telescopes do not detect these either. Additionally, the contribution from kaon decays becomes increasingly significant at higher energy [17].

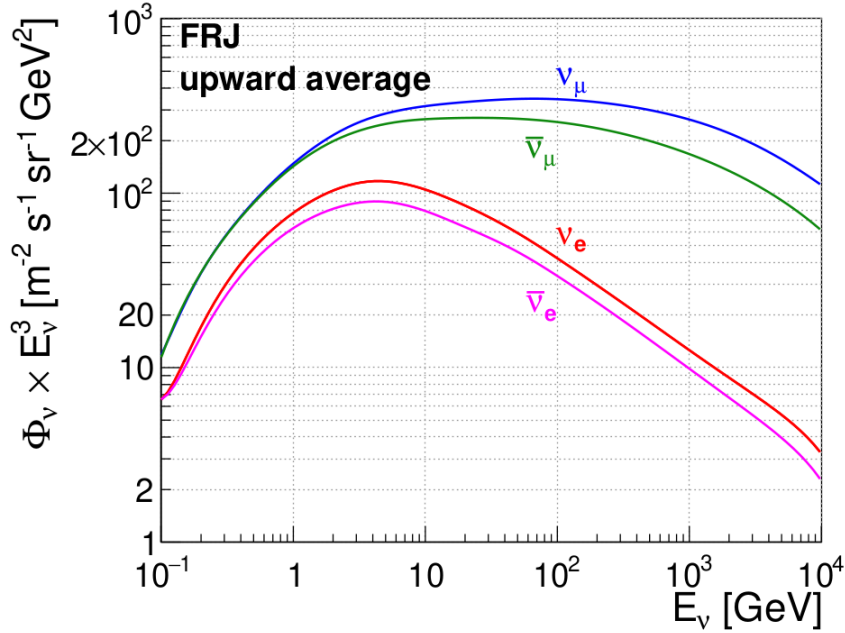


Figure 7: HHKM model prediction for the Frejus site averaged over an up-going direction as a function of energy. Taken from [7]

1.3 Neutrino Oscillations Phenomenology

The phenomena by which a neutrino can be detected after propagating a given distance from the interaction point in a flavor state distinct from the state in which it originated is known as "neutrino oscillation". In the late 1950s, Pontecorvo, hypothesized the presence of neutrino oscillations, a quantum mechanical process that is comparable to $K^0 - \bar{K}^0$ oscillations. Because of their extremely small mass differences, various massive neutrinos interact to produce and detect coherent oscillations [13]. It was 40 years after the hypothesis, in fact

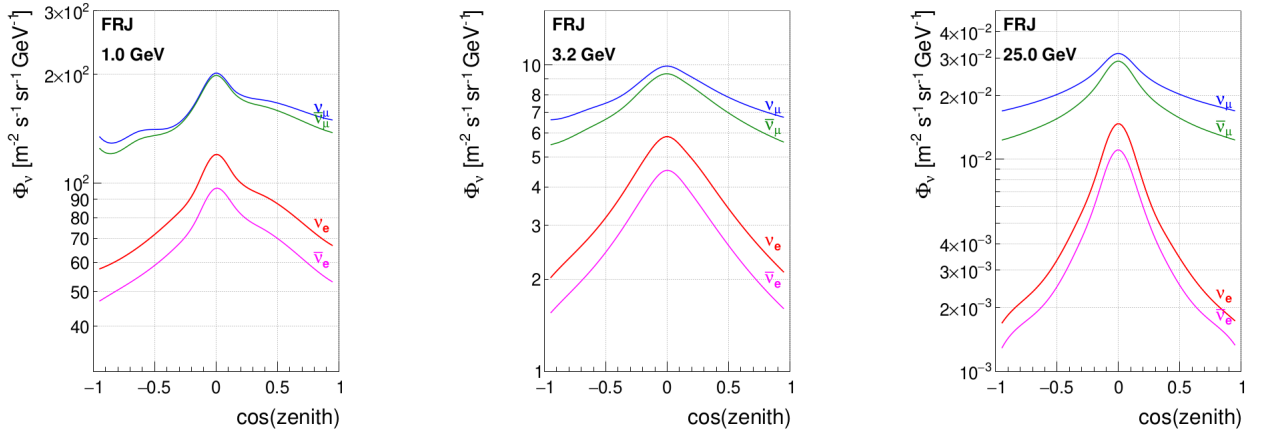


Figure 8: HHKM model prediction for the Frejus site for events at an energy 3.2 GeV as a function of the cosine of their zenith angle. Taken from [7].

in 1998, that the Super-Kamiokande (SK) experiment produced sufficient evidence for neutrino oscillations by figuring the deficit of the measured muon neutrinos by using the atmospheric neutrino flux [18]. It wasn't long after, that evidence of neutrino oscillations was found in the solar neutrino flux. It was, in fact, the Sudbury Neutrino Observatory (SNO) experiment, which not only provided additional evidence but gave an explanation to the "solar neutrino problem". This inconsistency with theoretical models was first detected by the Homestake experiment in the 60s [19]. Merely over 33% of the anticipated neutrinos were detected in this experiment. Other experiments also showed similar deficiencies, with measured rates varying from one-third to one-half of the expected values. The Homestake Experiment was only sensitive to electron neutrinos, despite the fact that estimations of the expected neutrino flux were accurate. Both NC interactions involving all three types of neutrinos and CC interactions involving only electron neutrinos might be detected by the SNO experiment. The fraction of electron neutrinos found was in line with previous experiment results, and the NC flux matched the estimated total neutrino flux, according to the results. It was the SNO, that finally gave a conclusive explanation to this deficit through neutrino oscillations. The discovery of neutrino oscillations was such a significant discovery, that was awarded even the Nobel Prize in 2015 to both collaborations [20].

Three basis vectors are superposed to define a neutrino state $|\nu\rangle$ in the most widely used model of neutrino oscillations at the moment. This image contains two bases that are worth mentioning. The basis vectors $|\nu_e\rangle$, $|\nu_\mu\rangle$, and $|\nu_\tau\rangle$ make-up the first, known as the flavor basis. In this base, a generic superposition takes the form

$$|\nu\rangle = \sum_{\alpha=e,\mu,\tau} c_\alpha |\nu_\alpha\rangle \quad (16)$$

where the complex values that meet the normalization criterion ($\sum_a |c_\alpha|^2 = 1$) are the coefficients c_α .

The flavor basis vectors match the Standard Model's neutrino states as mentioned in the first part and referenced in fig. 1. A neutrino is always created as a flavor basis state when it interacts via weak force such as beta decay or a deuterium reaction. Similarly, a neutrino measured through a weak interaction process (such as an inverse beta decay) will invariably be observed in a base state with a flavor. In the case of neutrino oscillations, a neutrino in a flavor state ν_a in its production point can be detected as another flavor state ν_β at another point in space. On the other hand, the mass states are eigenstates of the Hamiltonian H_0 , which determines the propagation of neutrinos. It should be noted, that even if the two basis states are correlated they are different. Thus for ease of understanding, the mass basis is denoted by the Arabic numerals like ν_i with $i = 1, 2, 3$ and the flavor basis ν_a is denoted by the Greek letters that are also used to denote the three different charged lepton generations $\alpha = e, \mu, \tau$.

Neutrinos oscillate between distinct flavors because their flavor eigenstates ν_α , are not equal to their mass eigenstates ν_i . Instead, a linear combination links these states together:

$$|\nu_a\rangle = \sum_i^N U_{ai}^* |\nu_i\rangle \rightarrow |\nu_i\rangle = \sum_{\alpha=e,\mu,\tau} U_{ai} |\nu_a\rangle \quad (17)$$

with

$$U^\dagger U = 1 \quad \sum_i U_{ai} U_{\beta i}^* = \delta_{a\beta} \quad \sum_a U_{ai} U_{aj}^* = \delta_{ij} \quad (18)$$

Likewise in the case of an antineutrino state, the complex conjugate has to be used

$$|\bar{\nu}_a\rangle = \sum_i U_{ai}^* |\bar{\nu}_i\rangle \quad (19)$$

This complex linear combination can only be achieved because these bases describe the same Hilbert space. More specifically equation 17 can be written as:

$$\begin{pmatrix} |\nu_e\rangle \\ |\nu_\mu\rangle \\ |\nu_\tau\rangle \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{pmatrix} \cdot \begin{pmatrix} |\nu_1\rangle \\ |\nu_2\rangle \\ |\nu_3\rangle \end{pmatrix} \quad (20)$$

In the 3-flavor model, the U_{ai} denotes the elements of the neutrino mixing matrix also known as PMNS. The parameterisation of the aforementioned matrix is done in the following manner:

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{13}s_{23}e^{i\delta} & c_{12}c_{23} - s_{12}s_{13}s_{23}e^{i\delta} & c_{13}s_{23} \\ s_{12}s_{23} - c_{12}s_{13}c_{23}e^{i\delta} & -c_{12}s_{23} - s_{12}s_{13}c_{23}e^{i\delta} & c_{13}c_{22} \end{pmatrix} \quad (21)$$

where s_{ij} stands for $\sin\theta_{ij}$ and c_{ij} stands for $\cos\theta_{ij}$ with θ_{ij} being the three real mixing angles and δ is a phase angle taking into account possible influence from CP violation. The PMNS is assumed to be unitary, resulting in the neutrino flux to be always conserved.

1.3.1 Neutrino oscillations in vacuum

Assuming that a flavor eigenstate $|\nu_i\rangle$ propagates as a plane wave, the following equation holds for the time evolution of the eigenstate:

$$|\nu_i(t)\rangle = \exp(-i(E_i t - p_i r)) |\nu_i(t=0)\rangle \quad (22)$$

The mass eigenstate's position is denoted by r , while its energy and momentum are represented by E_i and p_i . The expression can be simplified in the relativistic limit approximation as

$$|\nu_i(t)\rangle = \exp\left(-i\frac{m_i^2 L}{2E}\right) |\nu_i(t=0)\rangle \quad (23)$$

The neutrino energy, denoted by E in this equation, and the propagation length, denoted by L . The evolution of the flavor state during its propagation is given by

$$\begin{aligned} |\nu_a(L, E)\rangle &= \sum_{i=1}^3 \langle \nu_i(L) | \nu_a(L) | \nu_i(L) \rangle = \sum_{i=1}^3 U_{ai} \exp\left(-i\frac{m_i^2 L}{2E}\right) |\nu_i\rangle \\ &= \sum_{i=1}^3 \sum_{\beta=e,\mu,\tau} U_{ai} \exp\left(-i\frac{m_i^2 L}{2E}\right) U_{\beta i}^* |\nu_b\rangle \end{aligned} \quad (24)$$

where U^* is the complex conjugate of matrix U . Consequently, the probability of a neutrino in a flavor state ν_a after propagation of a distance L at an energy E can be obtained as a projection of the initial state $\nu_a(t)$ onto the flavor state ν_β in accordance to Eq. (17)

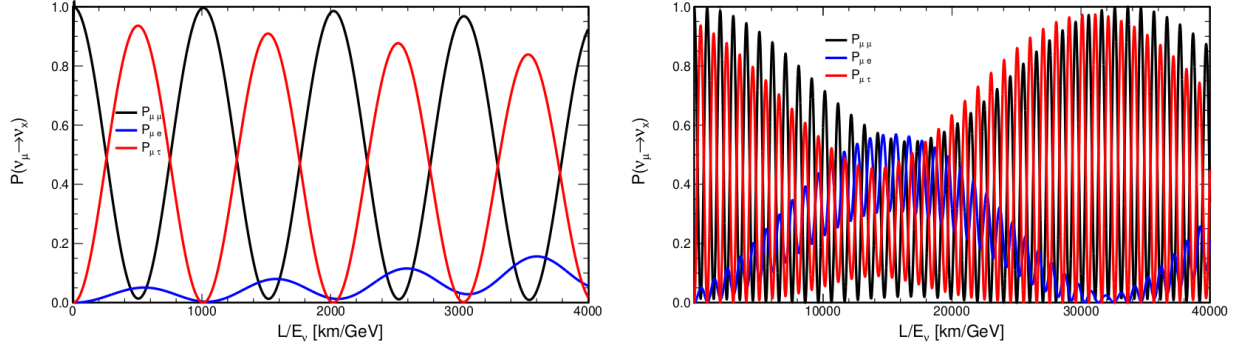


Figure 9: Muon neutrino ν_μ survival and transition probabilities in vacuum in the three-flavor model as a function of the neutrino path over energy (L/E). (Left) influence of Δm_{31}^2 in the shorter ranges of (L/E), (right) effect of both mass square differences. Taken from [21]

$$\begin{aligned}
 P_{a \rightarrow \beta}(L, E) &= |\langle \nu_\beta | \nu_a(L, E) \rangle|^2 = \sum_{ij} U_{\beta i}^* U_{a j}^* U_{a i} U_{\beta j} \exp \left(-i \frac{\Delta m_{ij}^2 L}{2E} \right) = \\
 &= \left| U_{a1} U_{\beta 1}^* e^{-i \frac{m_1^2 L}{2E}} + U_{a2} U_{\beta 2}^* e^{-i \frac{m_2^2 L}{2E}} + U_{a3} U_{\beta 3}^* e^{-i \frac{m_3^2 L}{2E}} \right|
 \end{aligned} \tag{25}$$

where Δm_{ij} is the squared-mass difference and is defined as $\Delta m_{ij} = m_i^2 - m_j^2$. From the exponent of Eq. (25), one can try to visualize it as an oscillation with the first term being L/E and the wavelength being given by the squared-mass difference Δm_{ij} , hence the "oscillation" justification in the term. It is clear that in the context of the mass eigenstate basis, two distinct states play the dominant factor. The first is the Δm_{21}^2 , also known as the "solar" one, with a value approximately of $7.5 \times 10^{-5} \text{ eV}^2$ [6]. The second is Δm_{31}^2 , also known as the "atmospheric" one, with a magnitude of $2.5 \times 10^{-5} \text{ eV}^2$.

1.3.2 Neutrino oscillations in matter

Up to this point only propagation in vacuum was discussed. However when a neutrino traverses a material medium CC and NC interactions that can occur as discussed in earlier sections. These in turn can include additional potential that affect the neutrino. Concerning neutral currents, the neutrino flavor has no separate impact on the effect, thus neutrino oscillation dynamics are not influenced by NC. Although, regarding charged current this is not the case. Since matter is only comprised of electrons and not the other leptons, there is an asymmetry; therefore, when an electron neutrino travels through the material medium an additional phase from forward scattering interaction occurs [21]. The phenomenon in question is known as the Mikheyev-Smirnov-Wolfenstein (MSW) effect.

Under the MSW only to electron neutrino an additional potential is included

$$V = \pm \sqrt{2} G_F n_e(x) \tag{26}$$

with G_F being the Fermi constant and $n_e(x)$ the electron density. The two possible signs account for neutrinos and antineutrinos, the positive being for the first and the negative for the latter. The "effective Hamiltonian" expresses neutrino oscillations in matter and is derived by including the potential term to the Hamiltonian in vacuum

$$H(x) = U \begin{pmatrix} 0 & 0 & 0 \\ 0 & \Delta m_{21}^2 & 0 \\ 0 & 0 & \Delta m_{31}^2 \end{pmatrix} U^\dagger + \begin{pmatrix} V(x) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{27}$$

In the case of a constant matter density, an analytic solution is feasible. Within the 3-flavor framework, the transitional probabilities are

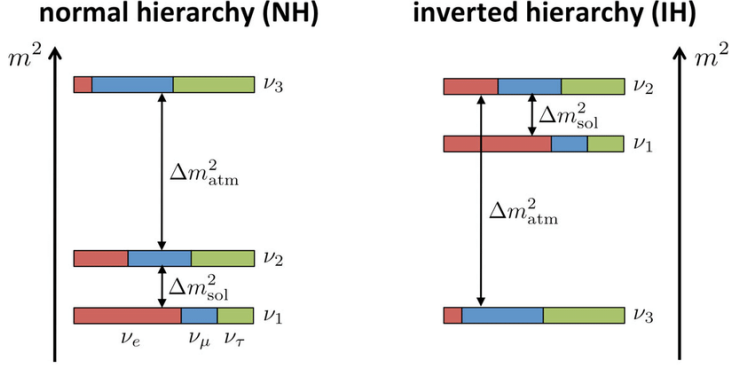


Figure 10: A visual representation of the two alternative mass orderings. (Left) the NMO, where the atmospheric mass-squared is positive, and (right) the IMO, where the mass difference is negative, is presented. The mixture of the flavor eigenstates to the mass eigenstates is colorized with ν_e (red), ν_μ (blue) and ν_τ (green). Taken from [23]

$$P(\nu_\mu \rightarrow \nu_e) = \sin^2(\theta_{23}) \sin^2(2\theta_{13}^M) \sin^2\left(\Delta^M \frac{L}{4E}\right) \quad (28)$$

$$\begin{aligned} P(\nu_\mu \rightarrow \nu_\mu) = & 1 - \sin^2(\theta_{23}) \cos^2(2\theta_{13}^M) \sin\left(\frac{(\Delta m_{31}^2 + \Delta^M)L}{8E} + \frac{VL}{4}\right) - \\ & - \sin^2(2\theta_{23}) \sin^2(2\theta_{13}^M) \sin\left(\frac{(\Delta m_{31}^2 - \Delta^M)L}{8E} + \frac{VL}{4}\right) - \\ & - \sin^4(\theta_{23}) \sin^2(2\theta_{13}^M) \sin^2\left(\Delta^M \frac{L}{4E}\right) \quad (29) \end{aligned}$$

Additionally, the effective mixing parameters can be expressed according to the normal mixing parameters

$$\sin^2(2\theta_{13}^M) = \sin^2(2\theta_{13}) \left(\frac{\Delta m_{31}^2}{\Delta^M}\right)^2 \quad (30)$$

$$\Delta^M = \sqrt{(\Delta m_{31}^2 \cos(2\theta_{13}) - 2VE)^2 + (\Delta m_{31}^2 \sin(2\theta_{13}))^2} \quad (31)$$

Equations and , provide important information on the effects of the MSW effect. The oscillation probability, θ_{13} , is the result of the effective angle's sine modulation by the magnitude M. The determination of the sign of Δm_{21}^2 has been accomplished through observation of matter effects in the Sun by putting constraints on the product $\Delta m_{21}^2 \cos(2\theta_{12})$ [22]. Notice that the sign of the atmospheric mass splittings is currently unknown. Ergo, there are two alternative nonequivalent orderings, the first possibility being NMO and the being latter Inverted Mass Ordering (IMO), for the mass eigenvalues given the observed hierarchy between the solar and atmospheric wavelengths:

- $m_1 \ll m_2 < m_3$ so $\Delta m_{21}^2 \ll \Delta m_{32}^2$ ($\simeq \Delta m_{31}^2 > 0$), referred to as NMO
- $m_3 \ll m_1 < m_2$ so $\Delta m_{21}^2 \ll -(\Delta m_{31}^2 \simeq \Delta m_{32}^2 < 0)$, referred to as IMO

For a visual representation of the two mass orderings/mass hierarchies see (Figure:10).

2 Tau Neutrino Appearance

The phenomenon of neutrino flavor oscillations and their experimental proof has already been explained in a previous section (see sec:1.3). The majority of the neutrino experiments [24] are mainly focused on measuring the appearance and the disappearance of electron neutrinos (ν_e) and muon neutrinos (ν_μ). By measuring the survival and transition probabilities of the various flavors, the 3-flavor oscillation parameters and the elements of the PMNS matrix can be derived. The leptonic mixing matrix and neutrino masses are estimated using an updated global analysis of neutrino oscillation measurements within the framework of the Standard Model with three massive neutrinos and some of its extensions by NuFit [25] with the current version being 5.3.

NuFIT 5.3 (2024)					
without SK atmospheric data		Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 2.3$)	
		bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range
	$\sin^2 \theta_{12}$	$0.307^{+0.012}_{-0.011}$	$0.275 \rightarrow 0.344$	$0.307^{+0.012}_{-0.011}$	$0.275 \rightarrow 0.344$
	$\theta_{12}/^\circ$	$33.66^{+0.73}_{-0.70}$	$31.60 \rightarrow 35.94$	$33.67^{+0.73}_{-0.71}$	$31.61 \rightarrow 35.94$
	$\sin^2 \theta_{23}$	$0.572^{+0.018}_{-0.023}$	$0.407 \rightarrow 0.620$	$0.578^{+0.016}_{-0.021}$	$0.412 \rightarrow 0.623$
	$\theta_{23}/^\circ$	$49.1^{+1.0}_{-1.3}$	$39.6 \rightarrow 51.9$	$49.5^{+0.9}_{-1.2}$	$39.9 \rightarrow 52.1$
	$\sin^2 \theta_{13}$	$0.02203^{+0.00056}_{-0.00058}$	$0.02029 \rightarrow 0.02391$	$0.02219^{+0.00059}_{-0.00057}$	$0.02047 \rightarrow 0.02396$
	$\theta_{13}/^\circ$	$8.54^{+0.11}_{-0.11}$	$8.19 \rightarrow 8.89$	$8.57^{+0.11}_{-0.11}$	$8.23 \rightarrow 8.90$
	$\delta_{CP}/^\circ$	197^{+41}_{-25}	$108 \rightarrow 404$	286^{+27}_{-32}	$192 \rightarrow 360$
	$\frac{\Delta m_{21}^2}{10^{-5} \text{ eV}^2}$	$7.41^{+0.21}_{-0.20}$	$6.81 \rightarrow 8.03$	$7.41^{+0.21}_{-0.20}$	$6.81 \rightarrow 8.03$
	$\frac{\Delta m_{3\ell}^2}{10^{-3} \text{ eV}^2}$	$+2.511^{+0.027}_{-0.027}$	$+2.428 \rightarrow +2.597$	$-2.498^{+0.032}_{-0.024}$	$-2.581 \rightarrow -2.409$
with SK atmospheric data		Normal Ordering (best fit)		Inverted Ordering ($\Delta\chi^2 = 9.1$)	
		bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range
	$\sin^2 \theta_{12}$	$0.307^{+0.012}_{-0.011}$	$0.275 \rightarrow 0.344$	$0.307^{+0.012}_{-0.011}$	$0.275 \rightarrow 0.344$
	$\theta_{12}/^\circ$	$33.67^{+0.73}_{-0.71}$	$31.61 \rightarrow 35.94$	$33.67^{+0.73}_{-0.71}$	$31.61 \rightarrow 35.94$
	$\sin^2 \theta_{23}$	$0.454^{+0.019}_{-0.016}$	$0.411 \rightarrow 0.606$	$0.568^{+0.016}_{-0.021}$	$0.412 \rightarrow 0.611$
	$\theta_{23}/^\circ$	$42.3^{+1.1}_{-0.9}$	$39.9 \rightarrow 51.1$	$48.9^{+0.9}_{-1.2}$	$39.9 \rightarrow 51.4$
	$\sin^2 \theta_{13}$	$0.02224^{+0.00056}_{-0.00057}$	$0.02047 \rightarrow 0.02397$	$0.02222^{+0.00069}_{-0.00057}$	$0.02049 \rightarrow 0.02420$
	$\theta_{13}/^\circ$	$8.58^{+0.11}_{-0.11}$	$8.23 \rightarrow 8.91$	$8.57^{+0.13}_{-0.11}$	$8.23 \rightarrow 8.95$
	$\delta_{CP}/^\circ$	232^{+39}_{-25}	$139 \rightarrow 350$	273^{+24}_{-26}	$195 \rightarrow 342$
	$\frac{\Delta m_{21}^2}{10^{-5} \text{ eV}^2}$	$7.41^{+0.21}_{-0.20}$	$6.81 \rightarrow 8.03$	$7.41^{+0.21}_{-0.20}$	$6.81 \rightarrow 8.03$
	$\frac{\Delta m_{3\ell}^2}{10^{-3} \text{ eV}^2}$	$+2.505^{+0.024}_{-0.026}$	$+2.426 \rightarrow +2.586$	$-2.487^{+0.027}_{-0.024}$	$-2.566 \rightarrow -2.407$

Figure 11: Three-flavor oscillation parameters from our fit to global data as of March 2024. The recent results are presented with the left side representing the NMO and the right side the IMO. Furthermore, the lower part includes data from SK, whereas the upper part doesn't [25].

The aforementioned oscillation parameters for the 3-flavor model are six independent parameters: θ_{12} , θ_{13} , θ_{23} , δ_{CP} and the two squared mass differences Δm_{21}^2 and Δm_{32}^2 (see fig:11).

	NuFIT 5.3 (2024)
$ U _{3\sigma}^{\text{w/o SK-atm}} = \begin{pmatrix} 0.801 \rightarrow 0.842 & 0.518 \rightarrow 0.580 & 0.142 \rightarrow 0.155 \\ 0.236 \rightarrow 0.507 & 0.458 \rightarrow 0.691 & 0.630 \rightarrow 0.779 \\ 0.264 \rightarrow 0.527 & 0.471 \rightarrow 0.700 & 0.610 \rightarrow 0.762 \end{pmatrix}$	
$ U _{3\sigma}^{\text{with SK-atm}} = \begin{pmatrix} 0.801 \rightarrow 0.842 & 0.518 \rightarrow 0.580 & 0.143 \rightarrow 0.155 \\ 0.244 \rightarrow 0.500 & 0.498 \rightarrow 0.690 & 0.634 \rightarrow 0.770 \\ 0.276 \rightarrow 0.521 & 0.473 \rightarrow 0.672 & 0.621 \rightarrow 0.759 \end{pmatrix}$	

Figure 12: The 3σ confidence intervals of the absolute values of the PMNS matrix according to NuFit5.3 with/out the measurements of SK [26, 27]

The largest uncertainties in the PMNS matrix come from the tau sector where a possible deviation from unitarity is one order of magnitude larger compared to the electron and muon sector. This prevents us from constraining unitarity better than 20-40 % [28]. These matrix elements are often determined by using the other elements and assuming unitarity. This way, unitarity is a prerequisite for the determination of these elements but is not tested and inferred. If unitarity is not assumed, the 3σ uncertainties on the matrix elements increase, especially in case of $|U_{\tau i}|$. The precision on $|U_{\tau 3}|$ can even vary roughly around 100%. Hence, to test the unitarity it is essential to precisely probe the tau neutrino sector.

2.1 Indirect detection of tau neutrino appearance

In general, three issues introduce complications in the measurement of ν_τ

- For the detection of the τ^\pm track great precision in millimeters is called for since the charged lepton has a low lifetime
- The visible energy is reduced since there are neutrinos produced by the tau decay, transferring part of the energy that the τ carried
- the cross-section is suppressed in contrast to the other flavors and the energy required for the production of τ is a few GeV

Long baseline experiment can conduct measurements for the θ_{23} with appearance from $\nu_\mu \rightarrow \nu_\mu$ probing $|U_{\mu 3}|$ and with appearance $\nu_\mu \rightarrow \nu_\tau$ probing into $|U_{\mu 3}|$ and $|U_{\tau 3}|$. Theoretically, for the measurement of $|U_{\tau 3}|$ three channels contribute $\nu_e \rightarrow \nu_\tau$, $\nu_\mu \rightarrow \nu_\tau$, $\nu_\tau \rightarrow \nu_\tau$, however in practice there are difficulties with some of the channels. In the case of $\nu_e \rightarrow \nu_\tau$, experimentally it is tricky since both charged leptons (e and τ) exhibit the same topology in the neutrino telescope and the magnitude of the oscillation is low due to the flavor composition of the third mass eigenstate [29]. For the channel $\nu_\tau \rightarrow \nu_\tau$, it is also experimentally extremely difficult due to a lack of statistics. Thus in practice, only $\nu_\mu \rightarrow \nu_\tau$ becomes the most viable option. In fact, the "Oscillation Project with Emulsion-tRacking Apparatus" (OPERA) experiment declared the detection of five tau neutrino candidates with a 5.1σ significance [30] in 2015. There was an update three years later with 5 more with a 6.1σ significance [31].

The detection of tau neutrinos from $\nu_\mu \rightarrow \nu_\tau$ oscillations of atmospheric neutrinos forms another important contribution to probing the unitarity of the flavor mixing matrix. Detectors that aim to measure atmospheric neutrinos are often Cherenkov detectors with a sparse instrumentation of PMTs. Although less densely instrumented than "Direct observation of the nu tau" (DONuT) [32] and OPERA [33], who have previously carried out tau appearance measurements, the gain from the much larger detector volumes justifies their use. In contrast to long baseline experiments, which use artificial neutrino beams, atmospheric oscillation experiments have an indirect detection in mind by utilizing the atmospheric neutrino flux. Since the atmospheric flux in the GeV range contains only ν_e and ν_μ , the presence of ν_τ can only stem from oscillations during the propagation of the neutrino.

However, at GeV energies, it is very challenging to directly identify the tau lepton in such detectors. This is because the tau decays practically immediately ($t_\tau = 2.9 \times 10^{-13}$ s), less than a millimeter away from its point of creation [6]. The oscillation probability depends on L/E. In ANTARES assuming a particle length of the entire Earth (~ 120000 km), the most relevant oscillation maximum is found for vertically up going events with energies between about 20 to 30 GeV (see figure:13). Furthermore, the detection of a tau neutrino is not as straightforward, the tau lepton that decays by producing additional neutrinos plus an electron (BR $\sim 17.8\%$), a muon (BR $\sim 17.4\%$) or hadrons (BR $\sim 64.8\%$)[34]. In these interactions, the tau neutrino scatters off the quark leading to a break up of the nucleon and a hadronic shower (vertex shower). Additionally, a charged tau lepton is created. However, this tau lepton decays within one millimeter or less after its creation. Thus, the different tau-channels cannot be all grouped into the same neutrino topology (see more about the topologies of a neutrino telescope in sec:3.5). The measurement of ν_τ is accomplished, on a statistical basis by finding the excess events on top of the expected distribution from the other flavors without ν_τ . From the branching ratios of the τ decay, the excess is evident in the event distribution of the shower-like signature.

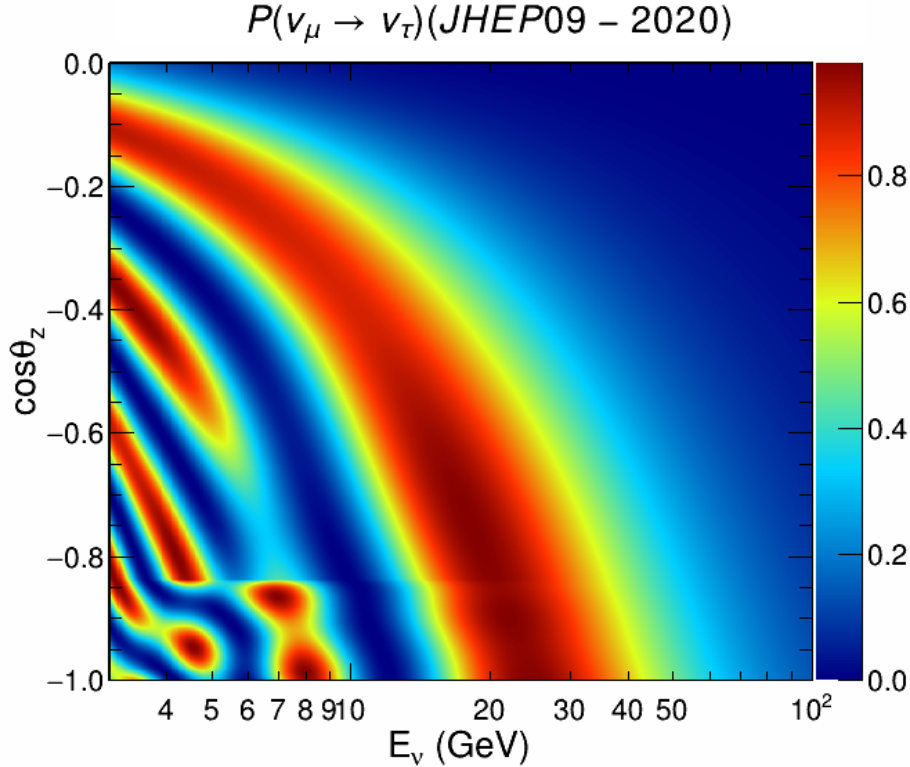


Figure 13: Transition probability from $\nu_\mu \rightarrow \nu_\tau$ as a function of energy in GeV and the cosine of the direction in the case of NMO. Calculated with OscProb Software [35].

Tau appearance is measured as the model independent normalization of the tau neutrino flux (tau normalization, TauNorm, n_τ), which is defined as the ratio between the number of measured and expected tau neutrino events like

$$n_\tau = \frac{N_{\text{meas}}}{N_{\text{exp}}} \quad (32)$$

Theoretically, presuming unitarity of the PMNS and that the cross-section and models stand, the TauNorm should be 1. However, in practice, the tau normalization is a scaling factor that can be accomplished in two ways for CC, or all current (CC+NC). The first approach, measures only tau events through the charged current component (CC-only). On other hand it scales all tau events with the neutral currents included. In these type of measurements there are two results, the value of tau normalization parameter and the exclusion

of the non-tau scenario. The tau normalization has been discussed, with the non-tau scenario being correlated to this parameter. The meaning behind this result is the significance to exclude the tau normalization converging to a value of zero. The non-tau scenario shouldn't be misconstrued with no oscillations, rather than no statistical excess of shower events was detected.

2.2 Tau neutrino cross-section

There are a lot of systematic uncertainties included in a tau appearance measurement. One of the most important uncertainties affecting the interpretation of it is the σ_{ν_τ} cross-section concerning the CC interactions. Due to flavor universality, the cross-section σ_{ν_i} is the same for all flavors $i = e, \mu, \tau$ [36, 37]. This has been verified for the case of NC, although for CC the σ_{ν_e} and σ_{ν_μ} is roughly and some conclusions can be drawn from these [38]. But for the case of the σ_{ν_τ} due to lack of statistics precise, measurements have not been accomplished.

In the GeV range relevant for "Astronomy with a Neutrino Telescope and Abyss environmental RESearch" (ANTARES), the possible interactions are QE, RES and DIS with the latter being the interaction that is the most dominant. The SM provides a theoretical prediction for the relevant cross-sections [4]. In the DIS regime, i.e. high energy section, the cross-section becomes proportional to the neutrino energy, however there are still some uncertainties, which might arise from the contributions of QE and RES. Although in the lower energies, i.e. the energy region where the other interactions become more prevalent, nuclear effects and lepton masses shouldn't be disregarded.

Regarding the measurement of the σ_{ν_τ} , there have been already some results from various experiments such as DONuT [32], OPERA [30] and SK [39]. Theory and measurements in principle agree with each other up to this point taking into account the current assumptions, but more precise constraints from future neutrino experiments might prove otherwise.

2.3 Status of current tau neutrino experiments

In this section, a brief synopsis of the tau neutrino appearance experiments is presented. A brief mention of the long-baseline experiment results has already been done in a previous part 2.1, with the main result being from the OPERA collaboration at $n_\tau = 1.1^{+0.5}_{-0.4}$ with a 6.1σ significance [30]. Apart from the long-baseline experiment, there are also atmospheric experiments that have presented some results such as SK, IceCube/DeepCore and KM3NeT.

In 2018, the SK collaboration [40] released an updated version of their measurement with 14.6 years of data detecting $N_{\nu_\tau^{CC}} = 185$ candidate events resulting in a TauNorm of $n_\tau = 1.47 \pm 0.32$ with 4.6σ significance for non tau scenario [39].

In 2019, the IceCube collaboration published results for atmospheric tau neutrino appearance in the lower-energy sub-array DeepCore. In that paper, the TauNorm had been measured as $0.57^{+0.36}_{-0.3}$ ($0.73^{+0.36}_{-0.24}$) and the case of the non-tau scenario had been rejected with a $2\sigma(3.2\sigma)$ significance for a data-taking period of 2.75 years for the CC(CC+NC) cases [29]. It should be noted that the IceCube is expected to get an extension, the IceCube Upgrade, aiming to improve in sensitivity by adding more densely instrumented lines [41]. According to their first estimates the aim for an event rate of about 300 ν_τ events per year in the first oscillation maximum.

In 2023, the first measurement from the KM3NeT collaboration had been presented using the "Oscillation Research with Cosmics in the Abyss" (ORCA) detector with the configuration of 6 detection lines (ORCA6). The duration of the measurement was 510 days or expressed to effective mass volume 433 kton-years. The corresponding result for TauNorm was $n_\tau = 0.5^{+0.46}_{-0.42}$ ($0.67^{+0.37}_{-0.33}$) for the CC and CC+NC case. As for the non-tau scenario there is the significance hasn't reached the 3σ significance yet, thus there is only a disfavor of $1.2\sigma(2.2\sigma)$ for the CC and CC+NC cases respectively.

Experiment	Livetime	Exclude non-tau scenario	Measurement (n_τ)
OPERA	4 years	6.1σ	$1.1^{+0.5}_{-0.4}$
Super-Kamiokande	14.6 years	4.6σ	1.47 ± 0.32
IceCube/DeepCore (CC)	2.75 years	2.0σ	$0.57^{+0.36}_{-0.3}$
IceCube/DeepCore (CC+NC)	2.75 years	3.2σ	$0.73^{+0.3}_{-0.24}$
KM3NeT/ORCA6 (CC)	1.4 years	1.2σ	$0.5^{+0.46}_{-0.42}$
KM3NeT/ORCA6 (CC+NC)	1.4 years	2.2σ	$0.67^{+0.37}_{-0.33}$

Table 1: Summary of the up-to-date results of the tau appearance measurements from the collaborations.

All the aforementioned results have been gathered in table as an overview before ending the chapter (see Table:1). Furthermore, there should be two more clarifications, the first being that all the results presented here are regarding the mass ordering (NMO), as for IMO the measurement and the exclusion scenario is higher. Secondly, the non-tau scenario is not the same the non-oscillation scenario, but only that $n_\tau = 0$.

3 ANTARES detector

3.1 Science Goals

The ANTARES used to be a novel underwater neutrino telescope located in the Mediterranean Sea 40km off the coast of the French city of Toulon [42] at a depth of 2.5 km. This neutrino telescope was a water-based Cherenkov telescope that detected neutrinos by the emitted Cherenkov radiation from secondary charged particles produced by weak interactions of neutrinos with the medium. ANTARES began operations in May 2008 and continued taking data for a whole of 15 years, until the 14th of February 2022 to be more precise, thus proving the concept of a functional deep sea-water neutrino telescope.

ANTARES's primary goal is to find and analyze high-energy neutrinos that come from astrophysical sources such active galactic nuclei, gamma-ray bursts, and supernovae. These neutrinos are able to shed light on some of the universe's most enigmatic and intense occurrences. Although its primary objective was neutrino astronomy, a multitude of other secondary multidisciplinary scientific objectives were the study of neutrino oscillation [43], dark matter [44] and earth sciences.

3.2 Detection principle

A prerequisite for a neutrino to be detected, the neutrino must interact with matter to produce relativistic secondary particles, which in turn these products will be the "signal" in the detector. For the secondary particles to be registered by the optical sensors, an emission of Cherenkov radiation must be present. A charged relativistic particle that propagates through dielectric medium polarizes the material in its vicinity. A time dependent dipole field is produced along the trajectory of the particle, with the condition that the particle speed exceeds the phase velocity of light in that medium. The asymmetric polarisation and reverting to its initial state results to the emission of Cherenkov radiation. The light that is emitted exhibits a specific angle θ_C in form of a conic angle front. The Cherenkov angle is derived by

$$\cos\theta = \frac{1}{\beta n} \quad (33)$$

Where β is the normalization of the particle speed to the velocity of light and n is the refraction index of its medium. As mentioned before in space the emission is in a form of a cone, but the projection in 2D is a circle.

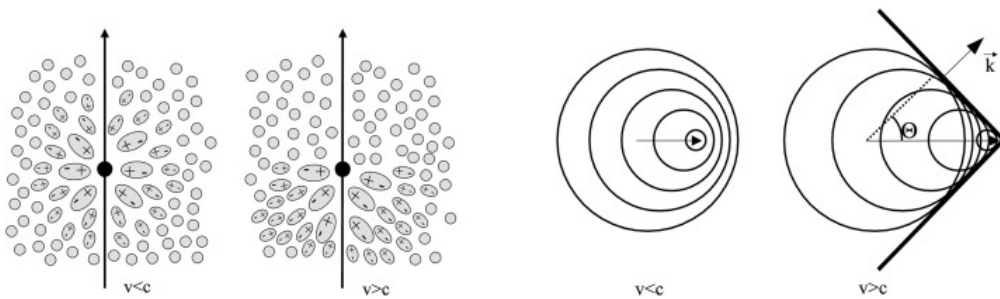


Figure 14: (Left) the polarization, (right) the schematic of the Cherenkov cone depending on the velocity of the particle compared to the speed of light in the medium. Taken from [45]

For the case of water-based neutrino telescopes, the average refraction index in deep-sea water is $n_{\text{water}} = 1.35$. In the paradigm that the secondary particles are ultra-relativistic ($\beta \sim 1$), which is a fair assumption since their kinetic energy is much higher than their mass ($E_{\text{kin}} \gg m_0 c^2$), the light is detected under an angle of $\theta_C = 42^\circ$.

The condition of the Cherenkov threshold, $\beta > 1/n$, can be derived as

$$T_{\text{thresh}} = E_{\text{thresh}} - m_0 c^2 = \left(\frac{1}{\sqrt{1 - n^{-2}}} - 1 \right) m_0 c^2 \quad (34)$$

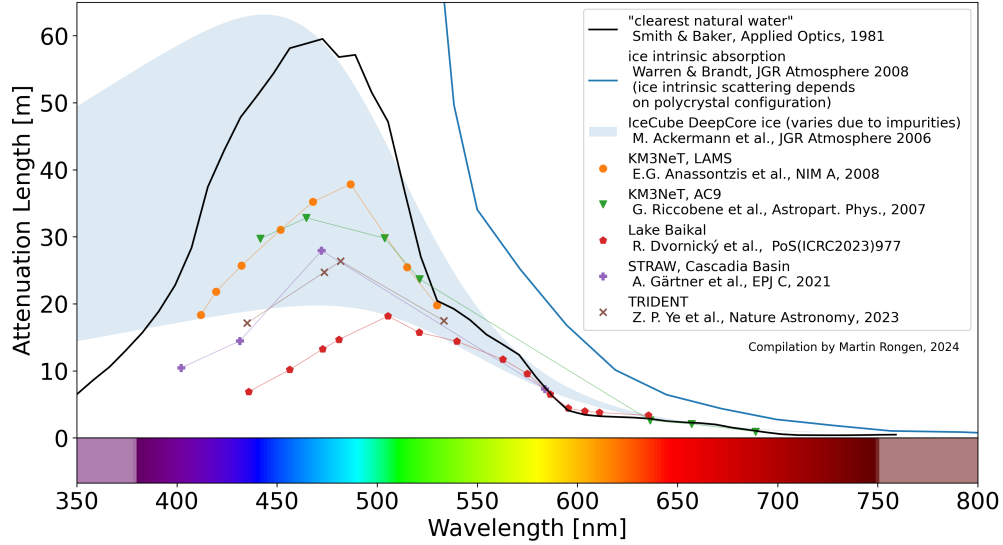


Figure 15: The attenuation length as a function of wavelength for the various sites of neutrino experiments

this being the kinematic energy threshold. Where T_{thresh} is the kinematic energy of the particle, E_{thresh} is the total energy and m_0c^2 being the mass of the particle in units of energy. From (Eq: 34), in sea water the kinematic energy required for the emission of Cherenkov light is $\sim 49^\circ$ [7].

The number of Cherenkov photons follows a behaviour of $dN/d\lambda \propto 1/\lambda^2$. More specifically an ultra-relativistic particle produces a number of Cherenkov photons per unit path length $d\lambda$ in a wavelength between $[\lambda_{\min}, \lambda_{\max}]$ as

$$\frac{dN_C[\lambda_{\min}, \lambda_{\max}]}{d\lambda} = 2\pi\alpha(1 - n^{-2}) \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\min}\lambda_{\max}} \quad (35)$$

with $a \sim 1/137$ being the fine-structure constant

It should be noted that seawater is transparent from 300 nm to 600 nm, otherwise, Cherenkov photons would be detected from the ultraviolet to the visible spectrum. Thus the total number of photons emitted in this wavelength range is $|dN_C/d\lambda| = 340/cm$

From the creation point until detection, the Cherenkov photons under propagation can result in absorption and scattering. The reduction of photons under absorption is exponential and the absorption length l_{abs} is the distance, where the number of photons becomes $1/e$ of the original ones. The maximum length, l_{abs} , is 70 m at a wavelength of $\sim 440\text{nm}$, whereas for the Mediterranean Sea the maximum for l_{abs} is about 60 m [46]. In the case of scattering, dust particles in water contribute to the scattering of light. In the site of ANTARES, the scattering length has been measured to be $l_{\text{scat}} \sim 55\text{m}$ at a wavelength of 470 nm [46]. The result of scattering blurs the Cherenkov signature at long distances.

3.3 Detector layout

The layout of ANTARES is illustrated by Figure 16. As shown in the illustration, the detector consisted of 12 vertical lines and an experimental one (IL 07) with a horizontal distance of about 60 meters. The lines were attached to the seabed and on top of the buoy was placed in order to keep it vertical. Each line was comprised of 25 storeys with a distance of about 14.5 meters from each other. In turn, each storey housed 3 optical modules (OMs) at an angle of 45 degrees downwards looking to up-going events. More details about the storeys and the detector itself will be explained in the next section (look section 3.4). In total, there are 295 OMs containing 885 PMTs for the entire detector, since the last line has only 20 storeys with OMs and the rest using acoustic sensors as part of AMADEUS [47].

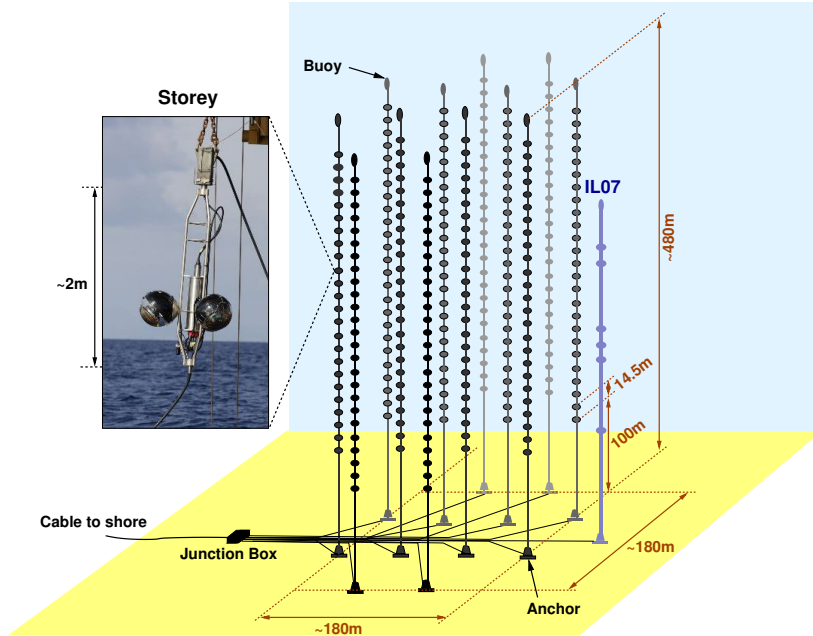


Figure 16: Illustration of the detector layout of ANTARES along with an enhancement of one storey created by Robert Lahmann.

3.4 Detector design

Here, a more detailed explanation of the storey's outline is presented in (Figure: 17). The various elements of the storey can be divided to Local Control Module (LCM), the Light Emitting Diode (LED) optical beacon, hydrophones and the OMs.

The LCM can be referred to the "brain" of the storey, since it houses all the read-out electronics. In this cylindrical box, there is a multitude of various electronics such as the Data Acquisition (DAQ) system, which hosts the operating system for handling all the data, the compass motherboard used for measuring the heading, pitch and roll, providing the orientation of the PMTs and, many other components responsible for communications and power supply of the storey.

The calibration can be divided into time, position and orientation. Pulsed light sources are employed for the apparatus's time calibration. There are two varieties of them: laser beacons and LED beacons [48]. They are dispersed throughout the detector in specific spots. An LED was a point-like light source emitted isotropically, used for the inner-storey and storey-to-storey calibration. The calibration in the last storeys is achieved with LEDs and laser beacons. Furthermore for calibration between lines an intensity of photons is required that can only be achieved through a laser, since the distances are so great that use of LED became inefficient. In contrast to the other method, lasers are only in two central lines. As for the position calibration, a system of acoustic receivers and transceivers had been decided. A transceiver had been anchored in every line of ANTARES and a total amount of five receivers was placed in specific positions same for each line. The location of the receivers along the line has been optimized, thus the upper part of the line has the larger density of the hydrophones, with their location being the 1, 8, 14, 19 and 25 storey.

The OMs is the cornerstone of a neutrino telescope and the main detection sensor. A simple description would be a PMT enclosed in a pressure-resistant glass sphere accompanied by the other components required for the proper function of the detector. The version of ANTARES hosts a 10-inch PMT from Hamamatsu (R7081-20) inside a 17-inch glass sphere [49]. The glass sphere and the gel-filled inside were selected to transfer Cherenkov photons to the sensors with high efficiency, with the transmittance being constant for $\gtrsim 350$ mm.

3.5 Events in ANTARES

In the energy ranges of ANTARES, neutrinos interact with the medium through DIS mainly, although the contribution from QE shouldn't not be ignored either. In the case of NC interactions, an "invisible" neutrino is

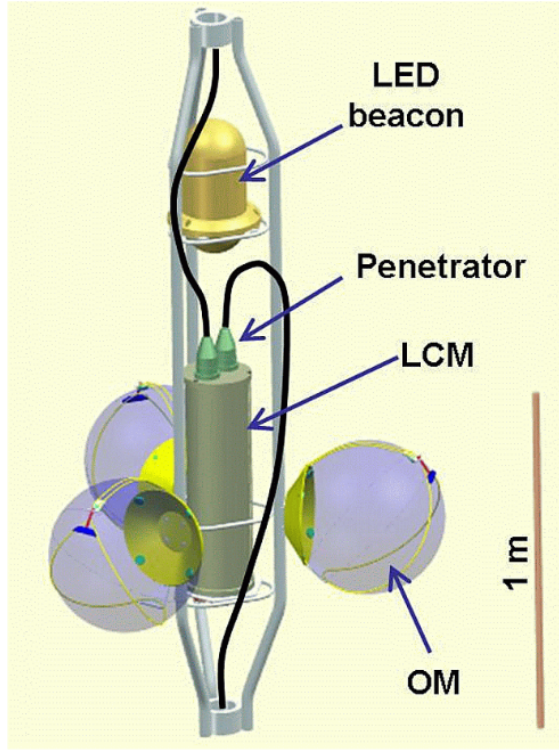


Figure 17: Illustration of the ANTARES storey. Taken from [42]

produced and a hadronic shower, whereas during CC a charged lepton and a hadronic shower are created, both detectable to a neutrino telescope (see sec:1.1.1). The possible event signatures are (1) muon tracks, (2) electromagnetic cascades and (3) hadronic cascades. Although the detector cannot differentiate between the nature of the showers, thus the neutrino topologies can be defined as (1) tracks and (2) showers (see fig:18). "Track-like" events, lose energy along the trajectory of the muon in a cylindrical volume with axis the path of the muon. In the case of "shower-like" events, the energy losses are more distributed in a spherical volume with the "sphericity" being influenced by the energy of the cascade [50]. Classifying neutrino flavours into topologies is more complicated than it seems. The "track-like" encompasses all ν_μ CC events and ν_τ CC events, whose taus decay into a muon. The "shower-like" events include all NC events, the ν_e CC events and the ν_τ channels not mentioned before.

3.6 Background sources

The average noise per PMT is 60-120 kHz with instances in short duration that can reach up to MHz. The three main sources are (a) dark counts, (b) decay of ^{40}K and (c) bioluminescence.

1. Dark counts

The dark count rate is an intrinsic ability of a PMT, which refers to hits registered in the detector without the presence of external light. These photons are produced thermally inside the dynode and can be reduced through fine-tuning. The background rate attributed to dark counts is constant, but considerably lower than the other sources [42].

2. Decay of ^{40}K in sea water

Radioactive ^{40}K exists in seawater in small traces ($\sim 0.04\%$) and is the main background of radioactive sources [7]. The ^{40}K decays under β -decay

$$^{40}\text{K} \rightarrow ^{40}\text{Ca} + e^- + \bar{\nu}_e \quad (89.3\%) \quad (36)$$

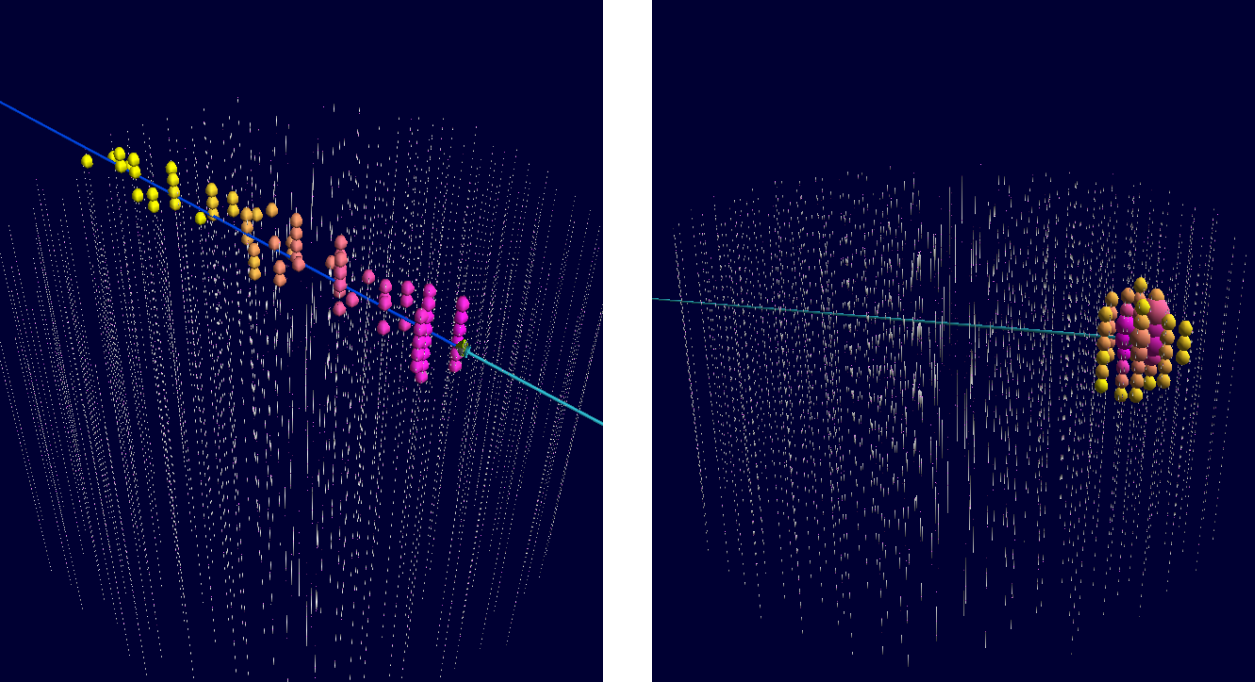
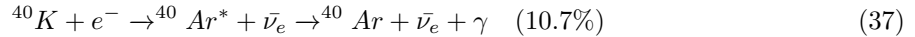


Figure 18: 3D visualization of both neutrino topologies of KM3NeT. (Left) visualization of a "track-like" event and (right) visualization of a "shower-like" event. Taken from [51].

and electron Capture (EC)



For the case of β -decay, electrons with energy above 1.31 MeV meet the criteria for Cherenkov emission. Whereas during an EC, after the de-excitation of Argon a photon with energy 1.46 MeV is produced. That in turn through Compton scattering can transfer its energy to electrons that can produce a Cherenkov cone. This background rate is also constant.

3. Bioluminescence

In the depths of the sea, sunlight cannot reach the detector, thus total darkness occurs. The organisms that inhabit these depths produce light innately from micro-organisms to larger animals. The chemical reaction for light production is due to two main ingredients, luciferin and luciferase with more details about the whole process in [52]. Bioluminescence can be divided into a diffuse emission produced by microscopic organisms and bright bursts from larger organisms with a duration of several seconds. It is usually triggered through turbulences in the lower parts of the OM or direct impact to the impact. Thus there is a correlation between the baseline rate and the prevalence of bursts with the speed of the sea current [53]. Regarding the bursts, the frequency is influenced by seasonal variations and exhibits a peak during spring [54].

4 Monte Carlo event processing

4.1 ANTARES MC production

The full production of Monte Carlo (MC) requires a series of consecutive steps, 3 main ones in total these being (1) the generation of the physics events, (2) the tracking of these particles and (3) the data acquisition processes of the detector. In the particle generation process, muons and neutrinos are produced with information such as the energy, the position and its direction in the proximity of the detector [55]. For the production of the physics events the codes that were used are the following for the case of neutrinos the main source-code is *GENerator of High-Energy Neutrinos (GEHEN)* [56], where interactions of NC and CC from all flavors are possible. The energy ranges of these neutrinos can be from sub-GeV energies up to 10^9 GeV with all main relevant processes considered such as the "Glashow resonance" [57]. It should be noted that also other packages apart from *GEHEN* like *LEPTO*, responsible for calculating DIS at high energies above 10 TeV, where extrapolation of the known models is required. The neutrino energies are sampled from a power-law spectrum that follows the following function ($dN/dE \propto E^{-\gamma}$), where γ is a user-defined parameter. The value for γ is chosen in order to achieve the required level of statistics for the entire energy range. The atmospheric muons being the main physics event, it is really important to accurately approximate its flux, since compared to other experiments no shielding exists. overlaying This lack results in a significant contribution of events in the up-going direction. There are two methods for generating the atmospheric flux, the full simulation approach and the parametric one. In the first one, the atmospheric muon bundles are created from reproducing a whole Extensive Air Showers (EAS) from a primary cosmic ray. For this approach, the *COsmic Ray Simulations for KAscade (CORSIKA)* software [58] is utilized, however it is a very computationally resource-intensive process since it tracks all particles from the primary interaction point to the sea level. A faster alternative is the second one, produced by the *MUon GENerator for neutrino telescopes based on Parametric formulas (MUPAGE)* software [59] and the method used in ANTARES in the latest version of MC v4.2. This method uses some parametric formulae derived from a full simulation of events in accordance with results being carried out in the Gran Sasso Laboratory by the MACRO experiment [60] and extrapolated under the sea. Even though the software depicts the energy and angle distribution of muons as a function of the muon bundle multiplicity at various depths, it lacks in flexibility by changing various input parameters that *CORSIKA* can, since it is simply a parameterization.

From the particle generation, Cherenkov photons are induced from the high-energy leptons passing through water at relativistic speed (assuming the Cherenkov criteria in sec:3.2). A dedicated software package was developed by the ANTARES collaboration called *KM3*, that simulated the photon production, the propagation of the particles and light in the water medium and an assessment of the probability of recording a hit on the PMTs. These three steps can be divided into the different programs used, the first one called *gen*, the second one *km3mc* and the last one *hit*. During *gen*, the "photon fields" are created from Cherenkov photons through the passage of leptons. Each photon is monitored from its interaction to the moment it is absorbed by the detector or leaves it. The second part of *KM3*, *hit*, uses simulations from *GEANT* [61] to calculate how likely is for a photon to be registered in the PMT given the effective area of an OM. The probability can be calculated to the ratio of the detected to generated photons multiplied by the cross-sectional area, or it can be equated to the projected area of the PMT multiplied by the quantum efficiency. The third part, *km3mc*, deals with the propagation of the light emission and the particle through a designated area, called CAN. The software is able to treat light emissions from both muons and electromagnetic showers, these being either derived from muon or electron cascades. In the case of muons, if the energy losses are less than 300 MeV/m they are classified as continuous, whereas for greater are discrete/stochastic [6].

As a last step, all the hits registered in the PMTs are parsed through the filtering algorithms of the ANTARES detector in order to assimilate the format and the environmental conditions of "real data". The hits don't only include physical events, but also the optical background such as bioluminescence and radioactive decay of ^{40}K (see more about in sec:3.6). As the hits are recorded and transformed into digital signals and transmitted to the shore without any modifications, the so-called "all-data-to-shore" approach [62]. Once in the on-shore data centres, the filtering algorithms are applied depending on the detector conditions. The software package used to simulate the trigger protocols is *TriggerEfficiency* [42].

4.2 AntDSTs)

The AntDST files represent the final stage of files of the MC processing chain and it is an internal storing format developed by the ANTARES collaboration [63]. More specifically, it is Rapid Object-Oriented Technology (ROOT) [64] based file format which can be used with all common data analysis frameworks: BBFit, CalReal, AAFit and SeaTray [65, 66, 67] and it stores the output of the event reconstruction for data and simulated events. The AntDST were designed to contain high-level variables needed for physics analysis and a fair amount of low-level data to facilitate the development of new data selection cuts and to debug the reconstruction.

The information stored inside the AntDST can be divided into Run Info & Quality parameters, MC truth parameters and reconstructed parameters. In the "Data Quality Tree", all the necessary run information is saved such as the run id, run duration, frame index and run quality criteria such as sea current, base rate, burst fraction, etc. For the "MC Tree", the parameters for identifying each event, kinematics, the position and direction of the particle. Of course additional information regarding the weights and other parameters associated to the weighting process are also saved there. Regarding the tree of the reconstructed parameters, for each algorithm, the position and direction of the particle along with its kinematics is stored. Lastly, variables regarding the quality of the reconstruction are also available, these quantify how well an event was reconstructed by that algorithm. The run quality is usually used in the selection criteria, since there needs to be confidence that the events are reconstructed properly.

The reconstruction algorithms are AAFit, BBFit and GridFit, which are mostly focused on "track-like" signature and Showeraafit, Showertantra and Showerdusj with a preference for "shower-like" topologies. It should be noted that due to the recent nature of the development of NNFit [68], this reconstruction has not been included in the latest official MC production, version 4.2. This version includes the events for the entire lifetime of the detector. This version includes the events for the entire lifetime of the detector. The files that are produced contain events from a single interaction type and are divided for neutrinos according to their energy spectrum and neutrino flavour, whereas for muons the entire spectrum is included in one. For the neutrinos, there is an extra separation according to particle and anti-particles. The events are split also according to every individual run of the detector, thus for every run the MC production is split into 29 isolated files, more importantly, each filename reveals the events contained inside.

The identifiers of the filename that contain neutrinos can be shown further below:

- particle or anti-particle distinction, with a specifier of a for anti-particles
- neutrino flavor, with names nue, numu and nutau for ν_e , ν_μ and ν_τ accordingly
- energy bracket, for event with true energy from 5 GeV until 20 TeV are contained in bracket a and from 20 TeV until 6 PeV in bracket b
- interaction type, with identifiers CC and NC
- run id, unique number identifying each run

So if someone calculates all combinations the neutrino files for every run equal to 28 individual files, 4 for electron neutrino, 4 for muon neutrinos and 6 for tau neutrinos flavors for every energy bracket. For the case of ANTARES, the tau neutrino files are split into 3 files, which comedirectly from the neutrino generator, for the case of CC interaction there is a further split in the case the tau lepton is decayed into particles that have a shower or track topology (see more in subsection:1.1.2)

It should be noted, for the case of the muon files, since all the muons are simulated with minimal weighting afterwards there are instances where for runs with longer duration due to the computational restrictions the files where split to two or three files. So an example of a ROOT file containing muons, could have a filename of *MC_065882_mupage_reco.root*, whereas in the case of longer duration runs it could look like *MC_075882_mupage_1_reco.root*, where the number 1 indicates that this file is the first of the subset of neutrinos. As an example the file named, *MC_065882_anumu_a_CC_reco.root*, the events contained can be revealed by following the guide above. So the file contains events generated are muon anti-neutrinos that interact with the detector through CC interaction with energy below 20 TeV for the run with the specific run id 65882.

4.3 Processing pipeline

The main analysis tool in the thesis for the computation of the sensitivities is *Swim*[69], which is a software framework internally developed from KM3NeT collaboration for the neutrino analyses oscillations. The input format is very specific requiring only ROOT, without any proprietary libraries. Given that to access an AntDST, you have to accomplish that inside the SeaTray singularity container, the need for a conversion is urgent. In that regard a pipeline has been created to process and adapt the AntDSTs, to a file that can be imported in the analysis pipeline. A flowchart of the various steps taken for the "Swim" ROOT files to be created (see Fig:19). There is a colouring scheme in order to visualize better the various components of the pipeline. By yellow the input files to the pipeline are indicated, with light red the processing steps, with light blue a logic gate and the final files are with light green. The number marker in each pipeline step and the arrows can reveal the sequence of the pipeline, where if the files contain neutrinos the atmospheric oscillation weights are calculated before the merging. The muons files skip this step since the branches that will be computed during this step have physical meaning only for neutrino events. From the previous section, Sec:4.2, the various types of files were discussed and with the naming convention that has been in place, the name can reveal the contents of the files. This was exploited to process only files directly relevant to the analysis, to that end neutrinos from all flavors were chosen in the relevant energy case and muon files. A small reminder, generated neutrinos with energy up to 20 TeV are stored in files with the naming convention *_a_* included in their name.

In total close to three hundred thousand files were processed by the pipeline with varied sizes. The decision to process only the necessary files, led to a 48% reduction of the number of files processed. This doesn't translate too the same level of computing power, because as a rule the of thump the muon files are more than an order of magnitude larger in size. While monitoring the jobs through the logs, it was revealed that a number of files were either missing or corrupted. This loss of files was during the transfer to the university cluster. As soon as it was realized the issue was quickly corrected. All the necessary steps were taken by members of the group to rectify the situation and however there are still a few missing files that are due to issues during the latest version of MC production. Once this was identified and the issue was referred to the person in charge, but until this moment no known steps have been taken. This however this doesn't have an overall effect given that it is less than 1% of the overall files.

- Total files: 565224
- Neutrino files in energy range a: 268887
- Muon files: 27943

The first step in the pipeline, see Fig:19, deals with the conversion of the summary files (AntDST), to a ROOT file. As mentioned, the reason for this step is to get rid the dependency from a singularity container, *Seatray*. For files to be accessed in *Swim*, the file must be able to be imported with standard ROOT libraries, since the structure that the file is expected to have is a **TTree**. Even though the AntDSTs have also ROOT format, their inner structure is more complex with custom developed libraries. In this part then, the file is accessed inside its singularity container and the necessary branches are read-out to a new file with a **TTree** structure. These branches that are extracted contain, quality and run information from the *DQ* named tree, such as environmental conditions, *RunID*, run duration, etc. In fact from this Tree, the rates of the PMTs is recorded and can be visualized for the duration of the detector, see in Figure:20.

In fact, many outcomes can be derived with the analysis of these environmental conditions in order to find, which runs are have gone through suitable environmental conditions for good quality of data taking. From Figure:20, the seasonal variations and their effect in the event rates is displayed as mentioned in the background sources (see Sec:3.6). This behaviour can be further exemplified, by gathering all the event rates as a function of the data taking month. The figures, Fig:20 and Fig:21, clearly indicate these seasonal variations and during an analysis with data events an investigation for which run selection should be included in the final dataset, but since given the exploratory nature of this analysis and the fact that a large data sample is desired for a first investigation it was decided not to include further analysis at this point of time.

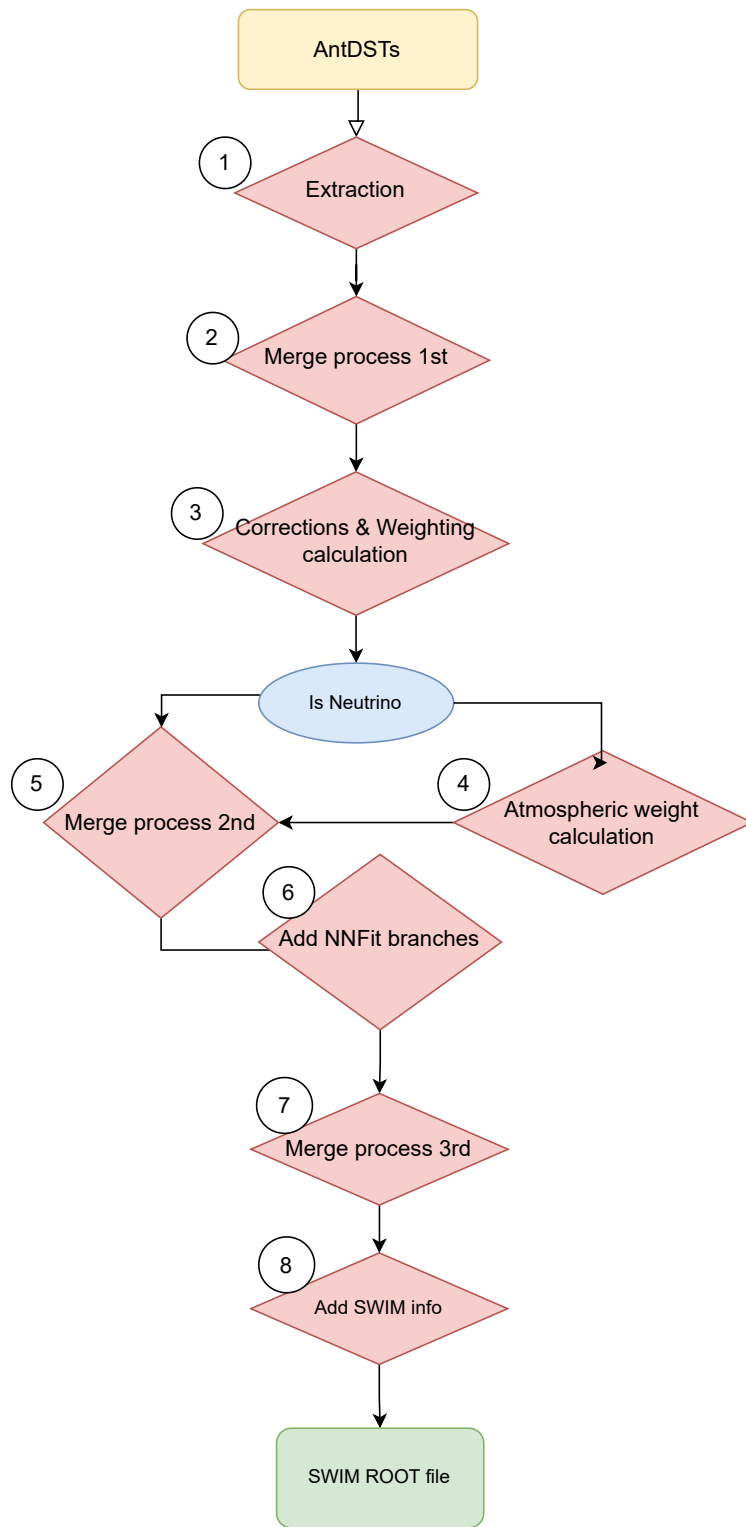


Figure 19: Flowchart for the processing pipeline. The colour scheme for this flowchart is the following, light yellow for the input file, light red for the processing steps, light blue for the logic gate and light green for the final file.

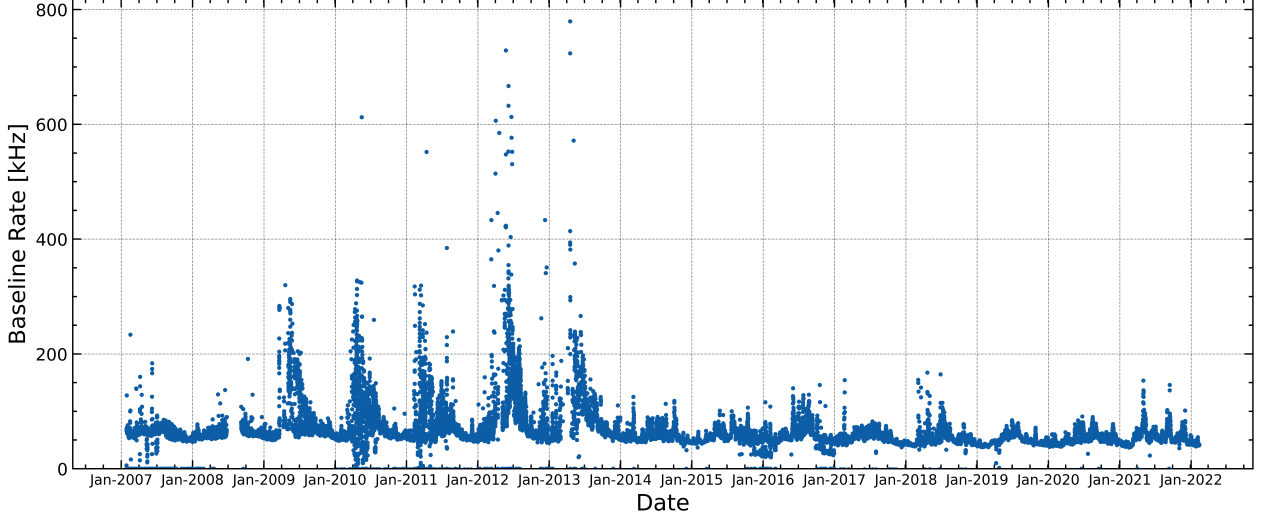


Figure 20: Event rate in ANTARES as a function of time, for the entire operation of the detector.

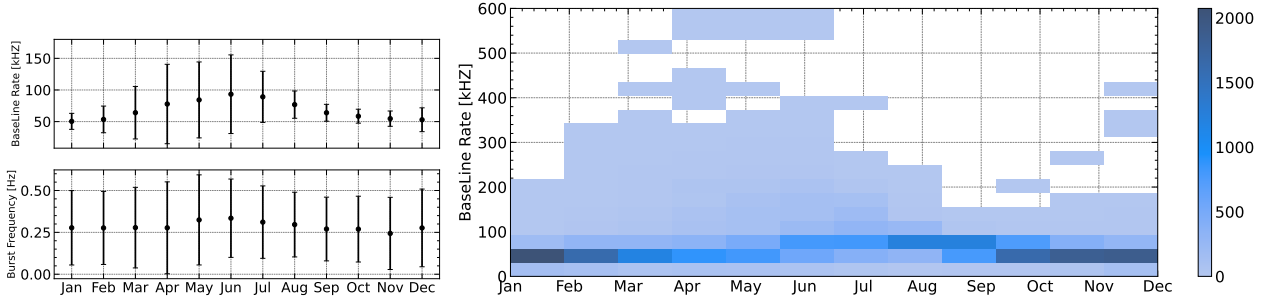


Figure 21: (Left) the fraction of burst-like events and (right) the detector rates as a function of the data-taking month.

Except for the data quality parameters, two other **Trees** are accessed and their information is extracted. The first one is the **MCEvent**, where all the true information for the generated events is stored. These parameters are the true energy, direction (θ, ϕ angles) and the vertex point of the interaction (x, y, z). Additionally, the weights that the events were generated with are stored, if the event was a neutrino. The generation weight, w_2 , is derived from the energy spectrum that the generation of neutrinos was based upon, see more details in Sec:4.1. The physical weight is also available, which corresponds to the multiplication of the generation weight to the **non** oscillated flux of the neutrino ($w_3 = w_2 \cdot \Phi_\nu$). Depending on the model used here three similar weights are provided named w_3 , w_{honda} and w_{bartol} . Derived from the Honda and Bartol atmospheric models. In the case of the tau neutrinos, since under the assumption of no oscillation there isn't an atmospheric flux, the weights are set to zero. The n_{gen} is a value related to the weighting that represents the number of generated events. The flavor/type of event and its interaction type is stored in this tree too. The conventions for these variables are the following

- type
 - $\mu = -13$
 - $\nu_e = 12$ and $\bar{\nu}_e = -12$
 - $\nu_e = 14$ and $\bar{\nu}_e = -14$
 - $\nu_e = 16$ and $\bar{\nu}_e = -16$
- interaction type
 - CC = 1
 - NC = 0
 - $CC^{\text{mu}} = 2$
 - $CC^{\text{shower}} = 3$

As seen above, all events are assigned a numerical value. For anti-neutrinos the assigned value is the absolute negative value of its particle counterpart and for muons the value given is minus thirteen. The

interaction type and its variations had been discussed, with one point that should be paid attention is the split in the CC interaction channel in the ν_τ . The last parameters that were extracted from this tree, are the event identification parameters. These are the frame index and the event id, with the first corresponding to which frame during the run the event belongs and the latter is an identification for an event to distinguish events that belong in the same frame. Frame is a time period of 104.858 ms, where the run is divided into. The frame index indicates to a specific frame of a run and the event id to the specific event in a frame.

The last type of parameters that we extract from the summary is the reconstructed information. Each reconstruction algorithm has a separate **TTree**, which includes the following reconstructed parameters the reconstructed energy, direction (θ, ϕ), reconstructed vertex (x,y,z). A parameter is quantifying the quality of the reconstruction is present, Lastly, there are two parameters denoting to the number of OMs and Detection Unitss (DUs) used for the reconstruction. The reconstruction quality parameters are specific to the reconstruction algorithm, with some having a parameters evaluating the overall quality of the reconstruction and there is also the case, where an estimation of the individual error of each parameter is stored. The reconstruction algorithms extracted where the following, (1) AAFit, (2) BBFit, (3) GridFit, (4) Showerdusj and (5) Showertantra.

The selection criterion used during the reconstruction process is for the event to be triggered and reconstructed by at least one of the aforementioned reconstruction algorithms. For each one of the algorithms, a boolean flag is introduced indicating if the criterion is satisfied or not by this reconstruction or not. This was introduced to easily, assess the efficiency of each method. There was however another reason for its adoption, with the current implementation, there were instances of wrongly assigned values in some branches. If the event passed the selection, even if one algorithm didn't reconstruct it, it was wrongly assigned the reconstructed parameters of another method This issue is corrected in the following step of the pipeline. The selection criteria that are used are very loose in order to maximize the detector volume before applying any sort of cut. There is a practical issue in applying a cut after the extraction, since this part of the pipeline is quite resource intensive in the university cluster. With some approximations between job types on average it lasts around 15 hours, using 150 cores in parallel. So it was pursued to limit the submission of this step only in the case of major changes, because this step comprises half of the duration of the processing pipeline and 70 % of the computational resources.

On the second step of the pipeline from Figure:19, an in-between merging is carried out in order to lessen the workload in the computer cluster. The process merges the files to that have the same event type, interaction type and last number of the run that they belong to. In total, there are fifteen combinations of event type and interaction type, those being $\nu_e^{CC}, \nu_e^{NC}, \bar{\nu}_e^{CC}, \bar{\nu}_e^{NC}, \nu_\mu^{CC}, \nu_\mu^{NC}, \bar{\nu}_\mu^{CC}, \bar{\nu}_\mu^{NC}, \nu_\tau^{CCmu}, \nu_\tau^{CCshower}, \nu_\tau^{NC}, \bar{\nu}_\tau^{CCmu}, \bar{\nu}_\tau^{CCshower}, \bar{\nu}_\tau^{NC}$ and μ . So for example, the files *MC_065882_anumu_a_CC_reco.root* and *MC_595282_anumu_a_CC_reco.root* would be part of the same merged file. From the ten files per type, make in total hundred fifty merged files. This not only makes monitoring a bit easier to handle, but it is also much healthier for the file system. Since a lot of open and exit operations slow the speed of the file system there were issues encountered, when there were a lot of users performing resource-intensive operations in parallel. There is also a physical explanation behind it, since when someone renders an analysis with data, only 10% is used to avoid including biases. This is easily achieved here, by using the zero-ending number runs for example.

From Figure:19, the third step of the pipeline is divided into two processes, the first being corrections to the reconstruction branches due to the issue encountered in the first step, and the second part that applies some corrections to the weight branches. As mentioned the reconstruction flag branches produced in the extractor step (step (1)) are crucial for this process. Depending on the value of the boolean flag, the branches associated with it are kept or not. If the value is False, then this indicates that the event was not reconstructed by that algorithm and the reconstructed parameters stored are erased. In the second part of the analysis, some corrections are applied into the weight branches, that take into account the efficiency of the detector through the entire 15 year duration of the experiment. For example, there was an overestimation in the MC production for runs below 30412 that took place into the early years of the experiment and thus a 20% reduction was applied to these files to approximate the conditions of the detector at that time. This correction was required, since the software that was used for the production can only take one file that simulates the photo-multipliers. Furthermore, after analysis from members of the collaboration data versus MC comparison, was carried out and additional correction were devised that are dependent on the year of the measurement. The corrections applied to the weight are gathered here

- run < 30412: if the run is before the 30412nd a 20% reduction need to be applied

- data/MC corrections: corrections to approximate closer the data. They are dependent on the event type, the interaction type and the year of the run that is part of

Additional branches are also calculated such as the non-oscillation weights and the *weight_one_year*, which approximates the muons in units of $years^{-1}$. This branch is used for Swim[69] to approximate the muon background. It should be stated, that the treatment of the weight corrections is standardised for the collaboration and the steps are well documented in the wiki.

The the weights that don't account for neutrino oscillations, can be calculated through

$$w_{non-osc} = \frac{w_3}{ngen} * run_duration \quad (38)$$

with w_3 being the physical weights, $ngen$ the number of generated events and *run_duration* the duration of each run. Although the weights (w_2, w_3) are already included in the branches of the AntDSTs.

The *weight_one_year* branch on the other hand is calculated as

$$weight_one_year = w_muon * \frac{365.25}{lifetime} \quad (39)$$

with *w_muon* being the muon weights and the ratio in days from the total number of days in year to the days that the detector was active. The *w_muon* is calculated as a "flat-based" factor according to the specifier of the muon file. More specifically, the factor can take the values of 3, 9 or 4.5. This is accomplished through a look-up table that stores a value for each run from zero to three and depending on the value, a weight is assigned. This parameter was calculated in step (1) of the pipeline. To sum up, the weight assignments depends on the number

- weight=3: if the number is zero or three
- weight=9: if the number is one
- weight=4.5: if the number is two

With the conclusion of step (3) of the pipeline from Figure:19, there are two possible paths depending on the content of the event files. If the files contain μ then the files are directed to initiate merging. In the case that there are neutrinos, then they must go through step number (4) process before any merging. For an initial study without the use of *Swim*, the weights that take into account neutrino oscillations need to be calculated for an expectation of the neutrino events. A specific library is required in this step named OscProb [35], developed by members of the KM3NeT collaboration, dedicated to the calculation of the neutrino oscillation probabilities for various scenarios, SM and Beyond Standard Model (BSM). Also, the neutrino fluxes from the Honda model [70] are used with some members having converted the tables of the paper into a graph saved inside a ROOT file.

The neutrino oscillated weights are calculated as

$$w_{osc} = \frac{w_2}{ngen} * (\Phi_{\nu_e} P_{\nu_e \rightarrow \nu_i} + \Phi_{\nu_\mu} P_{\nu_\mu \rightarrow \nu_i}) \quad run_duration \quad (40)$$

with w_2 being the generation weight, the weights produced from the generation step of the MC production that follow a power-law for more detail (see subsection:4.1), $ngen$ the number of generated events in the production file, *run_duration* the duration of the run in units of *years*, Φ the flux of the neutrino flavor according to the Honda model and $P_{\nu_x \rightarrow \nu_y}$ the oscillation probability from neutrino with flavor x to y, where x and y can be e, μ, τ . Depending on the neutrino flavors of x and y the oscillation probabilities are named either survival or transition probabilities. The survival probability assumes that the initial and the final flavor is the same, whereas for the transition probability the initial and final flavor must be different. The product $\Phi_{\nu_e} P_{\nu_e \rightarrow \nu_i}$ is one of the components of the oscillated atmospheric flux of the neutrino flavor i, which represents the electron neutrinos that oscillated to the neutrino flavor i and likewise for the muon neutrinos. These two components can describe the entire atmospheric flux, since tau neutrinos are not present in during the production of the atmospheric neutrinos for the standard case.

From an initial glance, neutrino oscillation weights appear to be straightforward but behind the flux and the oscillation probability lies a multi-dimensional problem, since both values are dependent on the energy and

the direction of the particle. The flux that is used derives from the measurement of the Frejus site without the mountain over the detector with the all-year average in order to account for the seasonal variations at a solar minimum. Furthermore, it was averaged in the azimuth direction thus the only free parameters in the flux equation are only the energy and the cosine of the zenith direction. The fluxes from the graphs that were presented in the latest paper from Honda (HKKM2014) [70], are turned into histograms with binning, 81 bins for the energy reaching up to 10 TeV and the zenith direction its values are averaged and binned into 20 bins. These histograms are then imported in the analysis and the flux value that is used is interpolated from the logarithmic version of the histogram due to its faster interpolation. The neutrino fluxes used during the calculation of Equation:17, can be seen in Figure:22.

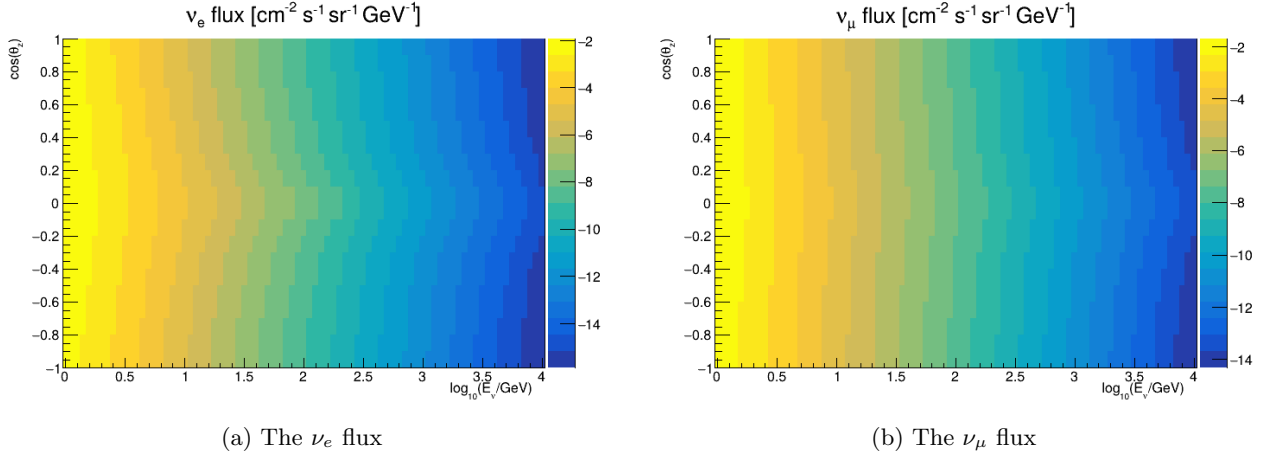


Figure 22: The atmospheric neutrino fluxes according to the Honda model. The logarithm of the flux is plotted as a function of the logarithm of energy to the cosine zenith direction.

As mentioned above, the oscillation probabilities are calculated with the OscProb software [35]. With this software, a symmetry spherical model of the Earth is simulated with several layers of various levels of constant densities and different properties, the Preliminary Reference Earth Model (PREM) model [71]. This is one of the classes provided by this library, which provides some models with different numbers of layers and properties, but also gives the possibility to import your own defined model. As a function of the cosine of the zenith angle with respect to the detector, the class can then generate path sequences through the various earth layers. The path sequences will account for the fact that some layers are above the detector, and the detector can be placed at any radius inside the model. By default, the site of the detector is situated at the bottom of the ocean and some already pre-compiled PREM models are already provided. Another part, is the definition of the PMNS matrix and the Hamiltonian of the problem. Since our problem is the standard scenario of a 3-flavor oscillation, several optimizations have been performed such as avoiding the computation of null terms [72], in a specific class called *PMNS_FAST*. For the case of other scenarios the situation is more complicated and the computational time longer. But dedicated classes for the main BSM such as Non-Standard Interactions (NSI), Lorentz Invariance Violations (LIV), Decoherence, etc.

Before the calculation of the oscillation probabilities, there are some prerequisites that need to be declared before starting the calculations such as the declaration of the PMNS matrix, which can be accomplished manually and with some predefined ones. The assumption for the oscillation parameters was for NMO and taken from PDG [6]. The PREM that is used is the default option that numbers 44 layers and will hold the neutrino paths that are calculated. After the definition of the necessary prerequisites, the approach of this analysis is to iterate through every event and calculate the probability of an electron neutrino that would transition to the flavor of the given event and for the muon neutrino the approach is the same. A speed-up is applied in the calculation where events that wouldn't pass through the Earth would be omitted from the calculation. By not parsing through the Earth, it is implied that the events are down-going and are coming to the detector from the horizon, assuming that the path length of the neutrino then would be around 15 km, it can safely be approximated that the survival probability of the event will be close to 1, thus no oscillation is likely. Therefore, for the case of the "down-going" events the oscillation and non-oscillation weights are

the same. On the other case, the path sequence of the neutrino for a given cosine zenith is filled from the upper-most layer to the innermost layer, until the outer layer again reaches the detector in a vector. From the neutrino path, then the oscillation probability is derived according to the neutrino cosine zenith and energy. After the calculation of the two oscillation probabilities, the oscillation weights are calculated according to Equation: 40.

The oscillograms of an initial ν_μ are calculated for the three other flavors as a function of energy and cosine zenith, by following the aforementioned steps with OscProb, with these being now presented below in Figure:23. All the necessary oscillograms required to for the calculation of the oscillation weights are gathered in Appendix:A.

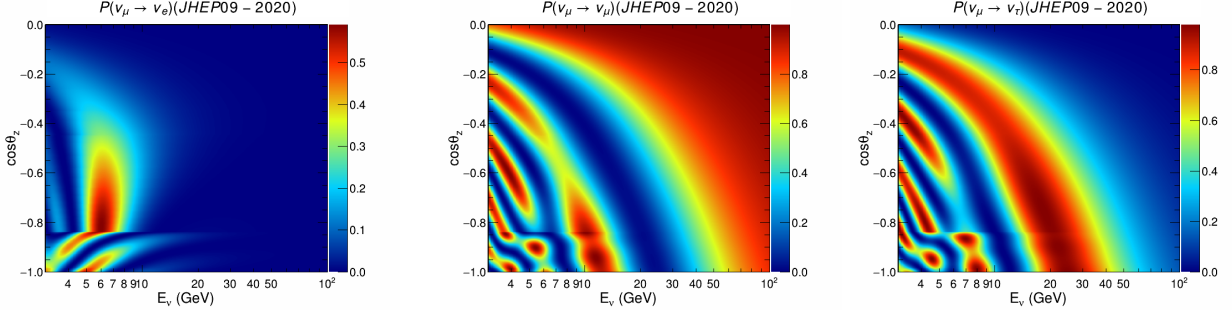


Figure 23: The two transition, $\nu_\mu \rightarrow \nu_e$ and $\nu_\mu \rightarrow \nu_\tau$, and the survival probabilities for the ν_μ according to the PMNS parameters from [73]. Calculated with OscProb.

Step (5) of the pipeline, commences the second round of merging. The purpose of this step is to merge the files in a manner that would be more efficient for step (6) of the pipeline, see Figure:19. The main reason is that the next process can be quite RAM intensive, in the order of more than 100 GeV per job. Before starting the explanation for the implementation of step (5), a small comment regarding the next step. The main task of step (6), is to include the reconstructed information from the newly developed method, NNFit. For a productive merging, the way the the H5 files that NNFit is stored, needs to be addressed. There are four grouping in their file, those being "Showers", "NuMuCC", "Muons" and "Taus". The "Showers" class includes events that would have a shower-like topology in the detector, such as ν_e CC, $\nu_e + \nu_\mu$ for NC and the corresponding antiparticles. The "NuMuCC" class includes only ν_μ and $\bar{\nu}_\mu$ under CC interaction. "Muons" include μ and "Taus" ν_τ and $\bar{\nu}_\tau$ for both CC and NC interactions. These four classes are offered as an individual run-by-run file or merged. Since the files in the pipeline are already merged, then using the merged files is the more productive way forward. How, should the merging be done? Should NNFit files become like the ones in the pipeline or the other way around? Both methods were investigated and the latter was less resource intensive. So the files had to take the approach of NNFit.

The H5 files combine a range of runs, with the common denominator not being the number of runs, but the size of the resulting file. This approach was chosen for the loading time to be roughly the same, since the loading time usually takes a significant role in the processes. This however approach isn't as straightforward to follow. The naming convention though of the merged files makes the task straightforward. Each merged file, contains the class that is part of, the id of the first run that contains and the id of the last. For example, the file *Taus_0164_85547_86343.hdf5* contains ν_τ and $\bar{\nu}_\tau$ events from the run ranges from 85547 up to 86343. The first number is a counter that indicates that is the 164th of this type. So with a script, the ranges of each file can be directly gathered. With the trickier part done, the files of the pipelines are split into their corresponding ranges and then merged into their file types that share the same class in NNFit. The full lists containing the run ranges of each NNFit merged file, gathered from the script are all included in Appendix:C. This method might be more complex than trying to merge a fully merged file, but the memory request would become "unrealistic" requesting multiple CPUs, without the computing need for all of them. This would result in a highly inefficient job and we want to avoid this at all costs.

In step (6), the merge between the files of the pipeline with the NNFit H5 files is commenced. The resulting file combines every available official reconstruction algorithm provided by the collaboration. To be precise, our file "absorbs" the NNFit information by creating new branches allocated to the "new" algorithm.

The methodology to accomplish that is to identify each event reconstructed by the algorithm and assign its reconstructed parameters. To distinguish an event three parameters are required run id, frame index and event index. The meaning of the frame and the event index have already been discussed in this section. During the process, both files are imported into data-frames and the merge is done based on the aforementioned three parameters,

- run id: indicates to which run the event is included
- frame index: shows which frame the event is part of
- event id: specifies the event in a frame

According to these three parameters, the NNFit is combined in our data frame. The instance of no match, is an indication that the event couldn't be reconstructed by this algorithm. For this step, a contact was made with the developer of NNFit and through discussions, it was identified that there is a mismatch in the files processed by both parties. The files that the developer was working with are *i3*, these files include both high and low level information, such as hit information. This is necessary for every reconstruction. However, the AntDSTs files are also created during the same process as the *i3*. This disparity in the number of files, was also verified on irods, a computing server that the production files are safely stored and access to it requires special admission. This verification was carried out by my supervisor and again as before it was forwarded to the person responsible.

A last merging, step (7), is executed to just one ROOT file. During this step a cut is applied, with the various types of selection criteria being discussed in the next section, Sec:4.4. The final step of the pipeline from Figure:19, correlates directly with *Swim* so in case this software is not used, the pipeline can end in the previous step with all the necessary information in a single file. There are two tasks completed in this step, firstly is the copy of the true information branches ($E, \cos\theta, y$) into new ones that would represent the case of perfect reconstruction. These branches will be colloquially be known as "MC-truth" for the next chapter of this thesis. The only reason for these additional branches is due to complications that arose during the calculation of the sensitivities assuming for "perfect" reconstruction of both energy and direction. These issues arose during the process of building of the detector response matrix, when the information from the true branches was used as part of the reconstruction during the configuration of the classes JSON file, see more details regarding *Swim* JSON configuration in section:5.1.1. So in order to calculate sensitivities with a "perfect" reconstruction new branches must be created to avoid this request conflict. The other process is necessary for any use of *Swim* and this is the inclusion of the detector dimensions. The format for this information needs to be a three-dimensional vector and more specifically a (**T3Vector**) and a simple script was compiled to check whether the file already has this vector, otherwise it would include the dimensions of ANTARES. Since the detector is simulated as a cylinder the parameters of this *AntaresCANDimensions* are the following

- $z_{\min} = -271.42$
- $z_{\max} = 357.94$
- $r = 279.45$

with z_{\min} being the minimum height, z_{\max} the maximum height and r the radius of the cylinder. All these values here are in units of meters. And with that completed, there is one file that can be parsed without issues directly to *Swim*.

4.4 Reconstruction algorithms

Since the start of ANTARES, a list of reconstructions has been developed by focusing on different energies and neutrino topologies. Since most of the studies performed and the main science objectives of the experiment target higher-energy neutrinos, they are optimized for energies in the TeV range. As a result, the dedicated shower-reconstruction algorithms *Showerdusj* and *Showertantra* do not cover the energy range relevant to neutrino oscillations. For the case of *Showerdusj*, it is even coded to estimate energies from the starting point of 100 GeV. In fact from the entire list of reconstruction algorithms that are contained in the AntDSTs only

BBFit, was focused on low-energy track-like events with an approximation for shower-like events via "BBFit Bright point". NNFit is a new generation Deep Learning (DL) reconstruction, very recently released in the collaboration, that focuses on low-energy reconstruction for both topologies. AAFit is a reconstruction, that is optimized to reconstruct high-energy tracks. Due to its focus, it's been widely used in main ANTARES papers. The reason it has been introduced in this analysis, despite the different target group of events, is its high efficiency. Allowing a wider sample was a compelling argument during the initial investigation of potential reconstructions.

For the performance of the following reconstructions, the energy and direction are investigated, wherever it is possible. For instance, in the current version of AntDST an energy estimator wasn't provided for the BBFit algorithm. Whereas for AAFit and NNFit, two estimators are investigated. Furthermore, the events are grouped into two classes, tracks and showers, according to the two their event topologies. For showers the events that are included are ν_e CC and ν_τ CC with a shower-like topology of the daughter particle after the decay of the lepton and for all ν NC, whereas for track the neutrinos classified are ν_μ CC and ν_τ CC with a track-like topology. Since, for the entire dataset there isn't a classifier provided, the analysis is carried out under the assumption of perfect classification. The three cut selections that are used in this study are presented, the (1) being a "purely" selection cut on the true energy and direction in order to find commonalities in these low-energy up-going events, the (2) and the (3) one focuses on the reconstruction algorithm selected after the initial analysis. The naming convention for the first cut is called, "**low-energy up-going**" events with the following cut selection on purely true direction, with those being

- $\cos \theta_{\text{true}} < 0$
- $E_{\text{true}} < 100 \text{ GeV}$
- $E_{\text{true}} > 10 \text{ GeV}$

where E_{true} is the true energy of the neutrino and with "low-energy" we mean from now on along this study as below 100 GeV and $\cos \theta_{\text{true}}$ the cosine of the zenith direction. The choice to pick only events that are up-going, derives the lack of a ν_τ flux from the down-going direction. This expected effect was discussed in Section:1.2.1. The energy cut is from the expectation that the biggest contribution to the tau appearance measurement would come from low-energy events, reminded Section:2.1. The decision to cut event below 10 GeV is due to the energy threshold of the detector. The two other cuts, (2) and (3), are presented in the discussion of the reconstruction algorithm that they are based upon (early indication Section:4.4.3).

4.4.1 AAFit

AAFit [74], a maximum likelihood fit, is applied on events with Cherenkov light detected by optical modules distributed on several detector lines (Multi-Line events (ML) events). It is considered the main track reconstruction algorithm of the collaboration. It is divided into various steps, each of which builds the input required for the success of the next step. This reconstruction chain's initial step preselects the hits associated with a single event, making sure that only L0 hits that satisfy the condition,

$$|\Delta t| \leq \frac{d}{u_g} + 100 \quad (41)$$

with Δt the time difference between the hit and the largest pulse amplitude in the sample, d the distance between OMs and u_g the group velocity of light in water. In this sense, it is anticipated that all signal hits will be retained, while background hits will cause the algorithm to become less sensitive. The initial reconstruction phase is called a linear prefit; it just needs the previously mentioned preselected hits as input [75]. It acts as a preliminary approximation of the track parameters, which are then fed into the succeeding fitting steps. The parameters fitted in order to approximate the track are 5 in total, the position \vec{q} in a given time t_0 and its direction \vec{d} , which can be calculated as a function of the local azimuth and zenith angles ($\vec{d} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$). In order for \vec{y} , the positions of the OMs, to come from straightforward translations of the starting point \vec{q} along the direction \vec{d} , the linear prefit first assumes that all recorded impact positions occur along the muon track, as $\vec{y} = \vec{q} + ct\vec{d}$. The next fitting step, which consists of maximizing probability via a reduced PDF that is a function of only the hit time residuals, is based on the

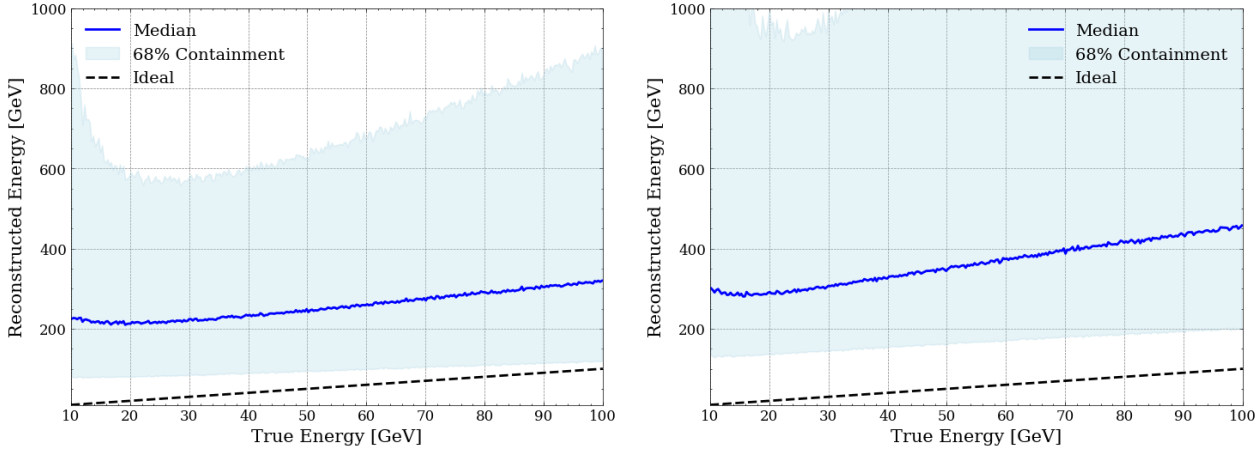


Figure 24: (Left) distribution of shower-like ν events and (right) the distribution of track-like ν events after applying the "low-energy true" cut for *AAFIt ANN Energy Estimator* with a dashed line representing the "perfect" reconstruction.

M-estimator results. The likelihood is maximized with regard to the track parameters $\vec{\Theta}$ and is the product of the probability to witnessing each individual hit with the specified time residual[76].

$$\mathcal{L} = \prod_i^{\text{Nhits}} P(t_i, t_i^{\text{exp}}, \vec{\Theta}) \quad (42)$$

where t_i^{exp} , is the expected time of the photons emitted by the muon to the PMT with the calculation of the time explained in detail in [74].

The Particle Distribution Function (PDF) function itself is acquired by the use of muons in the detector volume that produce a Cherenkov light yield, which is captured as a series of histograms that display the time distribution of the hits for various angles and separations from the simulated track. Following that, the histograms are fitted using continuous differentiable operations without considering the context. The resulting functions, which take into account the scattering of the Cherenkov photons by water molecules, peak at zero and have asymmetric tails for large time residuals.

Lastly, after the first fit, *AAFIt* repeats the M-Estimator and the previous step a total of 8 times, slightly modifying the initial direction estimation provided by the prefit. The solution with the best performing maximum likelihood of the 8 iterations, defines the initial point of the last fit step, which is again a likelihood maximisation. The quality parameters from the algorithm are β , the angular error estimator, and Λ , the likelihood over the number of degrees of freedom from the fit, with these being

$$\beta = \sqrt{\sigma_\theta^2 + \sin^2 \theta \sigma_\phi^2} \quad (43)$$

$$\Lambda = \frac{\log \mathcal{L}_{\text{max}}}{NDF} + 0.1(N_{\text{comp}} - 1) \quad (44)$$

The initial investigation includes, the event distribution with only the "low_energy_true" cut without any quality cuts are presented.

Without any quality cuts in the Λ and β , it is clear from both figures, Figure:24 and Figure:25, that energy estimation for low-energy events is a very complicated task. The fact that the algorithm is optimized for higher energies complicates the process even more leading to an energy overestimation. Most of the reconstructed events by *AAFIt* are higher-energies, thus the PDFs of the reconstruction favour high energy estimation.

From Figure 26, it is evident that "track-like" events even without any quality cuts in the reconstruction, has a good performance in the range close to -1 on their direction reconstruction. On the other hand, this is not the case for "shower-like" events, which are the main events of interest for us, in fact, the median values are almost constant for the entire up-going range. Thus given the initial challenges in the reconstruction

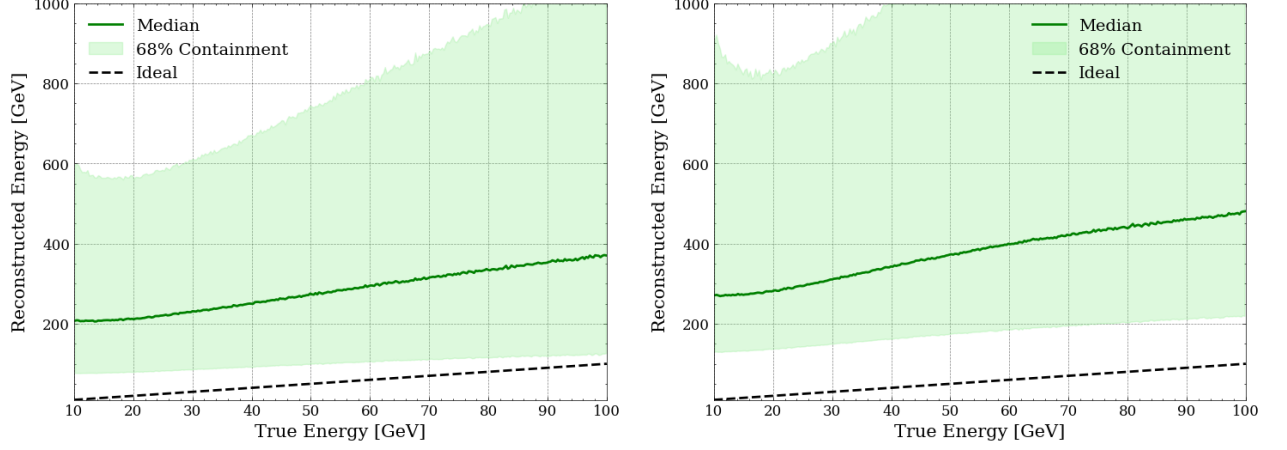


Figure 25: (Left) distribution of shower-like ν events and (right) the distribution of track-like ν events after applying the "low-energy true" cut for *AAFit dEdX Energy Estimator* with a dashed line representing the "perfect" reconstruction.

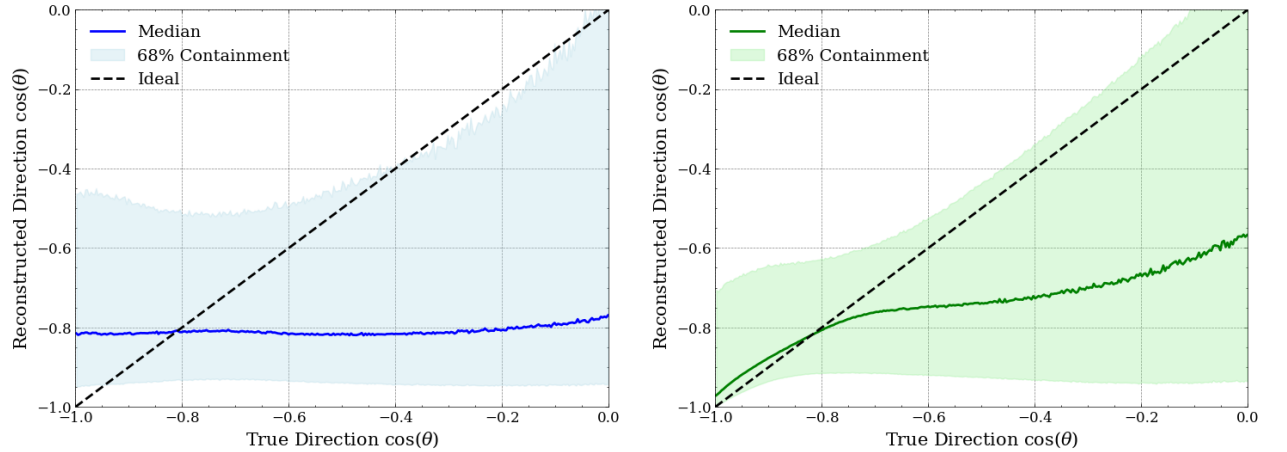


Figure 26: (Left) direction distribution of shower-like ν events and (right) the distribution of track-like ν events.

of both topologies, mainly shower-like events, in both energy and direction reconstruction, a quality cut is applied to check if there are some notable improvements. From the main paper of ANTARES for the measurement of the atmospheric neutrino fluxes [43], in the pre-selection criteria the Λ quality parameter is cut in $\Lambda > -5.7$, though it was meant for higher-energy fluxes and it is a bit more restrictive to for this study, therefore a more relaxed cut of $\Lambda > -6.2$. For events with lower lambda, the lack of gain in the sensitivities from the addition of these events, wouldn't justify it.

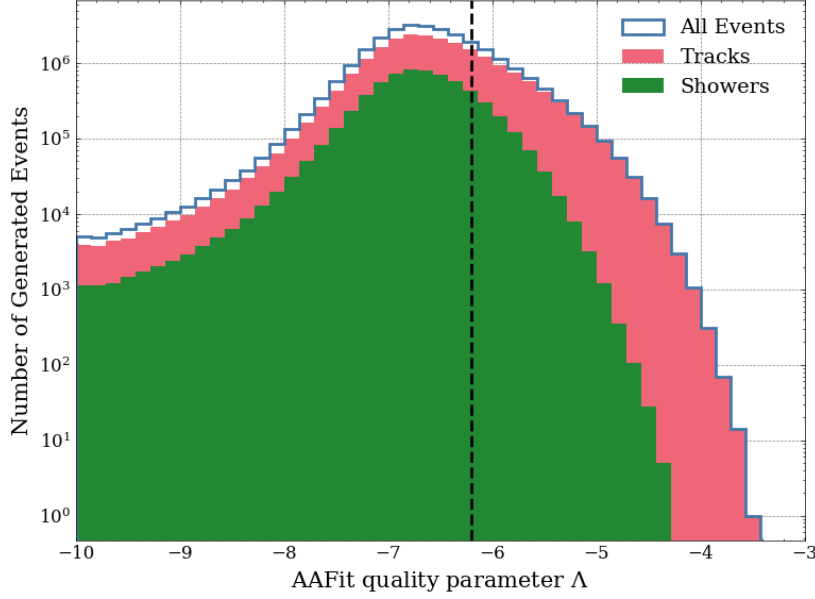


Figure 27: Distribution of the AAFit quality parameter Λ , for both topologies.

Event Topology	low-energy up-going(%)	AAFit Quality Cut (%)	NNFit up-going(%)
showers	65.98	9.01	7.62
tracks	62.62	16.04	8.57

Table 2: AAFit algorithm efficiency for low-energy up-going, well reconstructed by AAFit and reconstructed as up-going by both NNFit and AAFit.

Table 2 reveals the efficiency of AAFit for some selection cuts indicated on the header. From the initial "low-energy up-going" sample, AAFit exhibits the highest efficiency in reconstructing events. But with a "loose" quality cut, only 13% shower and 25% track remain. For this drastic reduction in the event sample, it was decided to search for alternative options. When imposing for the events to be reconstructed as up-going by both AAFit and NNFit, still a bit less than 10% remains. So in the case of introducing cuts from NNFit, it is still possible to use this algorithm as a possibility.

4.4.2 BBFit

The BBfit algorithm has been developed as a robust reconstruction method, which produces reliable results without precise positioning calibration [65]. The BBfit concept is based on the fact that both a detector line and a muon track can be considered as a straight lines in space. Apart from the special case that these two lines are parallel (as in the case of an exactly vertical track) one can define on the muon track a point of closest approach with respect to the detector line. Only hits due to direct Cherenkov photons are selected for the fit. Tails due to scattering and electromagnetic activity are ignored in the fit. The geometry of the three optical modules on a storey is ignored in BBfit. A storey is considered as space point centered on the detector line.

The procedure of the hit treatment, always starts from the list of snapshot hits and in case these are absent then triggered hits are considered. As a first step, time and charge calibration is applied. The calibrated hit times are given in ns with respect to the start of the run and the hit amplitudes are given in photoelectrons. In order to avoid large numbers during the fitting procedure the hit times are reset to time differences with respect to the first triggered hit. It should be noted that since the storey is "counted" as only one optical module, only one hit is used per storey.

The list of chosen hits is looked over before any fits are started. No fit attempt is done if there are fewer than five hits in this list. A single line fit method is initiated if every selected hit is on a single detector line. When the hits are dispersed across multiple detector lines, a multi-line fit process is used. A particle track is thought of as a straight line in space; phenomena such as multiple scattering and other potential deviations from the particle's straight path are taken into account. The code uses the assumption that the particle travels in vacuum at the speed of light (0.3 m/ns). Each and every space-time point on the track can be parameterized as

$$\vec{p}(t) = \vec{q}(t_0) + c(t - t_0)\vec{u} \quad (45)$$

At time t_0 , the particle travels in direction \vec{v} after passing through point \vec{q} . Redefining t_0 allows for the tracking of \vec{q} along the track. As a result, a total of five parameters define the track: two angles to define \vec{v} and three values to fix \vec{q} for a specific duration. We now examine a detector line that is precisely vertical along the z-axis, denoted as $(0, 0, z)$. We can determine the point of closest approach between the particle track and the detector line based just on geometrical factors. Along the z-axis, the closest approach point is

$$z_c = \frac{q_z - u_z(\vec{q} \cdot \vec{u})}{1 - u_z^2} \quad (46)$$

through which the particle passes at time

$$t_c = t_0 + \frac{q_z u_z - (\vec{q} \cdot \vec{u})}{c(1 - u_z^2)} \quad (47)$$

at a distance

$$d_c = \sqrt{p_x^2(t_c) + p_y^2(t_c) + (p_z(t_c) - z_c)^2} \quad (48)$$

A track and a single detector line are arranged in a way that is invariant to z-axis rotations. It is convenient to reparametrize the track in terms of z_c , t_c , d_c , and v_z (the latter being nothing more than the zenith angle's cosine). This indicates that just 4 parameters remain to completely specify the track. The information needed to construct a fitting function is

1. the Cherenkov photon's arrival time (t_γ) at the detector line position $(0, 0, z)$
2. the photon's associated travel route (d_γ)
3. the photon's inclination with respect to the detector line ($\cos \theta_\gamma$)

The four track parameters listed above can be used to derive all three values:

$$d_\gamma(z) = \frac{n}{\sqrt{n^2 - 1}} \sqrt{d_c^2 + (z - z_c)^2(1 - u_z^2)} \quad (49)$$

$$t_\gamma(z) = (t_c - t_0) + \frac{1}{c} \left((z - z_c)u_z + \frac{n^2 - 1}{n} f_\gamma(z) \right) \quad (50)$$

$$\cos \theta_\gamma(z) = (1 - u_z^2) \frac{z - z_c}{d_\gamma(z)} + \frac{u_z}{n} \quad (51)$$

with n being the refraction index and more specifically, it has been set to a value of 1.38.

Except for the method to fit the hits in a track, there is the option to reconstruct it as a "bright-point", shower-like. In contrast to the track, the bright point assumes a point-like source that emits a single light flash at a given moment (t_0). The light that is emitted, tends to propagate isotropically. The approximation of the

bright model, could interpret the light originating from hadronic and electromagnetic showers, presuming that the volume of the detector is much larger than the shower itself. The bright point can be described by four parameters, three for the position \vec{q} and the time t_0 . Similar to how particle tracks characterize the point of closest approach, it is easy to understand that for a bright point, $z_c = q_z$, $t_c = t_0$ and

$$d_c = \sqrt{q_x^2 + q_y^2} \quad (52)$$

In this case, the number of parameters is reduced by one, as the bright point is fully determined with respect to a single detector line at $(0, 0, z)$ by the three parameters z_c , t_c , and d_c . Analogously to a particle track, the photon arrival time t_γ , travel length d_γ , and angle with respect to a given arrival point z along the detector line can thus be calculated.

$$d_\gamma(z) = \sqrt{d_c^2 + (x - q_z)^2} \quad (53)$$

$$t_\gamma(z) = t_0 + \frac{n}{c} d_\gamma \quad (54)$$

$$\cos \theta_\gamma = \frac{(z - q_z)}{d_\gamma} \quad (55)$$

During the fitting process, apart from the signature assumption of a particle track or a "bright point", an additional fitting distinction for Single-Line events (SL) or ML is performed. For the fit, the χ^2 functions are defined and minimized for all four possibilities. Depending on the minimization process, the event is classified as single or multi-line. If the event is multi-line, the azimuth angle of the event is reconstructed.

methodology	single line	multi-line
particle track	4	5
bright point	3	4

Table 3: Number of minimum hits required for a BBFit fit.

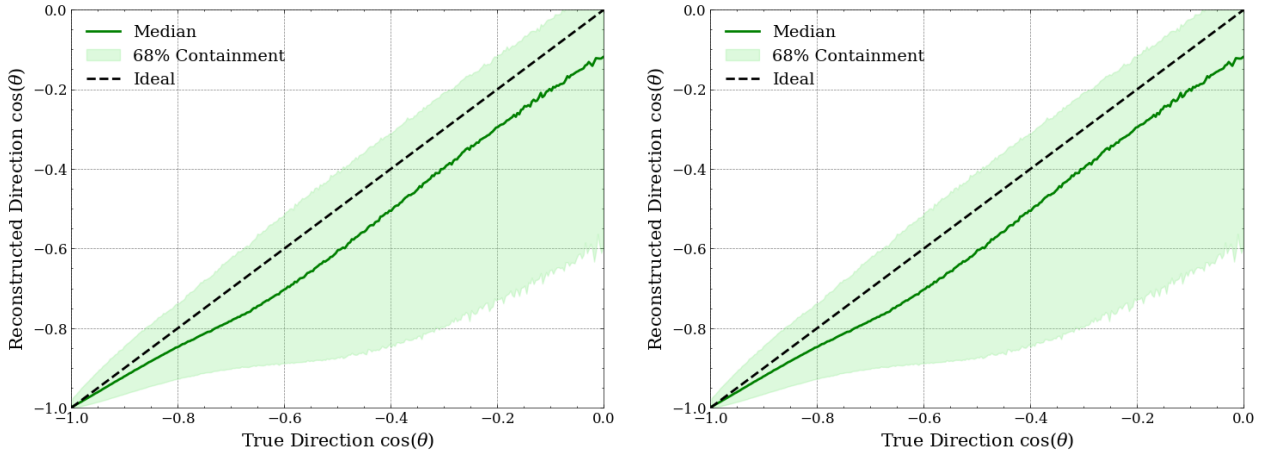


Figure 28: (Left) Direction distributions using the BBFit "particle track" approximation for reconstructing track-like events and (right) the "bright point" approximation for shower-like event reconstruction.

From figures 28 and 29, it is clear that BBFit has a better performance in both topologies compared to AAFit. Concerning tracks, the events are mostly concentrated along the identity, though no notable difference is seen, when applying the minimum events in each bin threshold. Only the approximations for the intended topologies are presented, since the differences are minimal in the direction reconstruction. Some slight

differences were identified in the reconstruction of the vertex or closest point depending on the approximation. Regarding the shower-like events, there is a definite improvement compared to AAFit, since AAFit has a non-insignificant contribution of events that are reconstructed as up going. When comparing algorithms

Event Topology	low-energy up-going(%)	NNFit Reconstructed(%)
showers	28.41	24.54
tracks	34.92	27.19

Table 4: BBFit algorithm efficiency for low-energy up-going, and reconstructed as up-going by both NNFit and BBFit.

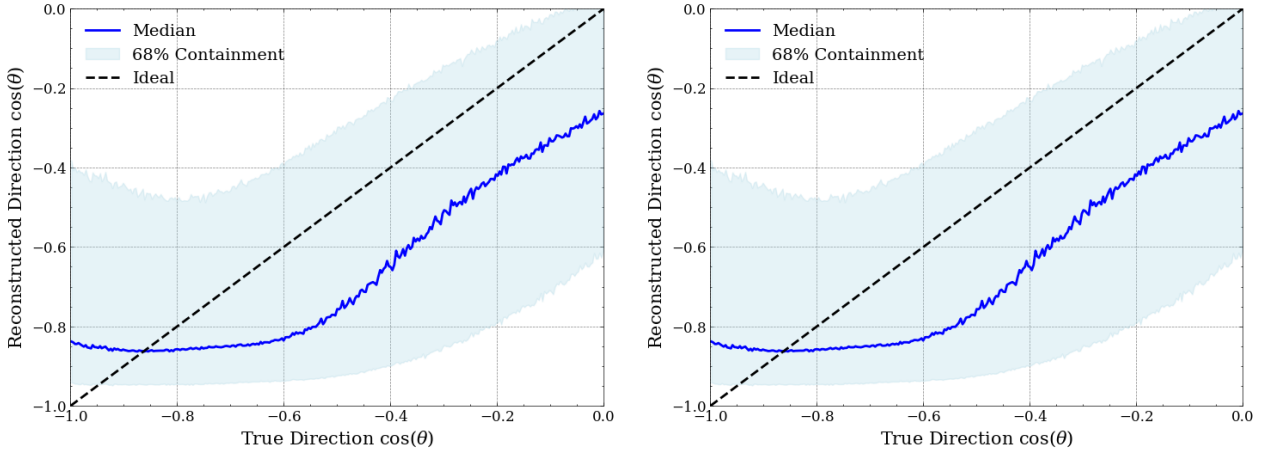


Figure 29: (Left) Direction distributions using the BBFit "particle track" approximation for reconstructing shower-like events and (right) the "bright point" approximation for shower-like event reconstruction.

However, from figures 28 and 29, no significant difference between the two reconstruction methods is detected in this low-energy sample.

4.4.3 NNFit

NNFit corresponds to the collection of Machine Learning algorithms, recently developed by the Valencia group from the ANTARES Collaboration [68]. It reconstructs neutrino properties of so-called SL events, these being events selected according to the BBfit criteria. NNFit algorithms are supervised methods and they combine Deep Convolutional Networks (DCNs) and dimensionality reduction approaches. NNFit is divided into two groups that have been trained separately: *NNFitTrack* and *NNFitShower*. The first has been trained with CC neutrino-(anti)muon Monte Carlo simulations (ν_μ^{CC}), while the second has been trained with both currents of neutrino-(anti)electron neutrinos and with NC neutrino-(anti)muon simulations (ν_e^{CC} , ν_e^{NC} , ν_μ^{NC}). Thus, *NNFitTrack* is optimized for track-like events and *NNFitShower* is optimized for shower-like events.

Mixture Density Network (MDN) together with DCN is the most popular machine learning algorithm used in NNFit. While MDNs enable the prediction of the uncertainty of the reconstructed parameters, DCNs enable the extraction of the spatial correlation of the input data. Principal Component Analysis (PCA) is another technique used to reduce the dimensionality of the data. With the least amount of information lost, this technique finds a smaller set of features in a lower-dimensional subspace that accurately represents the original data. The initial variables that may have been correlated are changed into a new set of principal components that are linearly uncorrelated.

The parameters reconstructed by NNFit are the following:

- NNFitTrack

1. Neutrino direction in terms of the zenith and azimuth angles, θ and ϕ respectively
2. Closest point of the track to the detector line in terms of horizontal distance (R_c) and vertical point (Z_c)
3. Energy of the generated muon

- NNFitShower

1. Neutrino direction in terms of the zenith and azimuth angles, θ and ϕ respectively
2. Interaction vertex point in terms of horizontal distance (R_v) and vertical point (Z_v)
3. Energy of the neutrino

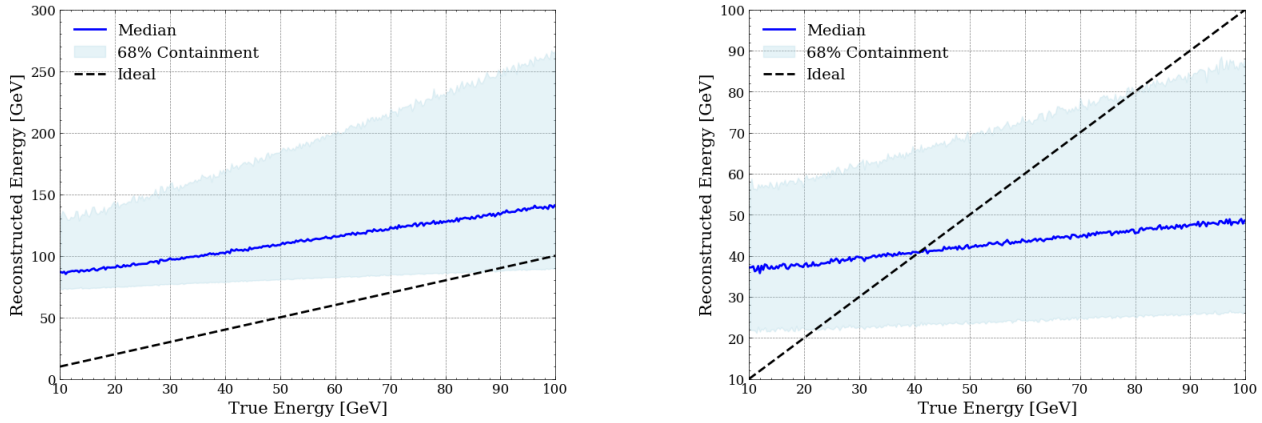


Figure 30: The energy reconstruction distributions for "shower-like" ν events, with (left) being the *NNFitShower* and (right) the *NNFitTrack* energy estimators.

For all reconstructed parameters an error estimation is provided, which can and will be used later to apply some quality cuts.

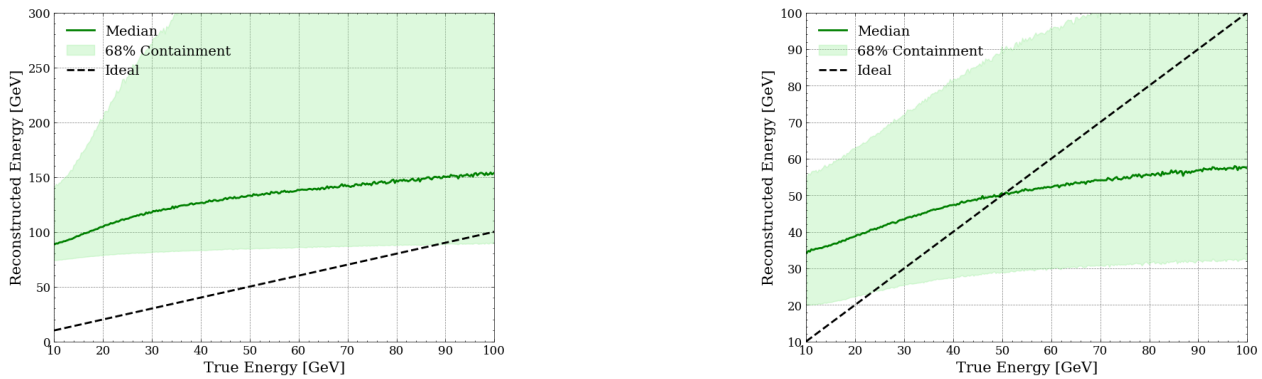


Figure 31: The energy reconstruction distributions for "track-like" ν events, with (left) being the *NNFitShower* and (right) the *NNFitTrack* energy estimators.

Figures 31 and 30, the difficulties in the performance of energy reconstruction are visible. Especially for the case of NNFitShower, the training that was performed was such that there is a plateau in the minimum reconstructed energy which is around 70 GeV. The fact that NNFitShower was optimized through its training

for above 100 GeV events, is something that be accounted for. A possible retraining could improve the energy estimation without significant changes. In the ranges above 80 GeV, NNFitShower in fact slightly underestimates the event's energy. As for NNFitTrack during training, no behaviour similar to NNFitShower is seen, but in the energy window that ν_τ appearance probability (reminder Figure 13), there is a widespread distribution of the events. So from both energy estimators, there are a lot of weaknesses.

The case of the direction resolution is more promising, similarly both models are presented for both topologies as before.

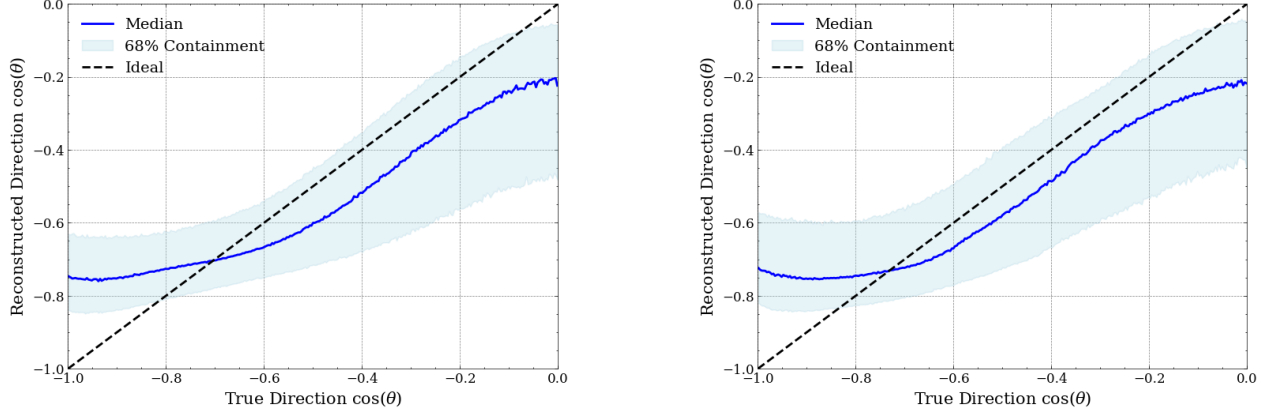


Figure 32: The direction reconstruction distributions for "shower-like" ν events, with the (left) being the NNFitShower and on the (right) the NNFitTrack direction estimators lie.

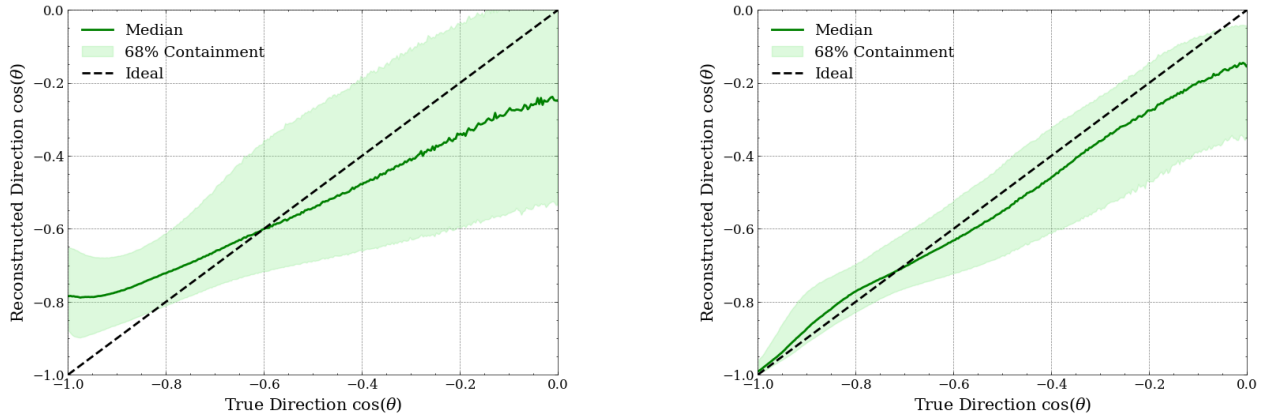


Figure 33: The direction reconstruction distributions for "track-like" ν events, with the (left) being the NNFitShower and on the (right) the NNFitTrack direction estimators lie.

From both Figures 32 and 33 an early indication is raised that, assuming for "perfect" classification, thus depending on the topology of the event and the designed model used, both models are the best-performing reconstructions in the direction reconstruction task. The improvement in the "shower-like" class is especially significant. However, it should be noted that, since NNFit is based on reconstructing only SL BBFit events. Only close to 19.5% of the ν_τ events remain, with slightly higher efficiencies for the other neutrino flavours, for ν_e around 32.1% and for ν_μ close to 31.3 %. Which is well below the efficiencies of the other reconstructions. This low efficiency, increases the complexity of the cut selection, given the already small detector volume assuming that the event must be reconstructed by NNFit. Therefore, it was chosen not to apply direct quality cuts in the reconstructed variables, but rather on their errors., the variables chosen to apply quality cuts are, the error in the position of the vertex (σ_R and σ_Z) for NNFitShower and the error in the closest point to the centre of the detector during the particle track (σ_R and σ_Z) and an error in the zenith direction. Close attention should be paid that the direction angles, are in the ANTARES coordinate system and are

directed towards the direction of the neutrino, not towards where it comes from, which is the strategy for the other algorithms. But quality cut will use only the error there isn't any need to apply some trigonometric corrections, though have already been done the the zenith branch. Hence the naming convention zenith is used, instead in the error estimation in the original theta is used (σ_θ).

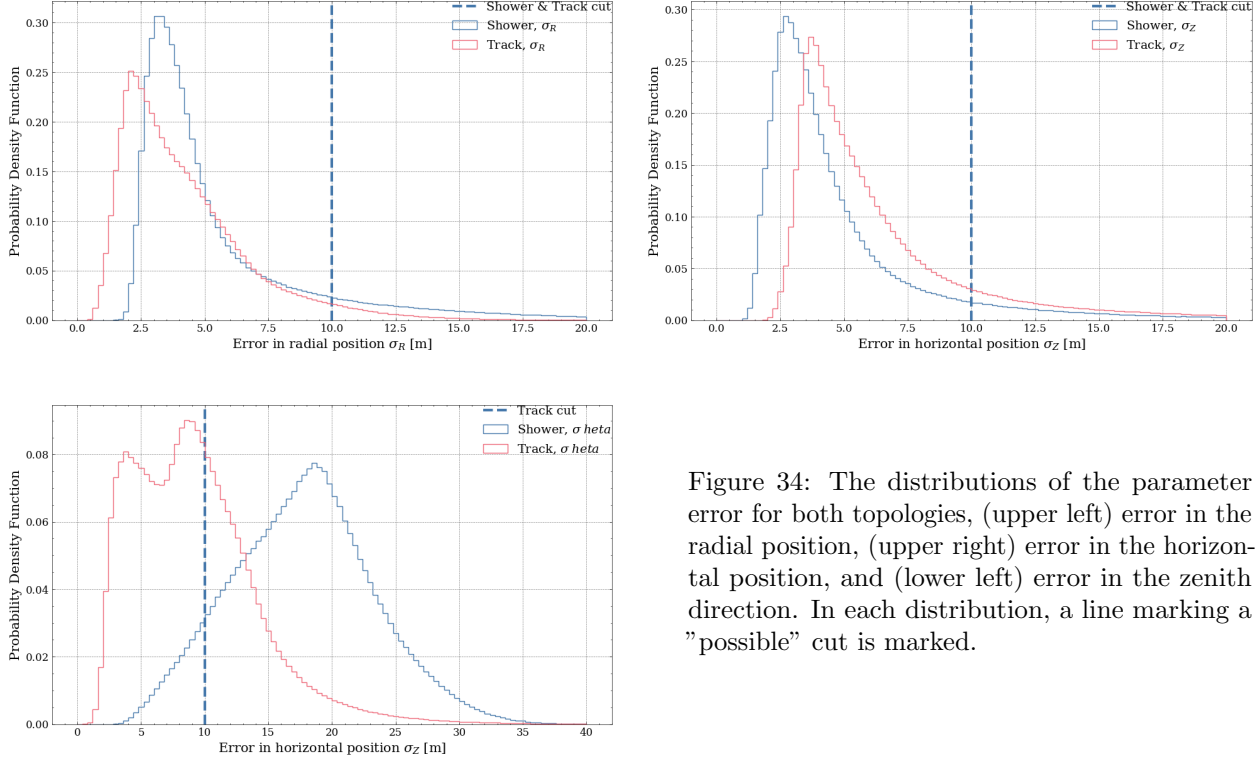


Figure 34: The distributions of the parameter error for both topologies, (upper left) error in the radial position, (upper right) error in the horizontal position, and (lower left) error in the zenith direction. In each distribution, a line marking a "possible" cut is marked.

By using the "proposed" quality cuts in addition to the events to be reconstructed as up-going encompass the so-called quality cuts. These are grouped together depending on the topology of the event the cuts are the following

- **tracks**

1. NNFitTrack_SigmaRClosest < 10m
2. NNFitTrack_SigmaZClosest < 10m

- **showers**

1. NNFitShower_SigmaRVertex < 10m
2. NNFitShower_SigmaZVertex < 10m

Therefore, after applying these quality cuts, the performance of the two models is presented below. Since, from Figures 32 and 33, it is already indicated that the respective algorithms perform better for the intended event topology, the plots presented will be from now on only to the intended particles. Thus NNFitShower will be used as an estimator for "shower-like" events (ν_e^{CC} , ν^{NC} and ν_τ^{CCshow}), with ν_τ^{CCshow} . Apart from the quality cuts, in order to approximate a realistic scenario some muon rejection cuts were investigated. However, since this part wasn't the main objective of the thesis and it was brought up to include also an iteration of cuts that could try to simulate more of an objective detector volume, no optimization was carried out and it consists of a selection of one-dimensional cuts "box-cut", which is inefficient compared to more sophisticated ones such as a Random Grid Search (RGS) method [77], employed by other members in other searches by both KM3NeT and ANTARES groups. Introducing these cuts, would definitely restrict the analysis statistics, but also provide an initial stepping stone for another study in this subject that would use the whole range of the ANTARES legacy-data.

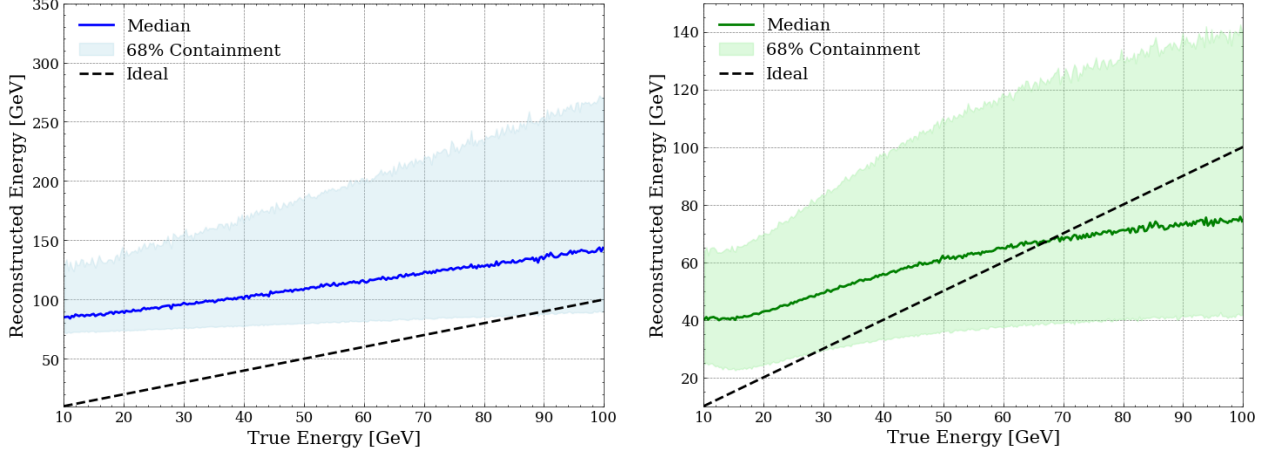


Figure 35: Energy reconstruction performance for both shower-like and track-like events. (left) *NNFitShower* estimator for showers and (right) *NNFitTrack* estimator for tracks after the quality cuts are applied.

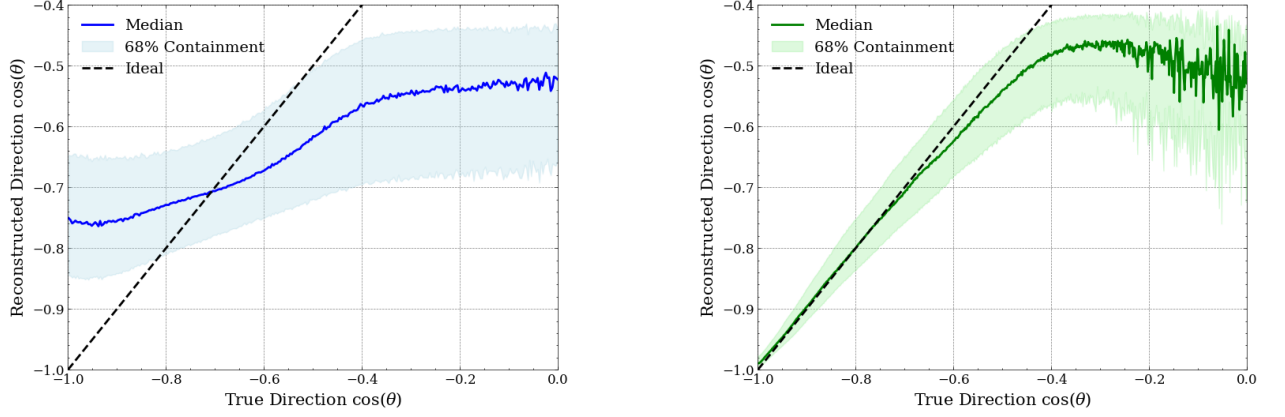


Figure 36: Direction reconstruction performance for both shower-like and track-like events. (left) *NNFitShower* estimator for showers and (right) *NNFitTrack* estimator for tracks after the quality cuts are applied.

With the introduction of the quality cuts the only thing remaining is the adoption of some initial muon veto cuts. Therefore for this very reason, the direction reconstruction and its error are probed. Already there is a significant rejection, by demanding for every event to be classified as up-going. That "loose" pre-cut, indeed provides a significant contribution to the rejection, but by no means is it enough. However, introducing stricter cuts to the reconstructed zenith, would inevitably decrease our statistics and even in case it is not handled with precaution part of the phase-space that the experiment is sensitive to might be lost.

After applying these selection criteria and combining them with the NNFit quality cuts, more than 99.99% of the remaining muons are rejected. Though given, the huge disparity between the signal/background fluxes, still around 1980 muons remain. Given still the high number of muons, further investigation is warranted, possibly introducing other parameters apart from the current reconstruction. Before continuing to the expected number of events, it is still notable that even with the cuts, we still have around 10% for both topologies from the original sample, as seen in Table 5. These two components make up the new cut selection criteria. These "nnfit quality" cuts, have some experimental cuts in the reconstruction quality and some muon veto cuts.

Again according to the event's topology, the following cuts on NNFit reconstruction are in order

- **tracks**

1. $\text{NNFitTrack_SigmaRClosest} < 10m$
2. $\text{NNFitTrack_SigmaZClosest} < 10m$
3. $\text{NNFitTrack_SigmaTheta} < 8m$
4. $\text{NNFitTrack_cos_zenith} < -0.4$

- **showers**

1. $\text{NNFitShower_SigmaRVertex} < 10m$
2. $\text{NNFitShower_SigmaZVertex} < 10m$
3. $\text{NNFitShower_cos_zenith} < -0.4$

Event Topology	low-energy up-going(%)	NNFit Up-going(%)	NNFit Quality cuts(%)
showers	25.55	24.54	17.82
tracks	27.46	27.19	9.66

Table 5: NNFit algorithm efficiency for low-energy up-going, NNFit reconstructed as up-going and after applying some selection cuts.

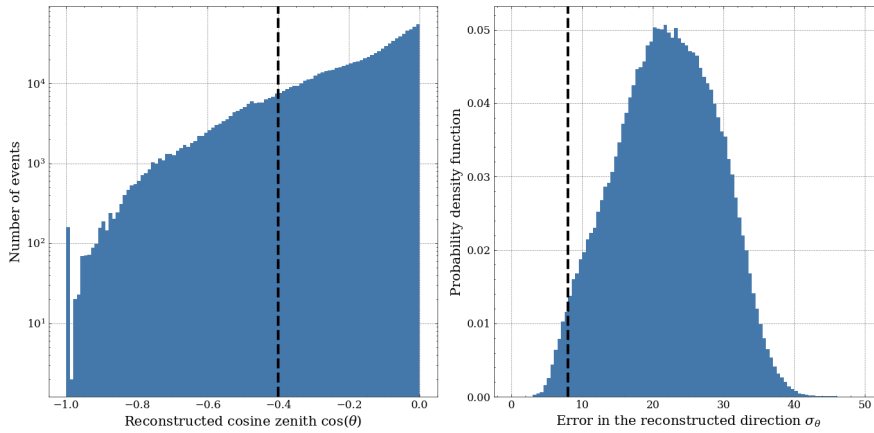


Figure 37: Distribution of the cosine of the zenith direction and distribution of the error of the zenith in degrees. With black dashed lines, the proposed cuts in these two variables, are indicated.

Since NNFit was developed to improve the reconstruction performance of SL BBFit events, the efficiency of the algorithm is quite poor, with only a quarter of the entire sample being reconstructed. However, 96.04% showers and 99.01% tracks are correctly reconstructed as up-going. This is quite encouraging, offering the possibility of a more detailed investigation without leading to a drastic loss of events.

4.5 Expected number of events

With the production of the atmospheric oscillation weights during the processing pipeline, in section 4.3, it is possible to give a first estimation of the expected number of events according to the various cut selection criteria used. Firstly, the "**low-energy true**" cut is applied in order to have an initial expectation of the total number of triggered and reconstructed by at least one algorithm. Afterwards, the first estimate of events reconstructed as up-going NNFit is provided, along with the distribution of (E, θ) phase space seen by the

detector in contrast to the ideal case of a "perfect" reconstruction. Lastly, the "**nnfit quality**" scenario is presented, where the muon rejection and quality cuts are combined, and tables with the remaining events are presented. All the numbers of events quoted in this section contain true low-energy events. So in case of the nnfit cuts, these numbers are not for the entire energy range, but for the energy region of 10 to 100 GeV. The figures contain the expected events of each topology since the detector does not have particle identification capabilities. Since this analysis's main focus is ν_τ , its distribution alone will be given to Appendix:B.

4.5.1 Low-energy up-going

As a reminder about the cut that is applied was discussed in a previous (see Section 4.4). This is a highly-permissive cut, which serves its purpose of providing an estimation of the total number of events included in each flavor. In Table 6, the expected number of events is presented after applying the weights.

Category	Events	Percentage (%)
ν_e^{CC}	7773	15
ν_μ^{CC}	41023	81
$\nu_\tau^{CC\text{mu}}$	1059	2
$\nu_\tau^{CC\text{shower}}$	3355	7
ν_τ^{NC}	4127	8
Total	57338	

Table 6: Expected number of events this low-energy , with percentages.

Figure , presents the distribution of events in the true phase space. The closer the reconstruction plots are to this one, the better the performance of the reconstruction. The binning that is used in this figure is the same across the entire section and the same binning is used in the next chapter of the sensitivities obtained from SWIM. The number of bins in energy and direction were chosen having in mind the energy and direction resolution of the detector.

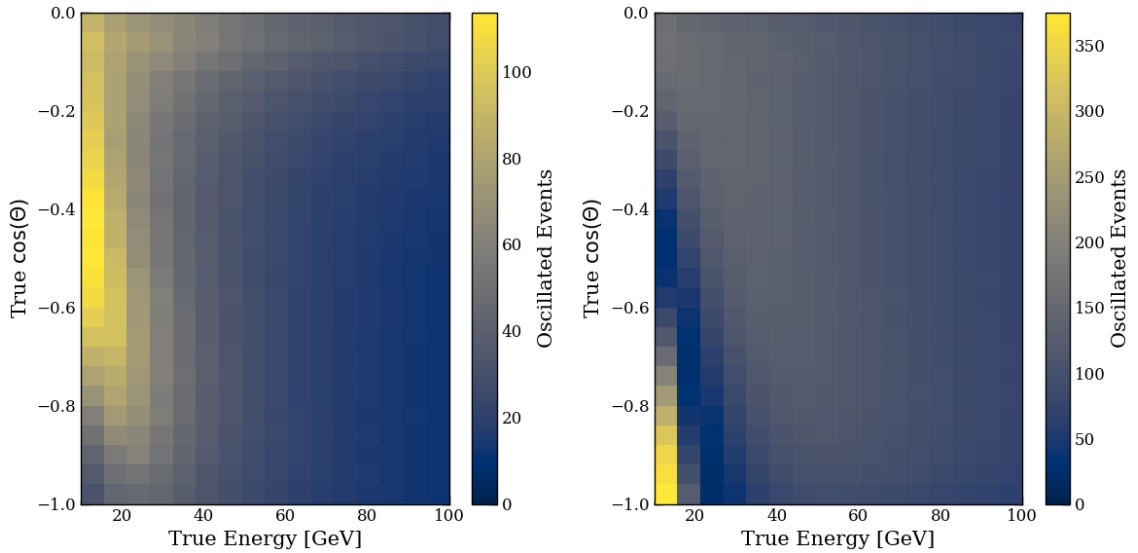


Figure 38: Expected number of events for (left) "shower-like" and (right) "track-like" neutrinos ν

In an ideal scenario that provides us with perfect reconstruction and classification, the distinctive structures from the ν_μ disappearance in the 20 to 40 GeV ranges would be possible to observe even with the current binning. For the physical interpretation behind these structures, see again (Figure 13).

4.5.2 NNFit up-going Events

The first expectations from reconstructed event are discussed in this part. From the previous study, regarding the various reconstruction performances NNFit was chosen as the logical candidate given its ability to sufficiently reconstruct both topologies with an emphasis on the shower-like events. The only cut in this sample is on the direction of the reconstructed zenith. For an event to be included, it needs to be reconstructed as up-going from its chosen reconstruction. Reminder that the model, from the two approximations, that is selected to give information regarding the event is chosen by the topology of the event, since it was already proved that the performance of each topology is better when the respective reconstructions are used.

The expected number of events are given in Table:7

Category	Events	Percentage (%)
ν_e^{CC}	1777	13
ν_μ^{CC}	10587	77
$\nu_\tau^{CC\mu}$	239	2
$\nu_\tau^{CCshower}$	476	3
ν^{NC}	722	5
Total	13701	

Table 7: Expected number of events for NNFit reconstructed events, with percentages.

The distributions of the reconstructed events are visualized as a function of reconstructed energy and direction, is presented in Figure:39.

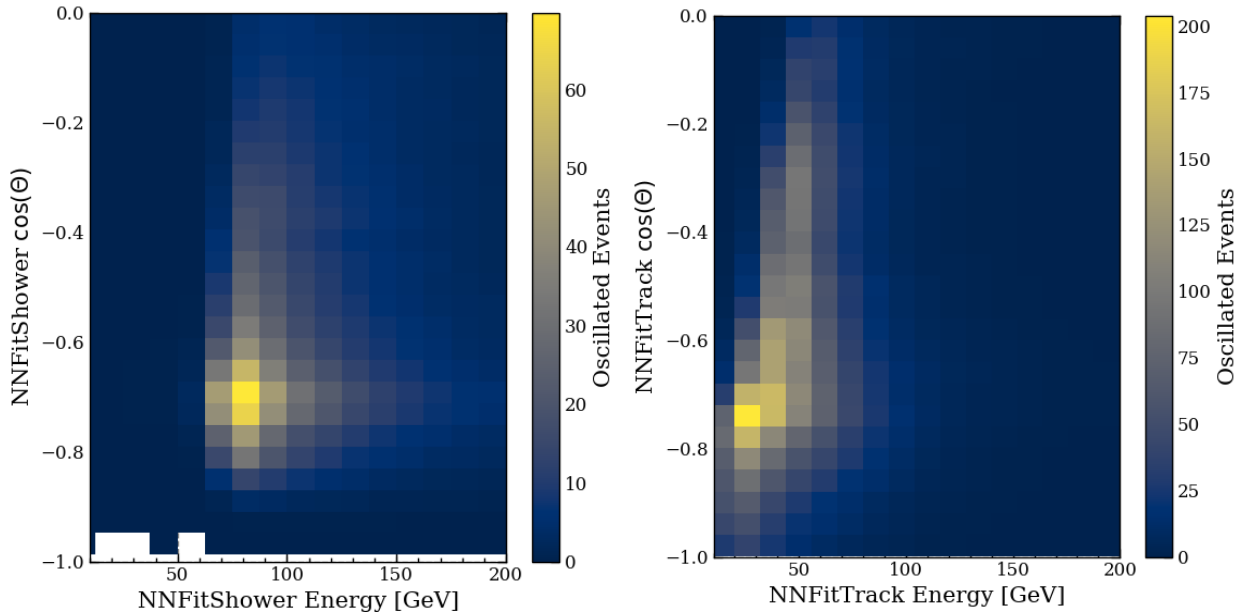


Figure 39: The expected number of NNFit reconstructed events, (left) shower-like events with *NNFitShower* and (right) track-like with *NNFitTrack*. for NNFit up-going events.

Figure 39, *NNFitTrack* is used to reconstruct . In comparison to Figure 38, the fine structures from the oscillation transition probabilities are not visible. Furthermore, in both topologies the main sensitivities come from events below $\cos \theta < -0.6$, without a clear separation between them.

Category	Events	Percentage (%)
ν_e^{CC}	885	13
ν_μ^{CC}	3107	45
$\nu_\tau^{CC\mu}$	67	1
$\nu_\tau^{CCshower}$	372	5
ν_τ^{NC}	447	7
μ	1982	29
Total	6860	

Table 8: Expected number of events for the "nnfit quality" scenario, with percentages.

4.5.3 NNFit quality

After applying the "nnfit quality" cuts, an expectation of the remaining events is presented in this section. This number of events as mentioned is not an optimized sample, but rather a representation of an entry point for a future analysis. In the Figure:40, the distributions after applying the quality and muon rejection cuts are presented. Even though there is a definite reduction in our statistics, the events are more consecrated on the phase-space, with the track class having its main sensitivity coming from below 50 GeV threshold, whereas the shower class has almost no sensitivity in that range.

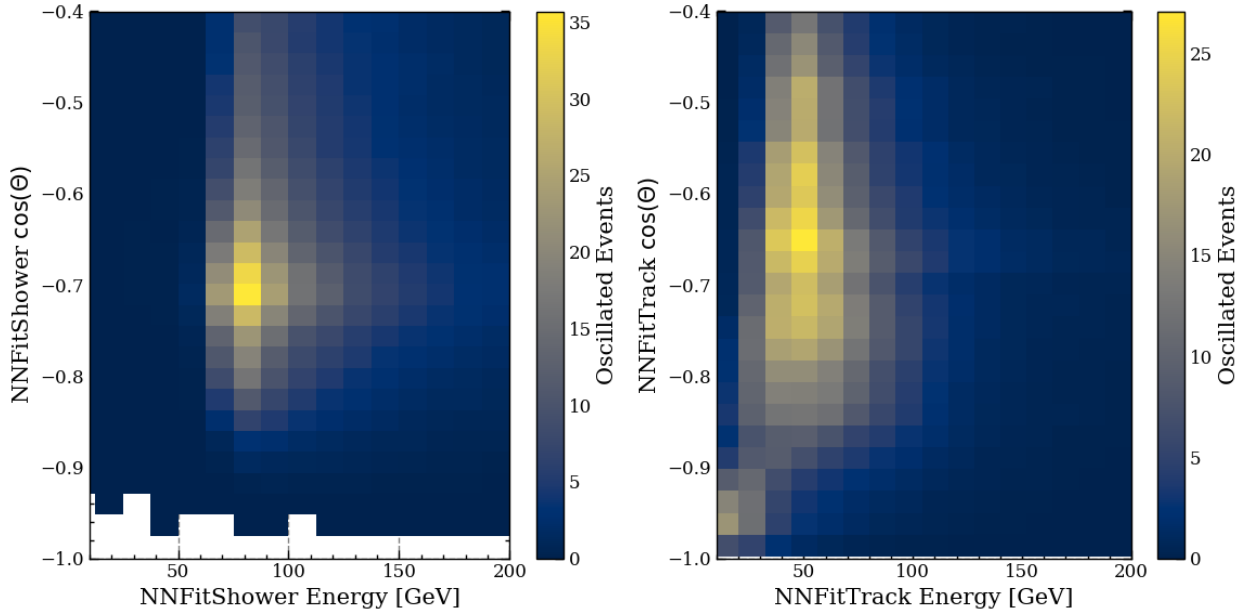


Figure 40: The expected number of reconstructed events in the "nnfit quality" scenario, (left) shower-like events with *NNFitShower* and (right) track-like with *NNFitTrack*.

5 Sensitivity of ANTARES to ν_τ appearance

This chapter focuses on the results of my investigation of the potential sensitivity of ANTARES on ν_τ appearance. The software being used in this analysis is called *Swim* and it is a framework developed by members of the KM3NeT collaboration. A short introduction regarding the in and outs of this tool is the first section of this part (see Sec.5.1). Apart from the mention of the various components encompassing *Swim*, the required files to perform this study are documented. The pipeline responsible for the production of the results in this chapter is discussed Sec. 5.1.2, from the production of the sensitivities when using the reconstructed parameters provided by the various reconstructions in Sec. 4.4 of the previous chapter, or when using smeared information in Sec. 5.3. In the next section, only sensitivities for a specific scenario were selected to be presented. This scenario is probing into the CC channel and assuming the neutrinos follow NMO. This is but a small subsection of the sensitivities produced during this analysis. However since the main task of the thesis is identifying the components of a future potential study, this can be already achieved by presenting one scenario. Lastly, an investigation regarding the potential gain in this study, by improving the current reconstruction capabilities of the detector.

5.1 Swim framework

Swim[69] is an official tool of the KM3NeT collaboration and the main tool used within the *Neutrino Oscillations* working group, the team that investigates neutrino physics phenomena with the ORCA detector. In truth, it is a collection of various packages, among them *OscProb* [35], that offer its users high-level tools for calculating sensitivities from both MC and real data. The analyses benefited that from it, are not only standard 3-flavor model oscillation, like this one, but also sterile neutrino, Lorentz invariance violation, non-standard interaction and other BSM analyses.

The tool is written with an Object-Oriented approach using the C++ programming language. An important dependency of the software is another framework also based in C++, ROOT. It is the only available data format for the event files. This occurs because *Swim* is based upon libraries concerning the data formats (**TTree**) and containers for histograms, those being standard or sparse higher dimensional ones (**THx**) and (**THnSparse**) respectively [78]. Another required tool, that was briefly mentioned before is *OscProb*, which is a library for the calculation of the neutrino oscillation probabilities initially written by Joao Coelho for long baseline experiments. Slight reminder, this tool was used during the calculation of the neutrino oscillation weights during the processing step for finding out the neutrino oscillation probabilities (see more in Sec. 4.3). In the context of *Swim*, *OscProb* provides the necessary tools for the computation of the detector response matrix. Lastly, during the minimization step of the fitting process, the *Minuit2* [79] library is also needed.

The main packages that comprise *Swim* are the following, each one containing classes for various functionalities

- OrcaSim: main classes for the experiment simulation, links the other classes
- OscProb: neutrino oscillation probability calculator
- OrcaDet: detector response simulator
- NuFitter: interface with *Minuit2*
- OrcaSyst: classes to handle systematic uncertainties
- Utils: common utilites classes

A figure illustrating the links between the various classes in *Swim* is presented below in Fig. 41. From the flowchart the User input is indicated, the first being a file containing the event information in ROOT format. For this analysis, the production and conversion of the event files are discussed in length in the previous chapter Chap. 4. For the neutrino atmospheric fluxes, tables from the Honda group the HHKM15 model is used from the paper [70]. Two tables are used from the measurements taken in Gran Sasso and Frejus. For ORCA and ANTARES, the Frejus approximation of the neutrino flux is used due to the site's proximity from the two experiments. These tables, are then converted to a graph that is saved into a ROOT file that is already provided inside the repository of the framework. Likewise, ROOT files containing the cross sections

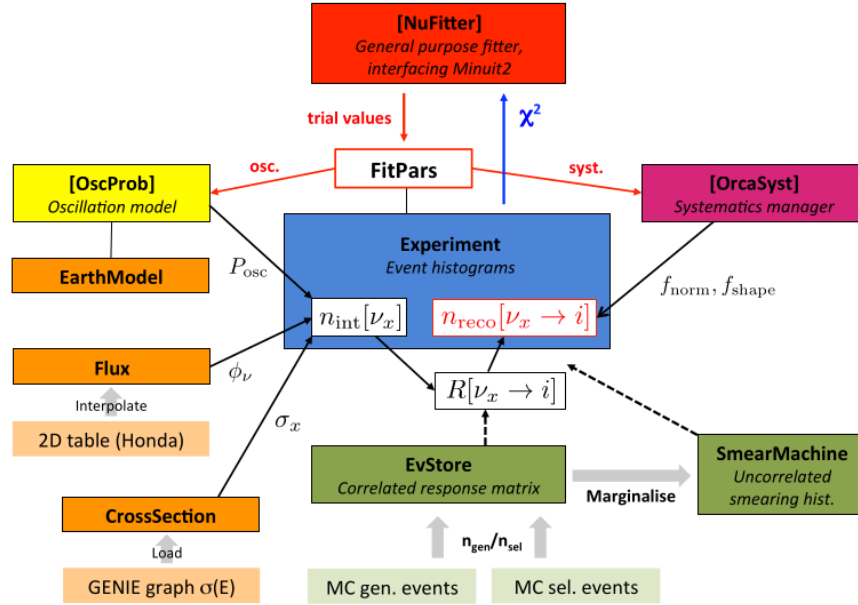


Figure 41: The interactions of each class of Swim are indicated within this flowchart. The lighter colour in the background point to User input. Taken from [78]

from several neutrino generators, such as GiBUU[80], GEHEN[56] and gSeaGen[81], are provided both for low and high-energy neutrino scenarios. For this analysis, the cross section file from the latest gSeaGen version available is used and the reason behind lies in the fact that it is the same model currently used by the KM3NeT oscillation working group. More specifically, the file used is *xsection_gsg.v7r2_10TeV.root*.

A typical analysis, when using this framework has the following steps:

1. Configuration with JSON file: this is achieved with the parsing of various JSON[82] files, that define various input settings, systematic uncertainties, classes and selection cut criteria and which parameters remain free to be fitted according to the model
2. Load User input: such as events files, neutrino flux model, cross sections, earth model, etc.
3. Detector response computation: according to the selection criteria chosen, which in turn are applied to the MC events
4. Expected rates of interacting neutrinos calculation: achieved by integrating the various input models taking into account the detector response
5. Event calculation: from the expected rates the number of reconstructed events is computed for a given exposure rate
6. Fit: the data are then fitted according to the chosen model and a negative log-likelihood is formed to quantify the quality of the fit

Before continuing to the user input files and the various configurations that need to be defined, when running the software, one class will be briefly mentioned due to its central role within *Swim* this being the *Experiment* class. In every analysis, there are two *Experiment* objects with one representing the pseudo-data and the other the model, according to which the data will be fitted during the minimization process. Each object contains oscillograms (two-dimensional histograms for each oscillation scenario, 18 in total), distribution of the event rates for each interaction type in units of $Mt * year$, and distributions for the reconstructed event numbers in each event class.

What follows is a summary of the phases in the computation chain that are specific to the *Experiment* class:

1. Calculation of oscillation probabilities
2. Computation of interaction event rates
3. Incorporation of the detector response using the *EvStore*
4. Combination of interaction channels histograms

Assuming that we measure our expected data—that is, TauNorm equal to 1 for normal mass ordering—the sensitivity analysis aims to determine how well it will be feasible to exclude specific values of the tau normalization. Finally, for the ν_τ appearance study, the fitting process used is a χ^2 profile, a scan of the tau normalization parameter, TauNorm defined in *Swim*, for values from 0 to 2 is performed and a likelihood value, χ^2 , is calculated. This χ^2 is an indicator how well of an agreement there is between the model and the data. For every point, a binary hypothesis test is performed, with the first hypothesis being the expectation from the models (atmospheric fluxes, earth models and oscillation scenario) which will be fitted to the pseudo-data. Here, the tau normalization is fixed to the desired value. The second hypothesis is obtained from our pseudo-data, that being MC or real data, that try to simulate reality. The steps of the fitting process are gathered below,

1. The pseudo-data are filled into a 2D-Histogram (E, θ) with the naming convention "data"
2. The expected events from the defined model are filled into a 2D-Histogram (E, θ) with the naming convention "model"
3. The "data" and "model" are fitted with a range of fixed values of TauNorm $N_{\text{test}} \rightarrow$ calculate χ^2_{fixed}
4. The fit is again run by releasing the TauNorm parameter \rightarrow calculate χ^2_{free}
5. For each value of TauNorm the $\Delta\chi^2$ difference is calculated $\Delta\chi^2 = \chi^2_{\text{fixed}} - \chi^2_{\text{free}}$

The χ^2 is a defined Poisson χ^2 that consists of two parts a statistical and a systematic one.

$$\chi^2(\text{model}, \text{data}) = 2 \sum_{ij} \left(n_{ij}^{\text{model}} - n_{ij}^{\text{data}} + n_{ij}^{\text{model}} \log \left(\frac{n_{ij}^{\text{data}}}{n_{ij}^{\text{model}}} \right) \right) + \sum_{\alpha} \left(\frac{\alpha_{\text{exp}} - \alpha_{\text{obs}}}{\sigma_{\alpha}} \right) \quad (56)$$

The first sum corresponds to the statistical, which runs over every bin of each individual histogram. The second section, known as the systematical part, contains all nuisance parameters that have restrictions, meaning that a "prior" has been defined for them. This term penalizes significant deviations from the expected value about this parameter's uncertainty. Since in this analysis, the focus lay on the quality of the data from ANTARES, no concentrated effort was invested in the study of the treatment of the systematic uncertainties. Instead, the template was taken from the oscillations working group. Apart from the previously stated reasoning, another point is that mainly the systematic uncertainties were derived from the environmental conditions of the experiment such as water properties, which are the same for both detectors due to their location proximity and environment resemblance. Lastly, the data set used for this analysis is a so-called Asimov set, where in this approach statistical fluctuations are neglected and each bin in the two histograms (data and model) is set to the expectation value given the defined configurations. From the nature of the Asimov set, it is expected that the TauNorm will converge to the expected value of 1, because it is a MC dataset.

From Figure: 42, the upper limit in the achieved sensitivity could be seen. The sigma confidence level is calculated as $\sigma = \sqrt{\Delta\chi^2}$, where the "model" can be exempted, for the given pseudo-data presented.

5.1.1 User configurations

As explained in Section: 5.1, offers a lot of configurations during its run, giving the user a lot of freedom with a wide choice of customization. This is accomplished with the *ParseFitter* library that enables the import of JSON files for the definition of some user-defined parameters. There is the possibility to include the necessary definitions inside the source code itself without the use of *ParseFitter*, however, it is not recommended since for every slight change in the import of some files a new compilation would be required. There are four main types of JSON files required to run a script with one optional, these are:

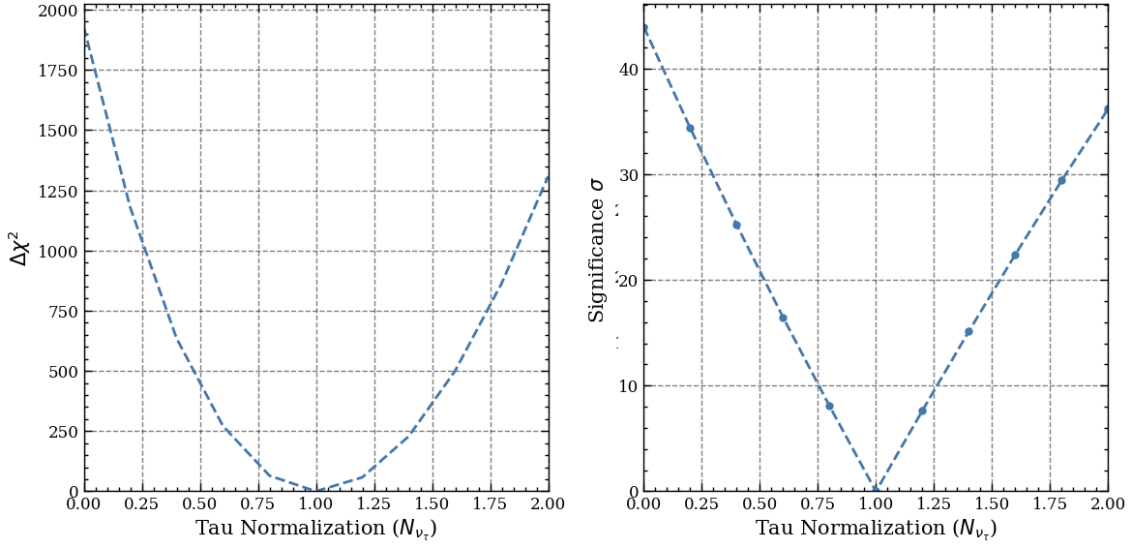


Figure 42: Under the assumption of perfect classification and reconstruction the theoretical all being "unrealistic" upper limit in the ANTARES sensitivity by freeing only the TauNorm parameter. (Left) the χ^2 profile is given and on the (right) the confidence level plot.

- **classes.json**: defining the separation of classes that events are grouped, the declaration of the cut selection criteria used and the selection of the reconstruction algorithm used for the parameters (E, θ, y)
- **variables.json**: defining the general types of the analysis, highly correlated with the *Experiment* object. Some of the variables defined here are the name of the file that contains the pseudo-data, the type of data set (Asimov or Pseudo-Experiments), the channel to probe into, etc.
- **parameters.json**: the values for each necessary parameter are defined for the two hypotheses ("model", and "data"), along with their priors. Apart from the assignment of values, the parameters that will be fitted are assigned with the "free" value, whereas by "fixed" the parameters will use their prior value during the fit..
- **binning.json**: the bins of the histograms created during the entire process for the three parameters (E, θ, y)
- **User.json**: parses information directly to the source code of the analysis and defines the parameter that would be scanned, i.e. TauNorm for our case, the range of the scan and many other user-defined parameters

The first four JSON files are necessary for the execution of every analysis with the latter being only used when χ^2 profile fit is selected. It is important to pay close attention to some of the definitions since there might conflict with globally defined parameters, such a case will be mentioned later during the discussion concerning the classes JSON file.

The proper definition of the classes plays very important in this analysis, since as already explained in Section:2 this measurement looks for a statistical excess of shower-like events with response to the theoretical expectation from the model. For this very reason, the events are grouped according to their topology. In case of ORCA[83] dedicated classifiers have been developed with different philosophies such a BDT and a RGS [84]. For ANTARES a BDT has been developed, but is based upon AAFit with high-energy neutrinos in mind [43]. There is no need to add extra classes since an error occurring during particle identification is impossible. Whereas in the case of ORCA, they assume three classes by splitting the "track" class into low and high-purity ones. The reconstructions used in this analysis are AAFit, NNFit and MC-truth. The reconstructions used in these analysis are AAFit, NNFit and MC-truth. More specifically for AAFit both estimators, *AAFit_ANN_ECAP* and *AAFit_dedX_CEA*, from Section:4.4.1 are included in separate sensitivity

and applied in both class's calculations. For the case of NNFit both models are used, with *NNFitShower* for the "showers" class and *NNFitTrack* for the "tracks" class, see Section:4.4.3 for more information on these two models. Due to the high variance and the vulnerabilities of the energy estimator, another possibility is introduced to include "perfect" energy reconstruction and only use the direction reconstruction of NNFit. Lastly, MC-truth corresponds to the values directly taken from the true branches, therefore when we use this reconstruction we imply that there is a perfect level of reconstruction. During the discussion of the processing pipeline, Sec, 4.3, for the last "reconstruction" to be implemented a workaround was devised. In the analysis, some sensitivities with smeared events will be presented. That is based upon several methods of smearing upon the MC-truth and are discussed in Section:5.1.2. In this JSON file, the cut selection is also defined these being "**low energy events**", "**nnfit up-going**" and "**nnfit quality**", which introduced and discussed in the previous chapter in Sec, 4.4. Excluding the smeared option, there are 5 scenarios per selection criteria. For the smeared options, the only difference between them is which events will be included so the "classes" JSON file is not influenced, thus coming to a total of eighteen for all cut options. The cuts that are defined in the files include two parts, the first being the event selection cut discussed in the last chapter and the "perfect classifier". For this classifier, true information of the events is used such as particle and interaction type, thus the resulting cut is presented:

- Tracks: ((abs(type) == 14 \vee interaction_type == 1)) \vee (abs(type) ==13) \vee (abs(type) ==16 $\&\&$ interaction_type == 2)) \wedge (selection cut)
- Showers: ((abs(type) == 12) \vee (abs(type) == 16 \wedge interaction_type == 3) \vee (interaction_type == 0)) \wedge (selection cut)

with the first part being the classifier and the second part being the selection criteria. Two implementations for the classifier have been done, first, the one shown above where every possible interaction type is mentioned and assigned to the proper topology. With the "tracks" being $\nu_{\mu}^{CC}, \mu, \nu_{\tau}^{CCmu}$, the muon neutrinos that interact through CC and the tau neutrinos, whose lepton decays into a muon for our signal and atmospheric muons that have passed through selection criteria as background. Whereas for the "showers" class, we have $\nu_e^{CC}, \nu_{\tau}^{CCShower}, \nu^{NC}$ electron neutrinos that interact through CC, the tau neutrinos, whose lepton decays into a shower and all neutrinos that interact with NC. The other implementation is simply to define one of the two classes and simply define the opposite, since we have only two classes. Both have off course the same effect, but the first one was chosen to explain easily which events are included in each class.

Most of the multi-dimensional histograms have the following parameters (E, θ ,y). In the context of the ANTARES experiment, the measurement of the Bjorken y parameter, which represents the fraction of energy transferred to a hadronic shower, is not possible, thus resulting in an unknown parameter, thus effectively the histograms become 2 dimensional (E, θ) . It is important though that the event statistics in each bin are not too small to avoid an overestimation of the χ^2 . The binning is similar to the standard from ORCA, with some notable changes. Since ANTARES is not sensitive in the ranges of a few GeV, the starting bin is 10 GeV. With 14 logarithmic bins up to 100 GeV. Due to the energy overestimation, there is an additional bin for energies above this threshold up to 1 TeV. For the direction, the binning applied to the $\cos \theta$ is assuming an up-going direction so from -1 to 0 with 20 bins. The initial binning started with 15, but given the detector volume that was higher than expected the decision was made to try to benefit from the performance of NNFit.

The parameters.json declares all necessary information about the oscillation parameters and systematical uncertainties. Since the scenario that has been chosen is for NMO, there is no need to have an additional parameter file that assumes the other model, however in theory it is possible to create all the plots assuming the other mass ordering. However, there are two different files depending on the number of parameters that are freed during the fit process. In the "restricted" scenario only the TauNorm parameter is freed, where the model assumes the prior values given for the other parameters. The more "realistic" scenario accounts for the statistical uncertainties for the other parameters and is also freed together with the TauNorm. The systematic uncertainties contain two types of parameters the first is related to neutrino oscillations. These are gathered on the upper part of the tables. The latter are correlated to some environmental conditions such as the water properties and PMT efficiency. The assumption is made systematic uncertainties of ORCA are the same as in ANTARES, given the close nature of the experiments. In the case of ORCA, an active effort is made for the determination of these uncertainties. Thus the same template is used. The two versions of the parameter.json are presented in Tables 9 and 10, just below:

Parameter	Initial value	Prior
θ_{12}	33.44 (NO) 33.45 (IO)	fixed
θ_{13}	8.57 (NO) 8.60 (IO)	fixed
θ_{23}	49.2 (NO) 49.3 (IO)	fixed
Δm_{31} [eV ²]	2.517e-03 (NO) -2.4238e-03 (IO)	fixed
Δm_{21} [eV ²]	7.42e-05	fixed
δ_{CP}	197.0 (NO) 282.0 (IO)	fixed
Energy Slope	0	10%
Zenith Slope	0	2%
ν_μ skew	0	5%
ν_e skew	0	7%
flavor skew	0	2%
NC scale	1	20%
Energy scale	1	9%
FluxNorm	1	free
TrackNorm	1	fixed
ShowerNorm	1	fixed
MuonNorm	1	fixed

Table 9: All systematic uncertainties and their treatment in the fit for the "restricted" scenario.

Parameter	Initial value	Prior
θ_{12}	33.44 (NO) 33.45 (IO)	fixed
θ_{13}	8.57 (NO) 8.60 (IO)	fixed
θ_{23}	49.2 (NO) 49.3 (IO)	free
Δm_{31} [eV ²]	2.517e-03 (NO) -2.4238e-03 (IO)	free
Δm_{21} [eV ²]	7.42e-05	fixed
δ_{CP}	197.0 (NO) 282.0 (IO)	fixed
Energy Slope	0	10%
Zenith Slope	0	2%
ν_μ skew	0	5%
ν_e skew	0	7%
flavor skew	0	2%
NC scale	1	20%
Energy scale	1	9%
FluxNorm	1	free
TrackNorm	1	free
ShowerNorm	1	free
MuonNorm	1	free

Table 10: All systematic uncertainties and their treatment in the fit for the "realistic" scenario.

In the variables.json file, some general parameters are set related to the Experiment object. Firstly, the name of the event file that is used needs to be specified if it is the same as the naming convention of the files that have already undergone mass production and conversion to Swim appropriate type (standard ROOT). For the case of ANTARES, no files had been prepared like ORCA and "Astroparticle Research with Cosmics in the Abyss" (ARCA), thus, of course, the need for the processing pipeline that was discussed in the previous chapter, 4.3. The exposure time of the experiment needs to be stated in units of years, where for this analysis since it uses the entire lifetime of the detector it was 12.433 years. This number came after a discussion with colleagues and the lifetime of the detector was measured in days each year from the fifteen years of detector operation. The paths of some models are defined here, such as for the cross section (gSeaGen version 7.1), for the composition of the earth interior PREM and the flux model (HHKM15 in the Frejus site). Another important role of this file is to define the treatment of the MC events during the computation of the detector response matrix, such as the type of weights to be included and options that enable the calculation of MC errors. These options are present to avoid an overestimation of the χ^2 due to finite statistics in MC. To that reason the "Analysis_type" used in the analysis is with an Asimov dataset, whose effect has been mentioned before in this section. Lastly, the "Experiment_type" defines the type of analysis declared in this file. This provides some already prepared classes that simulate various physics analyses such as standard oscillations, tau measurement and several BSM models. For this analysis, the first two options are applicable with the first measuring the CC channel and the latter both CC and NC. By choosing one of the two types of analysis, it is feasible to probe into various channels seamlessly.

Finally, the user.json file is used to parse some variables into the fitting source code. Depending on the type of fit used in the analysis this file might be optional. For the χ^2 profile, it is necessary to decide which parameter, will be scanned along with the number of points fitted during the whole process. The type of fit can be defined, with the *SimpleFit* performing a fit according to the loaded configuration and the *FitTwoOctants* executing two fits for each point one per θ_{23} octant and storing the one with the worst χ^2 . This includes, all the parameters that can be defined from this file that configures the fit, without the need to re-compile the

code each time, but alternating over the fit options without effort. In my implementation additional variables were included in order to define the exact path, where the results of the fits will be stored to the appropriate folder.

All the JSON files required for the execution of a χ^2 profile fit are stored in Appendix:D. These are not the only files created and needed for all the plots presented in Section:5.2, 5.3, and 5.4, but serve as a template. In practice, by having the templates as a guide and taking into consideration the discussion in the section all required files can be produced.

5.1.2 Analysis pipeline

In this section a detailed discussion of the analysis pipeline is presented here. The various steps are discussed and can be gathered in three main sections, (1) the file preparation phase, (2) the χ^2 fit and (3) the plotting phase. The various steps are visualized in a flowchart in Figure:43.

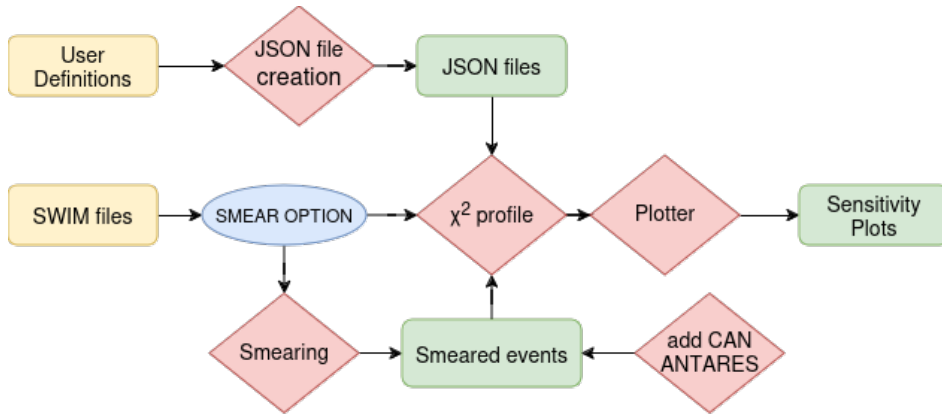


Figure 43: Flowchart of the analysis pipeline. The yellow refers to the files or variable definitions that act as input in the pipeline. The blue oval is a logic gate, where depending if I want to smear or not my data different paths are chosen. With red the various scripts used in the pipeline are defined and by green the files produced during the whole duration of the pipeline those being JSON files, smeared events in a ROOT format or the sensitivity plots.

The flowchart from Fig:43, acts as an individual job in the computing centre that acts completely independently from others. The lack of dependencies offers a high degree of freedom and can make the pipeline highly parallel being able to submit multiple jobs in parallel with minimal oversight. The only intervention afterwards is to create combined sensitivities plots, visualizing a range of different scenarios together. The last step is the only one that is run directly on the front end of the cluster, with all the other being submitted through SLURM. Every individual job requires minimal resources each one running on one CPU with just 3 GiB RAM. The computing time varies from scenario to scenario but can vary from around 2 hours when the only fitted parameter is TauNorm to up to 12 hours when the full range of systematic uncertainties is enabled. Despite the relatively computationally light workload of each job, due to the high number of different options, the toll on resources can be noticed if someone doesn't pay attention. Therefore, some combinations of jobs can be run in parallel and avoid unnecessary re-calculations that would needlessly take extra computational time. These pitfalls will be explained in the individual processes that someone should be aware, of to be efficient with his job submission.

In the first step, all the files are produced if not already present. One script is in charge of the JSON files that are required for a successful execution of the χ^2 profile. The types of JSON files required for the fit have already been mentioned in the previous section, see Sec:5.1.1, since there are a lot of version of the same type of JSON file, a list of variables is parsed into the script and the files needed for this particular scenario are produced if an already existing file doesn't exist. The arguments required by the script are the following:

- **channel:** corresponds to the *Experiment_type* from the variables.json file, with STD corresponding to the standard oscillations scenario that probes only into the CC case, whereas TAU correspond to the

tau scenario that measures both channels together

- **ordering**: corresponds to the selection of mass ordering with NMO and IMO the two options
- **reconstruction**: corresponds to the algorithm chosen to use its reconstructed values for (E, θ) , with the available options being MC, AAFit dEdX, AAFit ANN, NNFitand NNFitdirection only.
- **systematics**: for the systematic uncertainties scenario that will be chosen, the one that frees only the TauNorm parameter during the fit, the so-called "restricted" scenario. Or the scenario that frees the same parameters as the ORCA analysis.
- **smeared**: is a boolean variable that asks if the user wants to use smeared reconstructed events or not
- **smear level**: (optional argument) necessary if the smeared option is chosen and defines the level of smearing that will be applied to the MC-truth branches. The smearing strategy can either be a "flat-based" one or a more realistic one that takes into account the resolution of the detector.
- **asymmetry**: (optional argument) also a boolean argument that asks if the smearing that will be applied is on the same level for both energy and direction
- **asymmetric factor energy**: (optional argument) introduces a factor that will be applied to the energy branch to initiate an asymmetry to the level of smearing between the two parameters
- **asymmetric factor direction**: (optional argument) introduces a factor that will be applied to the direction branch

As mentioned above, not all arguments are necessary in each scenario, with the option of smearing needing four additional arguments depending on the smearing strategy. In every job, if the files are not already there, 6 JSON files are created and the question would arise if there are 5 types of JSON files are there, where the additional files come the reason is quite simple, in every job two fit are performed, as stated before the end goal is a $\Delta\chi^2$ plot. For this very reason, in the same job, first the χ^2_{free} is calculated and afterward the χ^2_{fixed} and each fit needs each individual User.json.

Depending on the scenario there might be a possibility that the initial event file undergoes smearing. In this process a new file is produced with two additional branches where the smeared variables are stored. It was chosen to create an entirely new file instead of appending the two new branches in the same file, due to some potential issues that might arise when using the *kOverwrite* class from ROOT, where the file could be corrupted and would result in the loss of the file. Thus it was safer for an entirely new file to be created and given that the initial size of the file is only around a few GiB. The approach of the smearing is quite straightforward, for each event, a Gaussian is created for each parameter separately with the mean value of the distribution being the true value and a error. Here it should be noted that there are actually two strategies for smearing the "flat-based" and a "resolution" one. In the case of the "flat-based", a smearing factor is introduced. In this analysis it is colloquially named as *smear level*. In this strategy the error of the Gaussian distribution is proportional to the true value and the level of smearing, with the Full Width at Half-Maximum (FWHM) being $\text{FWHM} = q_{\text{true}} \cdot \text{smear_level}$ with q being the parameter smeared either energy or cosine of the zenith direction. For example, during a smear of 20% an event with energy 50 GeV would have a standard deviation of $\sigma = \frac{0.2 \cdot 50}{2\sqrt{2\ln 2}} = 4.24$ GeV. From the distribution, with the help of a randomizer a value will be sampled to be the new smeared value. The library for the randomizer and the statistical distributions are both from ROOT, those being *TRandom3* and *TMath*.

From Figures 44, one can easily make a cross-check if the events are smeared following a normal distribution and if the standard deviation is proportional to the mean value of the distribution and the smear level. The strategy is the same for both energy and direction, but there are some issues if nothing else is done. There are some boundary conditions for both parameters, one being that there can be no negative energies. If the energy of an event becomes less than the detector's energy threshold of 10 GeV it would be lost from the sample. For the direction, there are two boundary conditions for the smeared cosine zenith to be between -1 and 1. One region needs to be paid attention to since the standard deviation of the Gaussian is proportional to the true value for events close to the horizon the cosine zenith would be around 0 and as a result the σ . This is also present in the direction distribution in Fig:44, where the gaussian is pruned for values below -1.

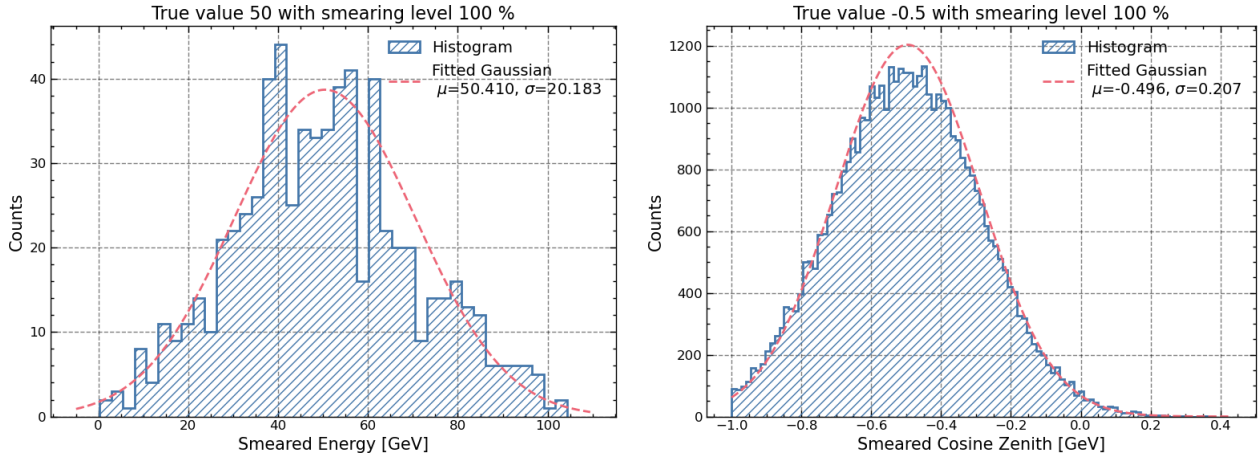


Figure 44: Distributions of the smeared parameters for a chosen true value with a smearing level of 100% . (Right) the distribution of smeared energy values for events with true energy 50 GeV and (left) the distribution of smeared cosine zenith values for events with true cosine zenith of -0.5.

Therefore a threshold was chosen, so even if the values were close to zero still a minimum error was chosen, in order to apply the smearing strategy to all events. The difference with the introduction of the minimum threshold is visualized in Figure:45, where the distribution of smeared cosine zenith was plotted for events that were exactly from the horizon. In the figure on the left without smearing, the distribution is the same as the true one. However, as it is evident there is a bump from below to above the horizon, the reason behind it is that these two types of events are approximated with two different functions and this behaviour comes directly from the generator. Before I mentioned these boundary conditions. However depending on the level of smearing some events might take these values when initially sampled due to the spread of the normal distribution, in this case, the code forces to do another sampling and if the new value is within the boundaries it is accepted and the analysis continues to the next event to be smeared. This occurs in both branches and depending on how "high" is the smear level, the need for additional sampling increases.

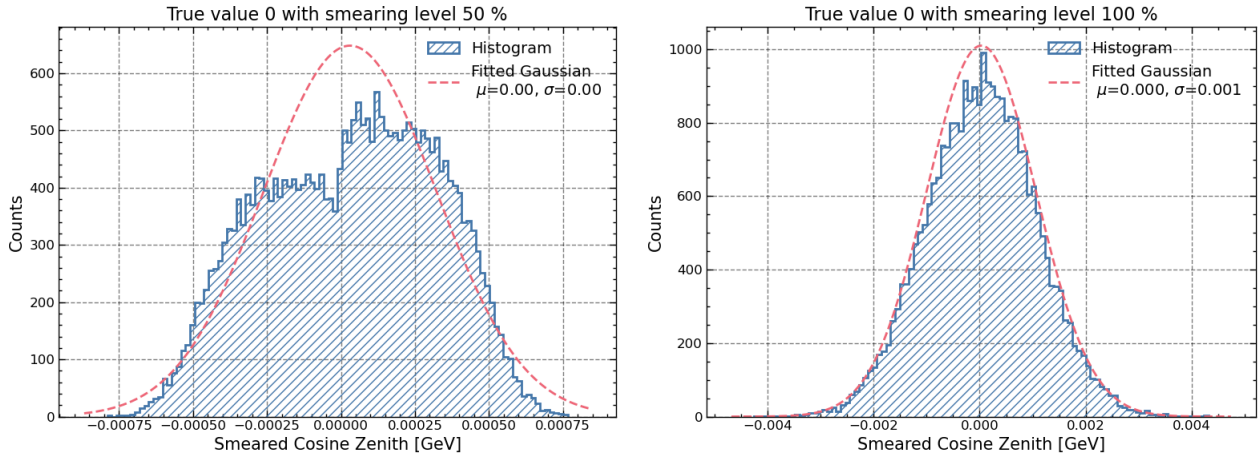


Figure 45: Distributions of the smeared cosine zenith for events from the horizon. (Right) distribution without the minimum threshold for an error, (left) distribution after the inclusion of the correction.

Since this point, the discussion focused only on one strategy, the "flat-based". The other approach uses the resolution of NNFit for the smear, since before the smear was only proportional to the true values without taking into consideration the detector at all. Here is, when the other implementation arises, by using the resolution of the algorithm to be the FWHM of the normal distribution. To that end, the resolution of the detector needed to be parameterized and fitted as a function of which values, true energy and direction for

both cases? To resolved this issue between the one or two-dimensional resolution for both energy and direction, the resolutions of the detector, where plotted as a function of both true parameters. By resolutions we define,

$$\text{res}_q = \frac{|\Delta q|}{q_{\text{true}}} \quad (57)$$

with q being energy or cosine zenith

The function used to parameterise the resolution for both cases is hyperbolic. For energies, it is expected since lower energies have worse performance, due to the lower amount of photons produced during the interaction. As seen in Section:4.4.3, up-going shower events with a direction parallel to the DUs perform the worst during direction reconstruction. Thus the hyperbolic function can too approximate the direction resolution. The exact constants fitted are the following:

$$y = \frac{\alpha}{(\beta x + \gamma)} + \delta \quad (58)$$

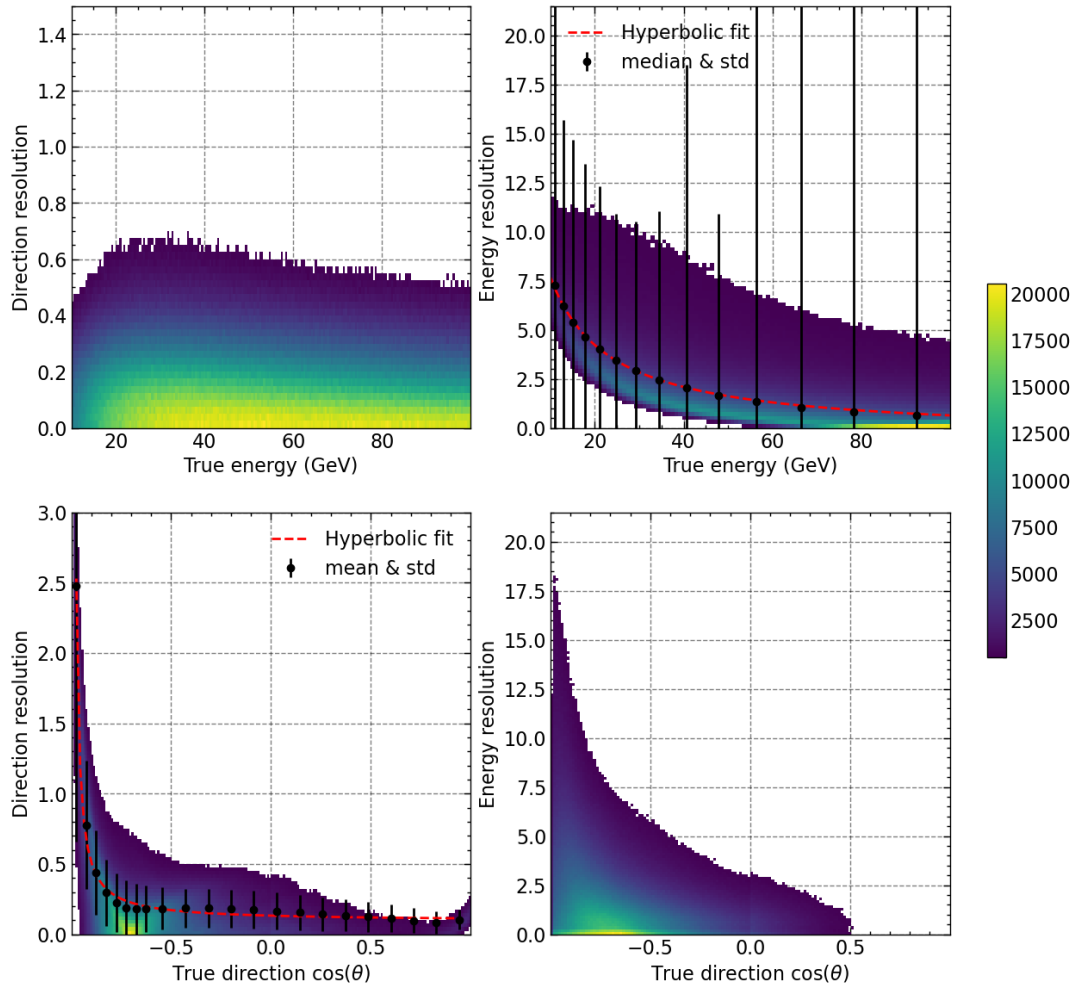


Figure 46: Resolutions of NNFit as a function of true energy and cosine zenith. (Upper left) direction resolution as a function of true energy, (lower left) direction resolution as a function of true direction with the hyperbolic approximation, (upper right) energy resolution as a function of energy with fit approximation and (lower right) energy resolution as a function of true direction.

These four constants were approximated, to parameterize the energy and the direction resolution, with the constants after the fit gathered in separate tables, Table:11 and Table:12. The reasoning behind the

Constant	Value
α	-640.34
β	-5.88
γ	-20.44
δ	-0.41

Table 11: Parameterization of the energy resolution

Constant	Value
α	-21352.43
β	-561954.61
γ	-556675.55
δ	0.09

Table 12: Parameterization of the direction resolution

one-dimensional approach of our hyperbolic function is after the visualization of the resolution distribution in Figure:46. In the direction resolution, a sharp drop is first seen with events directly up-going, their direction is parallel to the detector lines of ANTARES, having the worse performance, something that was known already. We don't see any influence from the energy in the direction resolution for this algorithm, which is not entirely the case for the energy resolution. The difficulty of the algorithm to reconstruct properly the event with that kind of directly is reflected in both energy and direction reconstruction. That being said, there is already such a significant variance in all bins where the resolution was calculated, that was decided that the energy resolution would too have a one-dimensional approximation. This assumption was achieved, because the influence of the particle direction, wouldn't have an important effect given the current level of energy reconstruction. As already mentioned previously, during the production of the JSON files, an asymmetric factor was introduced to account for the difference in the performance between energy and direction.

These are applied to the FWHM of the parameter when we want to introduce essentially different performance when using the "flat-based" implementation of smearing. Depending on the value of this factor, this can either increase or decrease the width of the distribution, thus either improving or worsening the quality of our events. From NNFit, it is evident that the direction resolution is better compared to energy, but since the algorithm is not optimized for these ranges, the reconstruction of ORCA was chosen as an example [85]. The median resolution for energies, in the region where ν_τ has the highest probability of appearing (reminder Figure:13) for the worst performing class is around 60%. For the case of direction it was around 20%. In the case of full ORCA with 115 DUs, the difference between the algorithm drops to 20% and 10% respectively. Although these values are rough estimates, still the fact that the direction resolution is better than the energy resolution is clear. Of course, the direction reconstruction uses the timing resolution of the detector, where even for the case of ANTARES, is sufficient to reconstruct this low energy sample. For energy estimation, the number of photons that arrive to the PMTs plays a crucial role. Therefore, the sparse instrumentation of ANTARES compared to ORCA hints this endeavour. Therefore during the asymmetric smearing, an extra factor of $\alpha_E = 3$, was chosen to roughly approximate this imbalance. The approach of symmetrical smearing has some issues due to the drastic difference between the energy and direction reconstruction tasks. Thus it is highly unlikely that both tasks will exhibit the same performance. So if someone tries to approximate the energy resolution, then there is an underestimation of the direction resolution. The approximation of the direction, would lead in an overestimation in energy. Lastly, the several options of smearing that were run and their sensitivities are included in the analysis are the following:

- ANTARES: smearing with the "resolution" approach
- ORCA 6DUs: "flat-base" smearing that uses the median values quoted above
- ORCA 115 DUs: "flat-base" smearing that uses the median values quoted above
- 10%: "flat-based" smearing with 10% direction and 30% energy resolution

- 30%: "flat-based" smearing with 30% direction and 90% energy resolution
- 50%: "flat-based" smearing with 50% direction and 150% energy resolution
- 70%: "flat-based" smearing with 70% direction and 210% energy resolution
- 90%: "flat-based" smearing with 70% direction and 210% energy resolution
- 100%: "flat-based" smearing with 100% direction and 300% energy resolution

For the first 3 cases, the asymmetric factor is not introduced since they already include an imbalance. An example of the distribution before and after smearing for the case of ANTARES and 70% are presented in Appendix E. Note, the script "add CAN ANTARES" is the same as in the processing pipeline, since a new file is produced in the three-dimensional vector that contains the dimension of ANTARES, must be present in the file during the computation of the detector response matrix.

After the production of the necessary files, they are parsed into the χ^2 profile analysis code. The several steps that undergo before the fit process have been discussed in Section:5.1.2. From the JSON files, discussed in Section:5.1.1, the analysis is configured to run to the requested scenario. Once the models have been imported the *Detector Response Matrix*, which corresponds to the probability of an interacting event with true parameters (E_t, θ_t, y_t) to be reconstructed with reconstructed variables (E_r, θ_r, y_r) . For the case of ANTARES, not only the reconstructed inelasticity, but also the version of MC events used in the analysis doesn't provide this information. Thus both values are artificially set to 0.5. The matrix is a multi-dimensional one, with 7 dimensions in total, 3 first ones for the true parameter space (E_t, θ_t, y_t) , 3 for the reconstructed space (E_r, θ_r, y_r) and one for the 8 different possible interactions, 6 through the CC channel and two for the NC. For the scenarios in this analysis, two versions of the response matrix are created, since these many classes were defined during the configuration. The efficiency of each neutrino flavor per class is calculated, as follows:

$$R(E_t, \cos \theta_t, y_t, E_r, \cos \theta_r, y_r)_{\nu_i}^{\text{class}} = \frac{n_{\text{sel}, \nu_i}^{\text{class}}(E_t, \cos \theta_t, y_t, E_r, \cos \theta_r, y_r)}{n_{\text{gen}, \nu_i}(E_t, \cos \theta_t, y_t)} \quad (59)$$

where R is the efficiency/probability, n_{gen} the number of generated events for the given neutrino flavor and $n_{\text{sel}}^{\text{class}}$ is then number of events per flavor that passed the selection criteria of the class. From the efficiency of the response matrix the histograms of the expected events are produced and stored in the *Experiment* object. The computation of the response matrix takes around half an hour, but once these matrices are stored in an *EvtStore* file, they can be used for every analysis from now on if the conditions that make up the response matrix stay the same. These parameters are the reconstruction algorithm, the cut selection, the make-up of the classes and the event file where the pseudo data are stored. From the response matrix, it is also very convenient to export the effective volume of the detector and can serve as a quick cross-check that the computation of the detector response didn't face any problems. ANTARES's effective assumed volume, assuming perfect reconstruction, is calculated and presented in Figure:47

Once the matrix is created or loaded the fit process begins. The first fit, is to find which TauNorm value is the most probable, according to our pseudo-data and model, i.e. the lowest χ^2 and this become the χ_{free}^2 . Once the minimum is found, a scan over the TauNorm parameter is performed and the χ_{fixed} is calculated for each point. The range of values and the number of points scanned are already defined by the User.json file. For this analysis, 20 points were chosen to be fitted during the "restricted" parameter case, where only TauNorm would be open, and 14 points for the "realistic" scenario. In the latter scenario, during the χ_{fixed} calculation step, the are still parameters free to be fitted. The type of fit, performed during the analysis is called *FitTwoOctants* executing two fits for each point one per θ_{23} octant and storing the one with the worst χ^2 . Once all the points are fitted, two separate ROOT files are created, one where the free point is stored and one with all the fixed points of the scan. Depending on the scenario, each run can last from one and a half hours up to 10 hours.

Once the fit results have been stored the plotting of the sensitivities can be commenced. There are two different approaches during plotting, the individual plot and the merged plot. The individual plot runs as the last step of the pipeline and can be a visual indicator in case there is an issue in the fitting process. However, just individual sensitivities don't offer much information in this analysis and direct comparisons are needed with other scenarios, which the merged approach provides. There are three options when plotting a combined

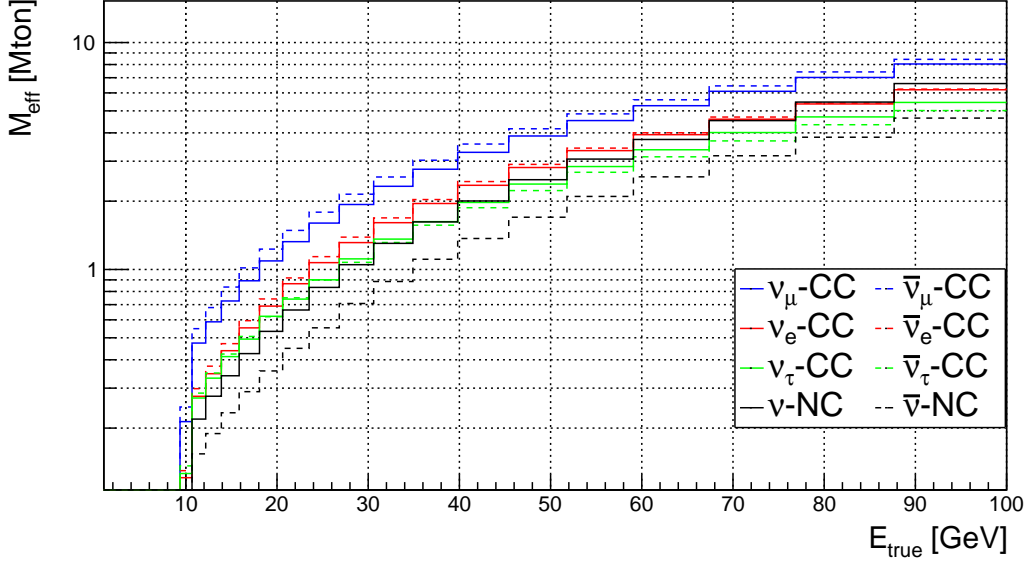


Figure 47: Effective mass volume of the ANTARES for triggered and reconstructed events.

plot, depending on the outcome that needs to be discussed. The first one, includes for the same scenario all the reconstructions for the same scenario (channel and neutrino mass ordering), with the reconstructions being *AAFit_ANN_ECAP* and *AAFit_dedX_CEA*, NNFit, NNFit direction reco only and MC-truth. In the second approach, compares the MC-truth, i.e. perfect reconstruction, with the smearing options and NNFit. The last option compares, the sensitivity gain from potential upgrades/improvements in the reconstruction algorithm, in the direction or the energy reconstruction component.

5.2 Reconstruction algorithms

In this part, the sensitivities of the available reconstruction algorithms are compared to the case of perfect reconstruction. As already mentioned, the algorithms compared are AAFit with each two energy estimators and NNFit with only the direction reconstruction and NNFit in its entirety. The comparisons, will be performed with the 3 sets of cuts discussed in the previous chapter, Section:4.4. And for each cut scenario, the sensitivities are computed twice depending on the parameters that are enabled to be freed during the fit, see more in Section:5.1.1. Before starting a reminder of the cuts that were defined gathered and presented below

- Low energy up-going
 - Tracks/Showers: $(\cos \theta_{\text{true}} < 0) \wedge (10 < E_{\text{true}} < 100)$
- NNFit up-going
 - Tracks/Showers: $\cos \theta < 0$
- NNFit quality
 - Tracks: $(\cos \theta < -0.4) \wedge (\sigma_{R_{Z\text{Closest}}} < 10) \wedge (\sigma_{R_{R\text{Closest}}} < 10) \wedge (\sigma_{\theta} < 8)$
 - Showers: $(\cos \theta < -0.4) \wedge (\sigma_{R_{Z\text{Vertex}}} < 10) \wedge (\sigma_{R_{R\text{Vertex}}} < 10)$

The unit of energy in the cuts is in GeV, the unit of distance for the vertex position or the closest point of the track is in meters, and the direction is in degrees. For the track class, the *NNFitTrack* model is used, and for showers *NNFitShower*.

5.2.1 Low Energy up-going

The first results assume that every triggered and reconstructed event is low energy and up-going. The cut is the loosest one, offering the largest of close to 1 Mt close to the energy range of interest for electron muons. The sensitivities for the "best" case scenario that ANTARES, utilizes the full detector volume with perfect classification is presented below in Figure:48.

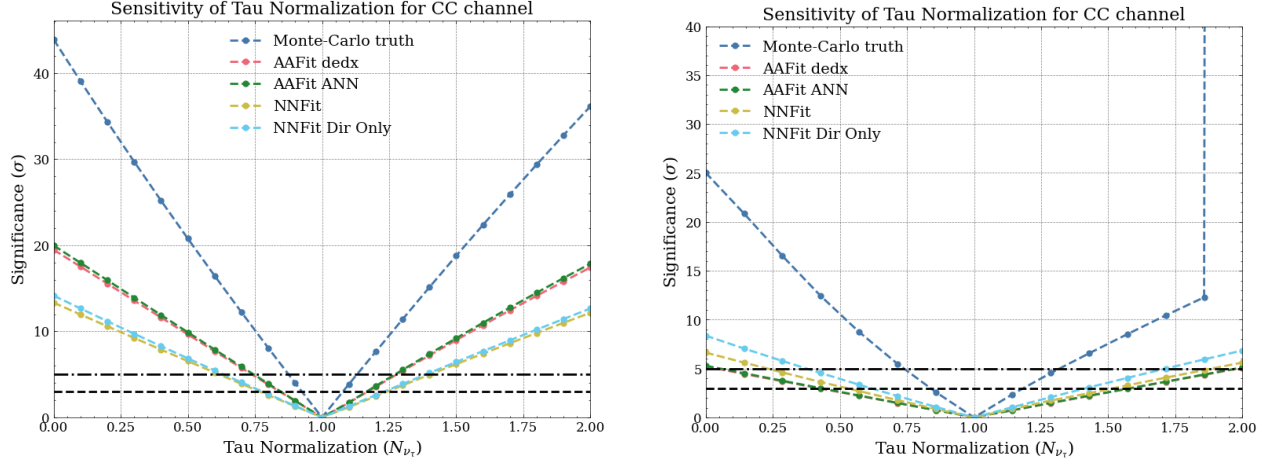


Figure 48: (left) Sensitivities for the "restricted" parameter scenario, (right) Sensitivities for the "realistic" scenario with cut selection criteria "Low energy up-going".

A very interesting phenomena arises from the sensitivity plots (Figure:48), in the case where only the TauNorm is fitted, both AAFit energy estimators perform better than NNFit, even in the case where only the direction reconstruction is used and we assume perfect energy reconstruction. This might seem counter-intuitive with the analysis of the previous section, where it was "supposedly" clear that the NNFit performed better in both topologies. But there are two things that need to be considered. Firstly, the efficiency of the two algorithms is different with AAFit having more than twice the efficiency in both topologies. If that was the only explanation, then in the case when systematic uncertainties are considered in the analysis, better performance would be again expected. However, this is not the case, in the "realistic" scenario both NNFit approaches have better performance. The explanation behind this is that when there are more parameters enabled to be fitted the resolution of the reconstruction plays a much bigger role. It was discovered that there is an effect behind this behaviour, where worse reconstruction with high variances would lose more sensitivity in the "realistic" scenario. Another clue for this behaviour will be seen in another case later on. One more intriguing fact when comparing the two AAFit reconstructions. In the "restricted" case, AAFit ANN seems to display better results, however, when the systematic uncertainties were enabled both seem to behave the same. This seems a bit puzzling given the fundamentally different implementation of the two algorithms. This behaviour is the same with the other selection criteria so at the moment is still unclear and no explanation can be offered, that would satisfy the reader and the writer. In the restricted case, the last point of the MC-truth has an "unreasonable" sensitivity. This is the influence in the fitting process from one of the parameters that were freed. To identify which parameter it was a systematic study is required by performing the same fit repeatedly with a varying number of free parameters. However, this study would be out of the scope of this analysis and would require a considerable number of CPU hours. So it wasn't followed through.

5.2.2 NNFit up-going

These sensitivities can be classified as the first that use reconstructed information for their cuts. By using NNFit, immediately only around 20% of the previous detector volume is used given the reconstruction's efficiency. Nevertheless, this is to be expected, since the phase-space (E, θ) of interest in the analysis is close to the detector threshold and no dedicated reconstruction has been developed in this regime. Thus the first sensitivities with NNFit, as the main reconstruction for cut selection are presented below

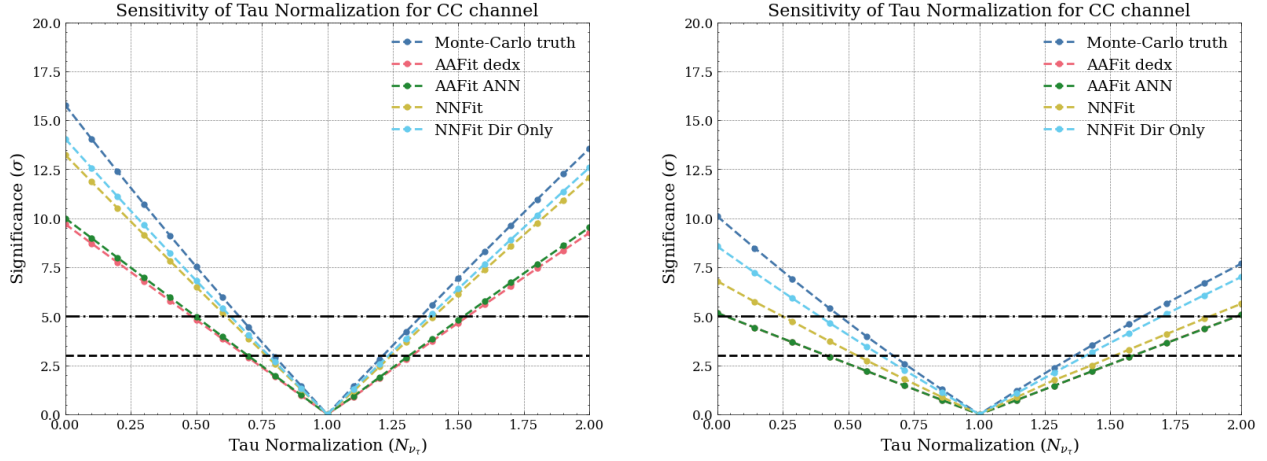


Figure 49: (left) Sensitivities for the "restricted" parameter scenario with cut selection criteria "NNFitup-going", (right) Sensitivities for the "realistic" scenario.

From Figures:49, even from the "restricted" scenario NNFit performs better with a 5σ difference. This is quite interesting and raises some questions regarding the type of events that were reconstructed by AAFit and not NNFit and gave the sensitivity in the previous cut. This comparison between algorithms is a bit skewed towards NNFit. In the AAFit sensitivity, the events need to be reconstructed as up-going by both reconstructions resulting in a considerable drop in statistics. A quick investigation had been performed for this very purpose and the loss of events was less than 10%. Comparing the sensitivities of NNFit with the two cut selections already mentioned, the performance is almost even and there is an explanation for this behaviour. The reason behind this is that the direction cut of up-going events has retained most of the events. So effectively, the only difference between these samples is true up-going events that were reconstructed as down-going and there aren't many. Hence the similarity in the performance between these two cuts. This is quite interesting and raises some questions regarding the type of events that were reconstructed by AAFit and not NNFit and gave the sensitivity in the previous cut. This comparison between algorithms is a bit skewed towards NNFit. In the AAFit sensitivity, the events need to be reconstructed as up-going by both reconstructions resulting in a considerable drop in statistics. A quick investigation had been performed for this very purpose and the loss of events was less than 10%. Comparing the sensitivities of NNFit with the two cut selections already mentioned, the performance is almost even and there is an explanation for this behaviour. The reason behind this is that the direction cut of up-going events has retained most of the events. So effectively, the only difference between these samples is true up-going events that were reconstructed as down-going and there aren't many. Hence the similarity in the performance between these two cuts.

5.2.3 NNFit quality

In this part, the sensitivities of "well"-reconstructed by NNFit events are presented, since large detector volume is not the only factor but also the quality of the reconstructed events. With the application of some quality cuts, that use NNFit reconstructed information to get rid of some outliers and strengthen the overall quality of our data. With the "stricter" quality cuts, the sensitivities of the reconstruction algorithms are presented below.

In Fig:50, there is an improvement in the sensitivity of all reconstructed algorithms, since the overall quality of the pseudo-data increases. The only outlier is the MC-truth, where there is a loss of sensitivity, due to loss of statistics. In the case of MC-truth, there is already a "perfect" reconstruction, so the quality of the sample is not a problem. Rather the main factor is the overall detector volume. When systematic uncertainties are included, there is no effective change in the sensitivities. This strengthens the fact the main gain of the sensitivity, derives from well reconstructed events. So possibly with a dedicated effort in a cut selection analysis through the more complex methods, the regions of the phase-space that give the main sensitivity can be included, for a more realistic approach in out-cut selection.

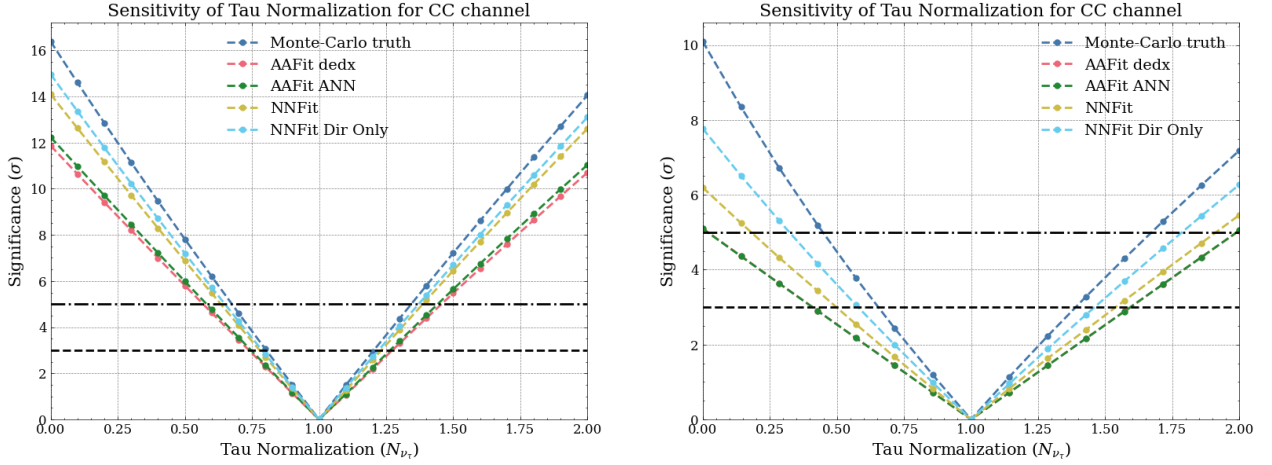


Figure 50: (left) Sensitivities for the "restricted" parameter scenario with cut selection criteria "realistic", (right) Sensitivities for the "nnfit quality" scenario.

5.3 Smeared events

In this part, the sensitivity of NNFit is tested to find to what level of smearing it equates. As mentioned already, the smear events originate from the true parameters that were smeared through a Gaussian with the width of the distribution being proportional either to a percentage value or the resolution of an algorithm. There are also some comparisons with the resolutions of other detectors such as ORCA6 and ORCA115. The comparison is made for both cuts that included reconstructed information, with the first cut the "low energy up-going" being rejected due to the difference in the detector volumes. The difference comes from the fact that the smeared branches come from the true ones, so the smeared branches have "essentially" 100% efficiency compared to the 20% of NNFit.

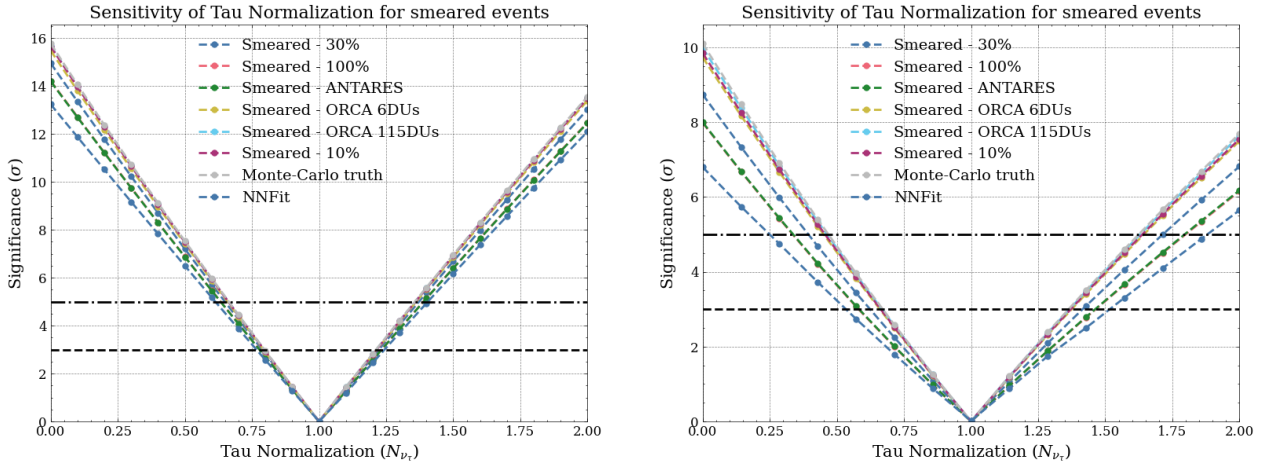


Figure 51: (Left) sensitivities of smeared event and NNFit for the "restricted" scenario and (right) Sensitivities for the "realistic" scenario with the cut selection criteria of "nnfit up-going".

The sensitivity plots for both aforementioned cuts, are presented in Figure: 51 and Figure:52. In the case of the "restricted" scenarios, events with low levels of smearing, such as both ORCA examples and the 10% smearing, are almost superimposed on each other, with an expected decrease in the sensitivity for a higher level of smearing. The ANTARES resolution has a very close performance of 90%. The 90% level smearing, which equals the 90% smearing in direction and 270% in energy, has better performance than ANTARES, confirming the expectation that the gain in the sensitivity from the energy is greater than the direction. If the

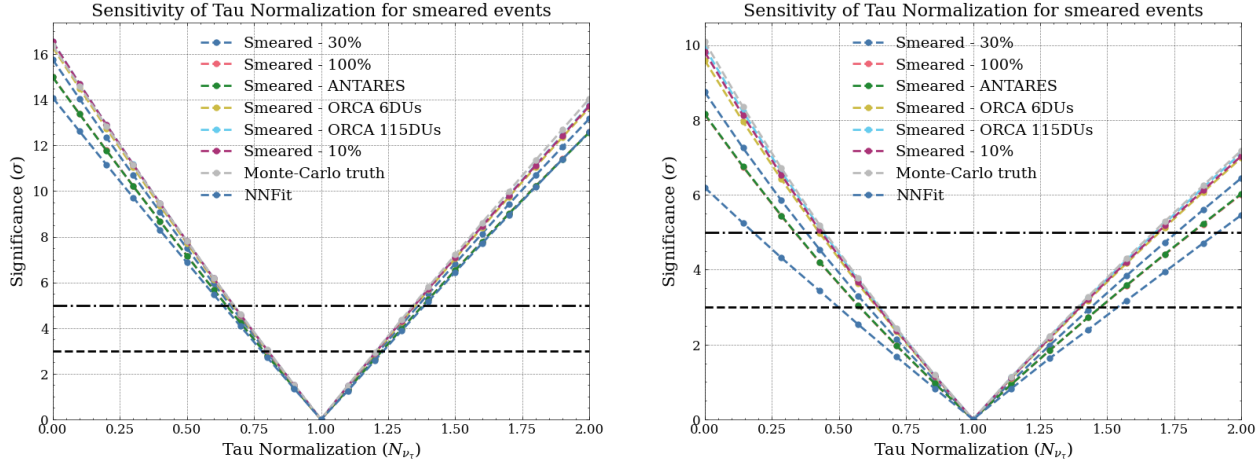


Figure 52: (Left) sensitivities of smeared event and NNFit for the "restricted" scenario and (right) Sensitivities for the "realistic" scenario with the cut selection criteria of "nnfit quality".

smeared ANTARES option is based on NNFit, then why it's not the same performance? Since this smearing option uses the resolution of the algorithm as the width of the normal distribution when sampling, someone would expect that. Well, this would be the case if NNFit, if the energy reconstruction didn't have such a huge variance, hence the difference in the sensitivities. From a degeneracy that was evident in the "restricted", it vanished when including all systematic uncertainties. There is a lack of coherence in the best-performing sensitivities now, being able to better distinguish the various scenarios.

5.4 Reconstruction upgrade

The sensitivities of the current algorithms have been presented already, with the current status being that NNFit is the best-performing for this analysis. From the study of the reconstruction performance accomplished in the previous chapter, Section:4.4, the difficulties in the energy reconstruction part were identified. So this section is dedicated to discussing the potential gain in the sensitivity from an upgraded reconstruction. As a starting base, the resolution from NNFit is given and theoretical upgrades in the energy and direction reconstruction are investigated. This analysis is very similar to the smearing in the previous section, but instead in this implementation during the "smearing" the width of the normal distribution is condensed, thus resulting in better-quality events. There is a discussion here whether to use NNFit directly or the smeared ANTARES data that use the resolution of nnfit and there is a distinct difference in these two approaches. By using the NNFit directly, the study focus shrinks only to the potential improvement of this reconstruction. Whereas in the other approach, only the performance is used, without the inherited limit in the algorithm efficiency. This might seem insignificant, but the difference is quite noticeable and it will be presented below comparing improvements with two cuts, the loose cut "low energy up-going" and "nnfit up-going". With this approach, a future algorithm that has a performance similar to NNFit can see the gain in sensitivity that this study would offer. The gain in the sensitivity would be proportional to its efficiency, so we propose the potential gain from a reconstruction with the efficiency ranges of (20%, 100%).

The algorithm improvement was split into two components, the two reconstruction tasks, energy and direction. In this study, three options were chosen,

- 40% direction improvement
- 20% direction & energy improvement
- 40% energy improvement

In Figure:53, the three cases of potential improvement are shown in comparison with the current status of the resolution of the best-performing algorithm. As expected the 40% increase in the energy provides the best gain in the sensitivity. Something that was also expected from previous discussions, regarding the influence of energy and track reconstruction in the ν_τ appearance study. Another point could be that taking into account the current status of these two reconstruction tasks, there is more room for improvement in energy reconstruction. As seen in the "restricted" case, regarding the fit process, the potential gain for a 40% increase in energy is from 1.5 up to 4.0 σ depending on the efficiency of the resolution. When comparing the other two improvement options, again it is evident that the an improvement in energy reconstruction is always better than in direction. The difference though decreases with lower statistics.

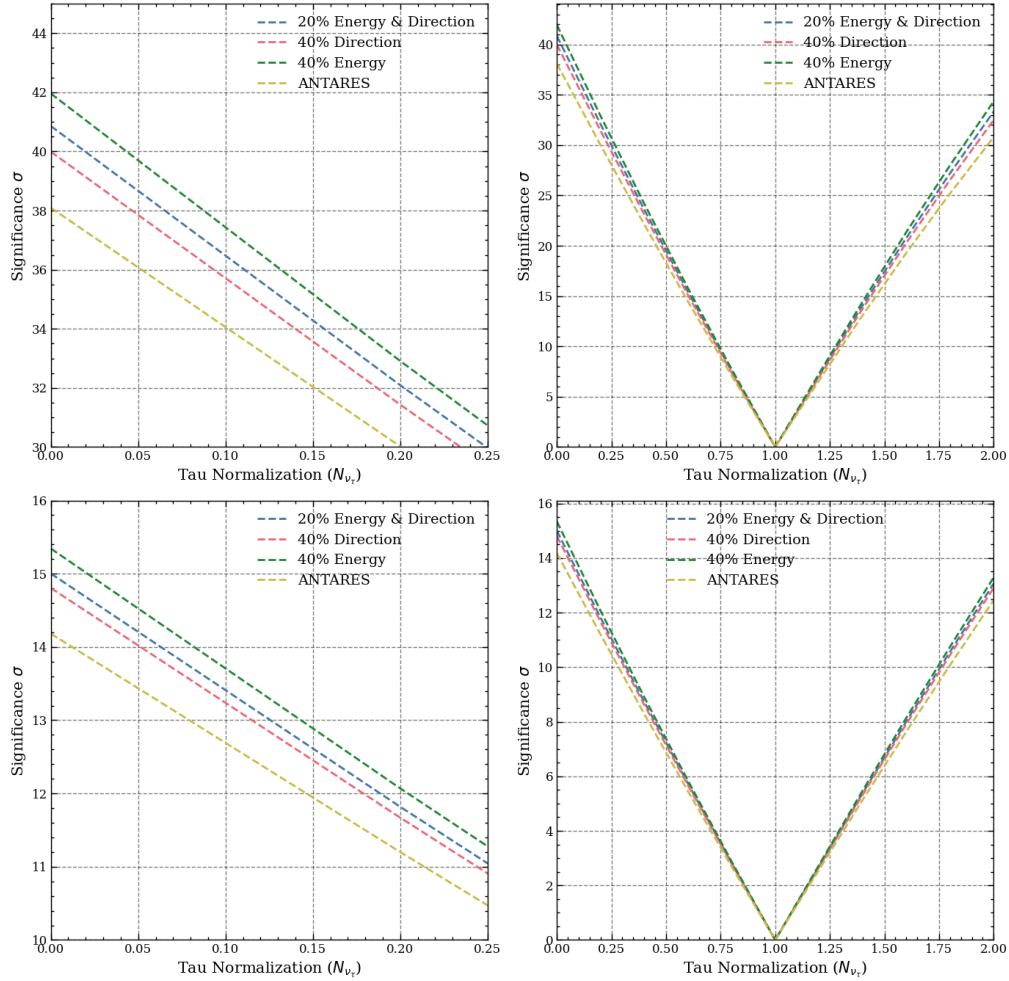


Figure 53: (Upper left) Potential algorithm improvements sensitivities with the cut selection criteria of "low energy up-going" and (upper right) zoomed version of the same plot in range (0, 0.25) of the TauNorm. (Lower left) Potential algorithm improvements sensitivities with the cut selection criteria of "nnfit up-going" and (lower right) zoomed version of the same plot in range (0, 0.25) of the TauNorm. These sensitivities are under the "restricted" scenario of systematic uncertainties.

The same investigation is carried out with the inclusion of systematic uncertainties to check if the behavior that was displayed before was influenced by the number of parameters freed. As seen in Figure:54, no unexpected behaviour appears. The best-performing "upgraded" algorithm remains the 40% energy upgrade and the 20% energy and direction upgrade being second.

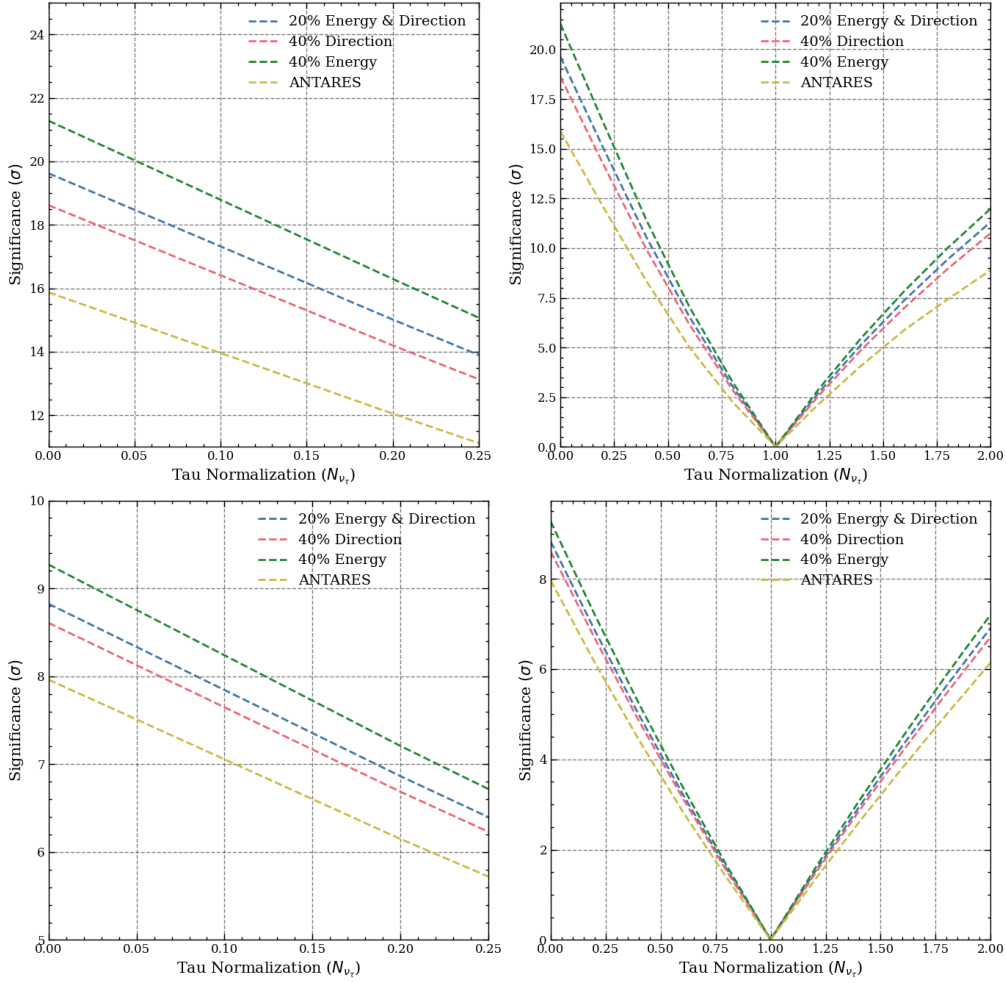


Figure 54: (Upper left) Potential algorithm improvements sensitivities with the cut selection criteria of "low energy up-going" and (upper right) zoomed version of the same plot in range (0, 0.25) of the TauNorm. (Lower left) Potential algorithm improvements sensitivities with the cut selection criteria of "nnfit up-going" and (lower right) zoomed version of the same plot in range (0, 0.25) of the TauNorm. These sensitivities are under the "realistic" scenario of systematic uncertainties.

With the 1.5 up to 4.0 σ upper limit of the sensitivity gain in a 40% reconstruction task improvement, the discussion changes to the worst performer. In a 40% direction improvement we expect sensitivity gains from 0.6 up to 2.5 σ . So any combination in track and direction that would have a combined improvement of 40% would be between these ranges. However, there should be a practical discussion, since the effort invested to reach this level might widely vary between the two tasks. Also, the upper limit in the improvement of each task is different, since they are based on different parameters. For direction reconstruction, the impressive timing resolution is already utilized. It is unclear though how more can be upgraded or the allocated time would have a return on investment. The complete opposite situation is for the energy reconstruction even with its current situation.

6 Conclusion

6.1 Summary

After an introduction on neutrino physics, the physics goal of this thesis is introduced, the ν_τ appearance. After briefly explaining the nature and the goal of such analyses, the status of current tau neutrino experiments is discussed and the most up-to-date results of on the tau sector are presented. The experiment that this analysis is based on is introduced, ANTARES. The detector design and principles of such water Cherenkov detectors are specified, with the events and backgrounds of these experiments being analyzed.

Chapter 4, is devoted entirely to MC events in ANTARES. Starting with introducing the several models used during the event simulation and production, until the final files, the AntDSTs. These files mark the entry point of this thesis, where through a dedicated pipeline these summary files are converted into a suitable format for the framework, Swim. It is the first pipeline, that provides the user with all available reconstructed information from AntDSTs and an additional reconstruction, NNFit, algorithm is added. The latter is a recently developed DL-based reconstruction that aims to reconstruct lower energy events, with dedicated models for both "shower-like" and "track-like" events, that could potentially benefit this study. With every potential information included, the search for a suitable reconstruction commences, with AAFit, BBFit and NNFit being compared. Through the investigation of the reconstruction algorithm, the difficulties of this study were given the spotlight, since none of the reconstruction algorithms has the centre of attention to the events that this study seeks. After the selection of the algorithm that provides the best performance for low-energy events, an initial collection of cut options is proposed. The expected number of events is calculated with the help of the oscillation weights, for each cut selection criteria.

Chapter 5, is assigned to provide all necessary information about ν_τ appearance sensitivities with *Swim*, the official oscillation analysis framework internally developed in the KM3NeT collaboration. A short outline of the key features and inner workings is provided to the reader, with a focus on the fit process utilized in the analysis and the possibility of user configuration through the use of JSON files. The analysis pipeline is then explained in detail with the purpose of each step clearly indicated. The first results presented are the sensitivities from the reconstruction algorithms, AAFit with two energy estimators *AAFit_dEdX* and *AAFit_ANN* and NNFit for each cut selection criteria. From the results, it was found that indeed NNFit provided the best performing sensitivity from the available reconstructions so far. In our current scenario, ANTARES when utilizing the entire 15 years of its operation provides up to 6.6σ in the non-tau scenario. This performance is then directly compared to smeared events in order to identify, in which level of smearing the resolution of NNFit is equal to. In this study, two different smearing strategies were devised for this purpose. The first one "flat-based", where the smearing is proportional to the value of the true parameter and based on a constant variable **smear level**. The second method, uses the resolutions of the best-performing reconstruction. It was revealed that NNFit in its current level provides the same performance as a smearing level of 90%. In this comparison, an estimate of the resolutions of ORCA6 and ORCA115 detector configurations was included highlighting the difference in performance compared to the reconstructions of ANTARES. Finally, the performance gain from potential work invested in the upgrade NNFit or the invention of a similarly performing algorithm is examined with a range of 0.6 up to 1.5σ depending on the reconstruction task, when including systematic uncertainties.

6.2 Outlook

An import work has been investigated in setting up a pipeline, where every possible reconstruction algorithm can be easily available. Apart from this, this pipeline extracts all triggered information, in an efficient computing cluster way and provides the maximum detector volume, before the inclusion of any analysis-specific cut selection criteria. This pipeline also provides reconstructed information from the recently developed NNFit algorithm, that could provide sensitivity to events that other algorithms lack. With the inclusion of the branch of oscillation weights, a more dedicated cut selection investigation on the tau sector in ANTARES can be further analyzed. The study includes also the first investigation for the tau appearance measurement with the ANTARES detector. In this analysis it was also possible to utilize the entire 15-year lifetime of the detector, reaching combined more than $1\text{ Mton} \cdot \text{year}$ of neutrino events in the measurement's sensitive region from 20 to 40 GeV. The sensitivity plots provide an initial look into potential sensitivity gains with

including ANTARES data in a combined analysis. Drawing comparisons with results from other experiments, can be done for the non-tau scenario only. The use of the "Asimov" dataset in our analysis, would converge the TauNorm to 1 in all cases, thus no relevant outcome can be extracted with such comparison. On the other side, this is not the case for the non-tau scenario, i.e. the TauNorm converges to 0. Even though it is not possible, we can draw conclusions from the significance that this scenario is excluded. Comparing the sensitivities produced under our assumptions, when utilizing the whole NNFit sample we can exclude the non-tau scenario of 6.6σ , which is even better than the OPERA with 4 years of data. In fact even when applying some quality cuts and getting rid of 26% of my shower sample there is a sensitivity of 6.1σ . In fact under this selection cut case the ORCA experiment under the assuming same detector constant detector volume of 6 DUs would take about 7.11 years to reach these levels probing into the same channel. This indicates that the large assumptions taken throughout this study inflate the sensitivity level provided. However, it shouldn't be underestimated the fact that there is a sample of 476 ν_τ^{CC} with a shower topology reconstructed with NNFit and assuming perfect classification and a comparable muon rejection background as ORCA would lead to a potential candidate for a combined analysis. Continuing with our assumptions, the entire lifetime of the detector has been included without taking into consideration the environmental qualities of each run. Furthermore, the treatment of systematic uncertainties is also not researched for this detector, even though there are a great deal of similarities with ORCA, some parameters are not the same such as the PMT efficiency. Thus some priors in the normalization parameters should be investigated if a combined analysis with two different detector configurations is in order. Having said that since the end of the analysis, another potential tool in this study has been developed. An initial version of a classifier developed by the Polytechnical University in Valencia that was the sole contributor of NNFit. Compared to the other version of the classifier, this is an extension of NNFit, which could provide a solution to this obstacle. Furthermore, the difficulties of the energy reconstruction have been documented, since none was optimized in this region. A future retraining of NNFit could provide another significant boost to this analysis. However, this study could try to enlarge its current ν_τ sample by not relying entirely on NNFit and using ML reconstructed by other algorithms. The potential gain of this approach hasn't been investigated rater and the quality of the data would be enough to improve our current levels of sensitivity. As a final outlook, personally, I believe given the early results and potential further steps, that there is room for improvement and time should be invested in such an endeavour. The tools for this study, that a year ago, were not present have been crucial for this study. Given that there have been other studies with lower energy events in mind, outside the main analysis topics of ANTARES, new tools are developed that benefit similar studies.

Acronyms

ANTARES "Astronomy with a Neutrino Telescope and Abyss environmental RESearch". 14, 16, 17, 18, 21, 22, 25, 30, 31, 34, 37, 39, 40, 46, 48, 49, 50, 56, 57, 58, 59, 65, 66, 83, 84, 85, 86, 87

AntDST "Antares Data Summary Tree". , 22, 23, 27, 30, 31, 65

ARCA "Astroparticle Research with Cosmics in the Abyss". 51

BSM Beyond Standard Model. 27, 28, 46, 51

CC Charged Current. , 2, 3, 7, 9, 13, 14, 19, 21, 22, 25, 26, 29, 31, 37, 46, 50, 51, 52, 57, 78, 84

CMB Cosmic Microwave Background. 4

CORSIKA COsmic Ray SIMulations for KAscade. 21

CR Cosmic Rays. 5

DAQ Data Acquisition. 18

DCN Deep Convolutional Network. 37

DIS deep inelastic scattering. 2, 14, 18, 21

DL Deep Learning. 31

DONuT "Direct observation of the nu tau". 12, 14

DU Detection Units. 26, 55, 56, 66

EAS Extensive Air Showers. 21

EC electron Capture. 20

ECAP Erlangen Centre for Astroparticle Physics. 96

FWHM Full Width at Half-Maximum. 53, 54, 56

GEHEN GENerator of High-Energy Neutrinos. 21

IMO Inverted Mass Ordering. 10, 11, 15, 53, 84

KM3NeT Cubic Kilometre Neutrino Telescope. , 14, 20, 23, 27, 40, 46, 47, 65, 84, 96

LCM Local Control Module. 18

LED Light Emitting Diode. 18

LIV Lorentz Invariance Violations. 28

MC Monte Carlo. 21, 22, 23, 26, 27, 30, 46, 47, 48, 49, 50, 53, 57, 58, 65

MDN Mixture Density Network. 37

ML Multi-Line events. 31, 36, 66

MSW Mikheyev-Smirnov-Wolfenstein. 9, 10

MUPAGE MUon GEnerator for neutrino telescopes based on PArametric formulas. 21

NC Neutral Current. 2, 3, 7, 9, 13, 14, 18, 21, 22, 25, 29, 31, 37, 50, 51, 57, 84

NMO Normal Mass Ordering. , 10, 11, 13, 15, 28, 46, 50, 53, 78, 84

NNFit Neural Network Fit. , 22, 29, 30, 31, 34, 37, 41, 42, 44, 49, 50, 53, 54, 55, 56, 58, 59, 60, 61, 62, 65, 66, 73, 78, 85, 86, 88

NSI Non-Standard Interactions. 28

OM optical module. 17, 18, 20, 21, 26, 31

OPERA "Oscillation Project with Emulsion-tRacking Apparatus". 12, 14

ORCA "Oscillation Research with Cosmics in the Abyss". 14, 46, 49, 50, 51, 53, 56, 65, 66, 78

PCA Principal Component Analysis. 37

PDF Particle Distribution Function. 32

PDG Particle Data Group. 3, 28

PMNS Pontecorvo-Maki-Nakagawa-Sakata. , 8, 11, 12, 13, 28, 29, 84

PMT photomultiplier tube. , 12, 17, 18, 19, 21, 23, 32, 56

PREM Preliminary Reference Earth Model. 28, 51

QE quasi-elastic. 2, 14, 18

RES resonance production. 2, 14

RGS Random Grid Search. 40

ROOT Rapid Object-Oriented Technology. 22, 23, 27, 30, 46, 52, 53, 57, 85

SK Super-Kamiokande. 7, 11, 12, 14, 84

SL Single-Line events. 36, 37, 42

SM Standard Model. 1, 2, 14, 27

SNO Sudbury Neutrino Observatory. 7

UHE Ultra-High-Energy. 4

A Oscillograms

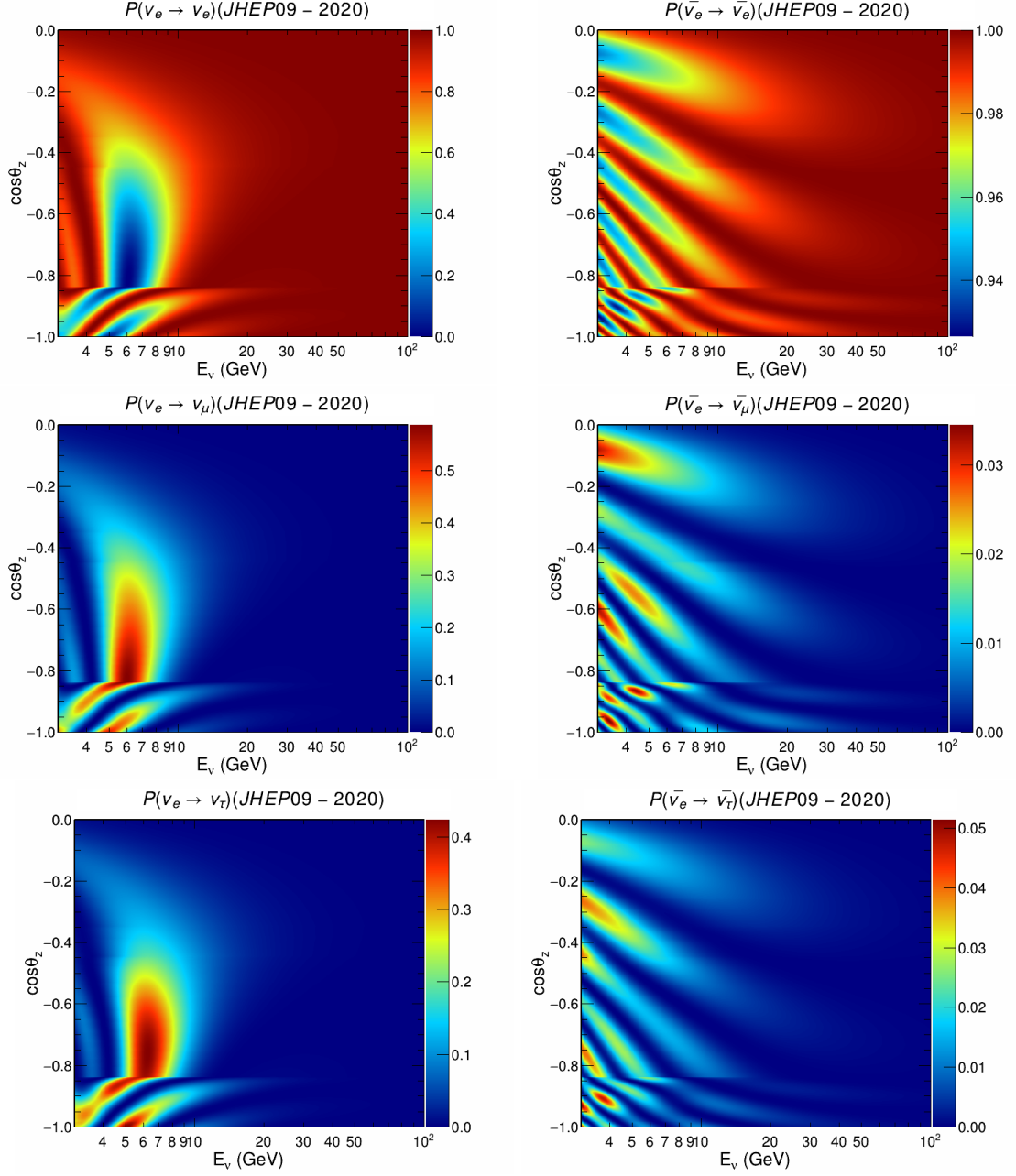


Figure 55: The two transition, $\nu_e \rightarrow \nu_\mu$ and $\nu_e \rightarrow \nu_\tau$, and the survival probabilities for the ν_e according to the, on the first column and for anti-particle probabilities on the second one. Calculated with OscProb.

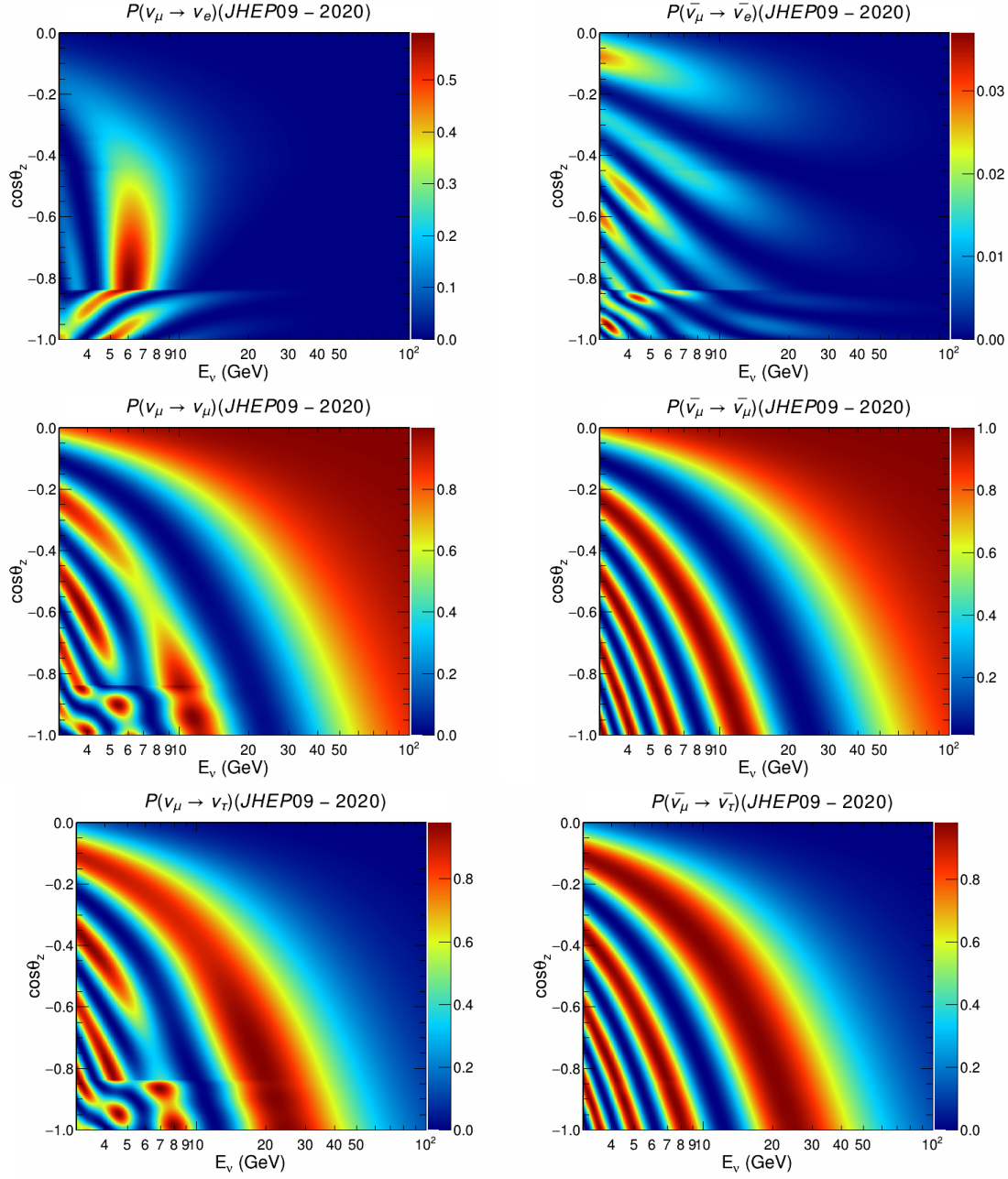


Figure 56: The two transition, $\nu_\mu \rightarrow \nu_e$ and $\nu_\mu \rightarrow \nu_\tau$, and the survival probabilities for the ν_μ according to the, on the first column and for anti-particle probabilities on the second one. Calculated with OscProb.

B ν_τ weighted distributions

The ν_τ event distributions are presented, as in the main body, with a split in topologies. A series of plots are created for each cut selection criteria, such as the "low-energy up-going" 57. For the first cut, the true event distribution is exhibited to get an insight into the events. For the other two cuts, NNFit distributions are displayed, in Figures 58 and 59.

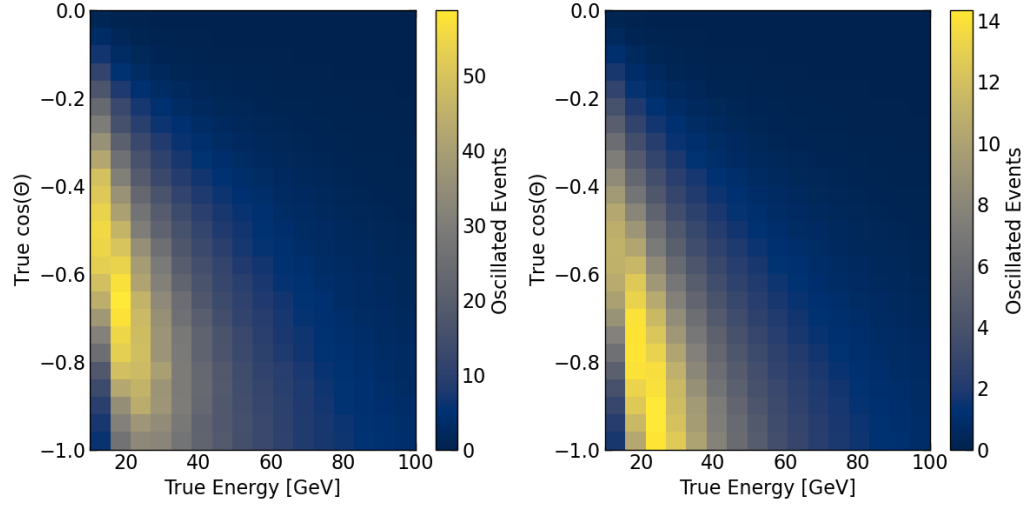


Figure 57: True ν_τ distribution with (left) "shower-like", and (right) "track-like" events.

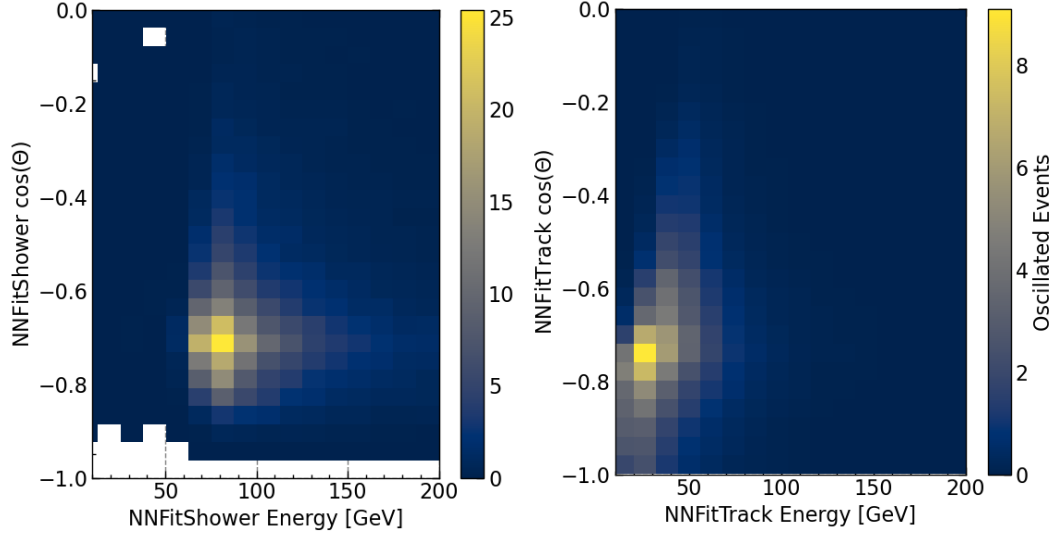


Figure 58: NNFit reconstructed neutrinos with "nnfit upgoing" cut. (Left) "shower-like" ν_τ with NNFitShower, and (right) "track-like" ν_τ with NNFitTrack.

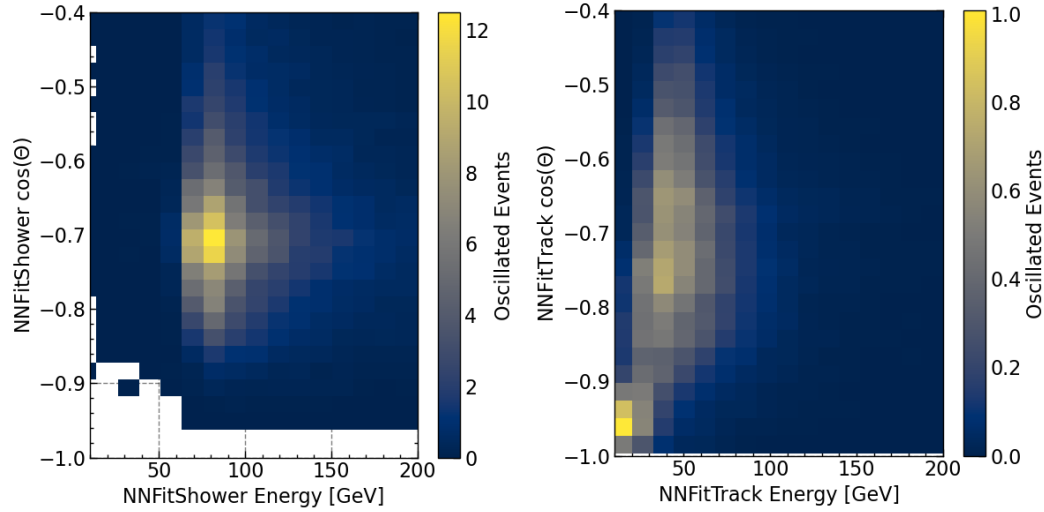


Figure 59: NNFit reconstructed neutrinos with "nnfit quality" cut. (Left) "shower-like" ν_τ with NNFitShower, and (right) "track-like" ν_τ with NNFitTrack.

C NNFit file merging

The files are gathered into 4 main classes, called *NuMuCC*, *Showers*, *Taus* and *Muons*. This merged version contains events over the ranges of runs that each file contains. This is indicated by the id of the starting and ending number runs. The following tables contain this information, being splitted into multiple columns to save space. Note: the last run listed is not contained in that file but the next one.

Table 13: Each one of the 239 merged files of the *Showers* class. The first element contains the *Run ID* of the first run contained in this run and the second row contains the element of the last run.

First	Last	First	Last	First	Last	First	Last
25800	26405	26405	26820	26820	27094	27094	27558
27558	27831	27831	28247	28247	28569	28569	28777
28780	28982	28982	29122	29123	29287	29287	29558
29559	29871	29871	30117	30118	30318	30318	30508
30508	31074	31074	31368	31368	31999	32008	32346
32346	32655	32655	33067	33067	33382	33382	33632
33632	34452	34452	34661	34661	34890	34890	35243
35243	35515	35515	35751	35751	35983	35983	36276
36276	36481	36481	36812	36812	36978	37015	37218
37220	37448	37448	37746	37746	38118	38119	38329
38329	38537	38537	38687	38687	38850	38850	39044
39044	39218	39218	39573	39573	40225	40225	40795
40795	41283	41283	41491	41491	41803	41803	42178
42178	42689	42689	43091	43091	43512	43512	43866
43866	44188	44188	45232	45232	45466	45466	45717
45717	45964	45964	46217	46217	46546	46546	46765
46765	47257	47257	47992	47992	49255	49255	49601
49601	50012	50012	50416	50416	50859	50859	51127
51127	51359	51359	51690	51690	51926	51931	52197
52197	52471	52473	52737	52737	53052	53052	53490

Continued on next page

Table 13: (Continued)

First	Last	First	Last	First	Last	First	Last
53490	53664	53664	53830	53830	54043	54043	54209
54209	54351	54351	54519	54519	54733	54733	54893
54893	55078	55078	55326	55326	55923	55923	56555
56555	56743	56743	57291	57291	57461	57461	57684
57684	57926	57926	58176	58176	58334	58334	58524
58524	58692	58692	58896	58896	59082	59082	59249
59249	59453	59456	59656	59656	59908	59908	60110
60110	60334	60334	60621	60621	60951	60953	61216
61216	61415	61415	61617	61617	61789	61789	61964
61966	62130	62132	62321	62321	62506	62506	62703
62703	62915	62915	63093	63093	63390	63390	64117
64117	64558	64558	64803	64803	65010	65010	65203
65203	65350	65350	65506	65506	65649	65649	65796
65796	65981	65981	66134	66134	66304	66304	66474
66474	66642	66642	66807	66807	66980	66980	67242
67242	67397	67397	67578	67578	67761	67761	68160
68160	68723	68723	68895	68895	69084	69084	69254
69254	69459	69459	69671	69671	70794	70794	71309
71309	71754	71754	72237	72237	72747	72747	73177
73177	73644	73644	74110	74110	74562	74562	75018
75018	75479	75479	75998	75998	76485	76485	77097
77097	77550	77556	78083	78083	78509	78509	78950
78950	79367	79367	79804	79804	80235	80240	80669
80669	81072	81072	81514	81514	81941	81941	82237
82237	82415	82415	82628	82628	82736	82736	82826
82826	82916	82916	83019	83019	83107	83107	83195
83195	83282	83282	83375	83375	83466	83466	83590
83590	83700	83700	83791	83791	83875	83876	83987
83987	84086	84086	84201	84201	84293	84293	84408
84408	84517	84518	84614	84615	84723	84723	84819
84819	84971	84971	85086	85088	85222	85222	85327
85327	85449	85449	85614	85614	85735	85735	85841
85841	85969	85970	86112	86120	86272	86273	86392
86392	86527	86528	86757	86757	86874	86874	86984
86984	87102	87102	87230	87230	87330	87330	87444
87444	87564	87564	87662	87662	87774	87774	87968
87968	88082	88087	88189	88189	88342	88342	88434
88435	88626	88626	88741	88741	89099	89099	89208
89208	89307	89307	89424	89424	89549	89549	89815
89815	89949	89949	90044	90044	90101		

Table 14: Each one of the 29 merged files of the *Muons* class. The first element contains the *Run ID* of the first run contained in this run and the second row contains the element of the last run.

First	Last	First	Last
25657	29761	29762	33325
33327	36386	36387	39204
39205	44064	44070	48512
Continued on next page			

Table 14: (Continued)

First	Last	First	Last
48514	52408	52409	54673
54675	58170	58172	60466
60468	62585	62587	66058
66060	68883	68887	72077
72077	74348	74348	76975
76983	79417	79417	82563
82566	84516	84516	85220
85221	85922	85923	86529
86529	87285	87285	88006
88006	88472	88472	89084
89084	89520	89520	89520
90007	90101		

Table 15: Each one of the 185 merged files of the *Taus* class. The first element contains the *Run ID* of the first run contained in this run and the second row contains the element of the last run.

First	Last	First	Last	First	Last	First	Last
34346	34579	34579	34813	34813	35168	35170	35472
35472	35685	35685	35904	35904	36139	36139	36385
36385	36572	36572	36893	36893	37091	37091	37292
37292	37470	37470	37765	37767	38156	38156	38402
38402	38571	38571	38722	38722	38926	38926	39072
39072	39249	39250	39638	39640	40603	40603	40922
40924	41367	41367	41586	41586	41947	41947	42500
42500	42842	42851	43215	43215	43663	43663	44003
44003	45071	45071	45373	45373	45753	45753	46418
46418	46921	46924	47997	47997	49497	49497	50365
50365	50949	50949	51230	51232	51538	51541	51988
51988	52370	52371	53028	53028	53505	53505	53764
53764	54125	54125	54384	54384	54623	54623	54869
54869	55119	55119	55703	55703	56063	56063	56605
56605	57052	57052	57396	57396	57668	57671	57935
57935	58195	58195	58385	58385	58550	58550	58721
58721	58927	58927	59118	59118	59282	59282	59550
59550	59947	59947	60224	60224	60449	60449	60730
60730	60999	60999	61387	61387	61588	61588	61777
61777	61969	61969	62150	62150	62345	62345	62533
62533	62745	62745	62939	62941	63152	63152	63413
63413	64138	64138	64611	64611	64892	64892	65047
65047	65319	65319	65472	65472	65643	65643	65801
65801	65990	65990	66195	66195	66361	66361	66532
66532	66691	66691	66853	66853	67119	67119	67310
67310	67480	67480	67651	67653	67826	67826	68585
68585	68907	68909	69095	69095	69264	69264	69468
69468	69678	69678	70809	70809	71332	71332	71763
71763	72250	72250	72751	72756	73182	73182	73633
73633	74099	74099	74539	74539	74978	74978	75432
75432	75955	75955	76429	76429	77048	77048	77507

Continued on next page

Table 15: (Continued)

First	Last	First	Last	First	Last	First	Last
77507	78021	78021	78451	78451	78876	78876	79321
79321	79753	79753	80188	80188	80615	80615	81011
81011	81456	81456	81877	81881	82206	82206	82381
82381	82586	82586	82720	82720	82806	82806	82899
82899	82984	82984	83087	83087	83179	83179	83269
83269	83360	83361	83449	83449	83547	83548	83650
83650	83740	83740	83825	83825	83927	83927	84019
84019	84132	84132	84230	84230	84340	84340	84437
84438	84546	84547	84640	84640	84797	84798	84953
84953	85115	85115	85337	85337	85547	85547	86343
86343	86470	86472	86730	86730	86865	86865	87053
87053	87177	87177	87394	87394	87977	87977	88123
88126	88259	88259	88434	88434	88707	88707	89082
89082	89195	89195	89292	89292	89409	89409	89522
89522	89626	89626	89754	89754	89912	89912	90009
90009	90101						

Table 16: Each one of the 82 merged files of the *NuMuCC* class. The first element contains the *Run ID* of the first run contained in this run and the second row contains the element of the last run.

First	Last	First	Last	First	Last	First	Last
25800	27095	27095	28280	28280	28983	28984	29561
29561	30319	30320	31675	31675	32700	32700	33638
33638	34893	34893	35753	35754	36480	36480	37211
37217	38113	38113	38671	38671	39215	39215	40788
40788	41784	41784	43073	43073	44174	44174	45707
45707	46536	46538	47949	47952	49954	49954	51104
51104	51909	51909	52712	52712	53644	53644	54194
54194	54712	54712	55230	55230	56605	56605	57502
57502	58145	58145	58656	58656	59202	59203	59770
59770	60288	60288	60803	60803	61292	61292	61835
61835	62355	62355	62949	62949	64144	64144	65030
65031	65530	65530	65999	65999	66502	66504	67004
67004	67605	67605	68745	68745	69274	69274	70857
70857	72299	72299	73708	73708	75058	75058	76529
76537	78117	78117	79402	79402	80699	80704	81974
81974	82646	82646	82922	82922	83201	83201	83476
83476	83798	83798	84096	84096	84419	84419	84732
84732	85105	85105	85466	85466	85852	85852	86283
86283	86768	86768	87110	87111	87450	87452	87780
87780	88195	88195	88631	88631	89216	89216	89549
89549	90041	90041	90101				

D JSON Files

In this section some templates for the five types of JSON files that are used during the run of the χ^2 profile fit with the *Swim* framework. The case that will be discussed here are using the NNFit reconstruction for the "NNFit Upgoing" cut scenario. For the fit, the "realistic" parameter scenario is run as always we assume NMO scenario with the same parameters freed as in the ORCA ν_τ analysis. The channel that is measured is only CC consistent with all the plots presented in the main body of the thesis, see Sec: 5.2. As for the

```

1 {
2   "classes": [
3     {
4       "name": "tracks",
5       "general_cut": "((abs(type) == 14) || ((abs(type) == 13) || (abs(type)
6       == 16 && interaction_type == 2)) && (NNFitTrack_cos_zenith < 0)",
7       "muon_loose_cut": "((abs(type) == 14) || ((abs(type) == 13) || (abs(
8       type) == 16 && interaction_type == 2)) && (NNFitTrack_cos_zenith < 0)",
9       "reconstructions": {
10        "E": "NNFitTrack_Energy",
11        "cosT": "NNFitTrack_cos_zenith",
12        "By": "NNFit_Bjorken_y"
13      },
14      "ClassNorm": "TrackNorm"
15    },
16    {
17      "name": "showers",
18      "general_cut": "(! (abs(type) == 14) || ((abs(type) == 13) || (abs(type)
19      == 16 && interaction_type == 2)) && (NNFitShower_cos_zenith < 0)",
20      "muon_loose_cut": "(! (abs(type) == 14) || ((abs(type) == 13) || (abs(
21      type) == 16 && interaction_type == 2)) && (NNFitShower_cos_zenith < 0)",
22      "reconstructions": {
23        "E": "NNFitShower_Energy",
24        "cosT": "NNFitShower_cos_zenith",
25        "By": "NNFit_Bjorken_y"
26      },
27      "ClassNorm": "ShowerNorm"
28    }
29  ]
30
31  "binning": {
32    "nEbinsTrue": 30,
33    "EminTrue": 1,
34    "EmaxTrue": 100,
35    "nEbinsReco": 20,
36    "EminReco": 10,
37    "EmaxReco": 100,
38    "ncosTbinsTrue": 40,
39    "ncosTbinsReco": 20,
40    "nBybinsTrue": 1,
41    "nBybinsReco": 1,
42    "custom": true,
43    "custom_EbinsReco": [
44      10.0,
45      11.7877,
46      13.895,
47      16.3789,
48      19.307,
49      22.7585,
50      26.827,
51      31.6228,
52      37.2759,
53      43.9397,
54      51.7947,
55      61.054,
56      71.9686,
57      84.8343,
58      100.0,

```



```

55         10000.0
56     ]
57 }
58
59 "user": {
60     "parname": "TauNorm",
61     "npoints": 15,
62     "parmin": 0.0,
63     "parmax": 2.0,
64     "type": "fixed",
65     "ordering": "NO",
66     "channel": "STD",
67     "reco": "NNFit",
68     "both_octants": true,
69     "systematics": "systematics",
70     "cut_option": "is_nnfit",
71     "is_smeared": false,
72     "smear_level": 0
73 }
74
75 "variables": {
76     "MClabel": "ANTARES",
77     "SelectedEvents_filename": [
78         "antares_w_nnfit_FINAL1.root"
79     ],
80     "exposure_nyears": 12.433,
81     "output_path": "output",
82     "flux_path": "flux/Honda2014_frj-solmin-aa_ORCA6_hist.root",
83     "by_path": "xsection/dummy_by_ORCA6.root",
84     "crossfile_InteractingEvents_path": "xsection/xsection_gsg_v5r1_SWIM.root",
85     "crossfile_RespMatrix_path": "xsection/xsection_gsg_v5r1_SWIM.root",
86     "PREMTable": "prem_default.txt",
87     "ExtensiveOutput": true,
88     "EnableMCError": true,
89     "UseW2Method": true,
90     "PlotEffMass": false,
91     "Verbose": false,
92     "SetMuons": false,
93     "EnableSmearMachine": false,
94     "Analysis_Type": "Asimov",
95     "Experiment_Type": "Std",
96     "PseudoExperiment_seed": 11,
97     "SetBootstrap": false,
98     "Bootstrap_seed": 1,
99     "Bootstrap_Fraction": 1
100 }
101
102 "parameters": {
103     "Dm21": {
104         "vData": 7.42e-05,
105         "vModel": 7.42e-05,
106         "fixed": true,
107         "prior": false,
108         "prior_mean": 7.42e-05,
109         "prior_sigma": 2.1e-06
110     },
111     "Dm31": {
112         "vData": 0.002517,
113         "vModel": 0.002517,
114         "fixed": false,
115         "prior": false,
116         "prior_mean": 0.002517,
117         "prior_sigma": 2.1e-05
118     },
119     "DeltaCP": {
120         "vData": 197.0,
121         "vModel": 197.0,
122         "fixed": true,
123         "prior": false,

```

```

124         "prior_mean": 197.0,
125         "prior_sigma": 27.0
126     },
127     "Theta13": {
128         "vData": 8.57,
129         "vModel": 8.57,
130         "fixed": true,
131         "prior": true,
132         "prior_mean": 8.57,
133         "prior_sigma": 0.12
134     },
135     "Theta12": {
136         "vData": 33.44,
137         "vModel": 33.44,
138         "fixed": true,
139         "prior": false,
140         "prior_mean": 33.44,
141         "prior_sigma": 0.77
142     },
143     "Theta23": {
144         "vData": 49.2,
145         "vModel": 49.2,
146         "fixed": false,
147         "prior": false,
148         "prior_mean": 49.2,
149         "prior_sigma": 2.0
150     },
151     "EnergyScale": {
152         "vData": 1.0,
153         "vModel": 1.0,
154         "fixed": false,
155         "prior": true,
156         "prior_mean": 1.0,
157         "prior_sigma": 0.05
158     },
159     "ZenithSlope": {
160         "vData": 0.0,
161         "vModel": 0.0,
162         "fixed": false,
163         "prior": true,
164         "prior_mean": 0.0,
165         "prior_sigma": 0.07
166     },
167     "EnergySlope": {
168         "vData": 0.0,
169         "vModel": 0.0,
170         "fixed": false,
171         "prior": true,
172         "prior_mean": 0.0,
173         "prior_sigma": 0.3
174     },
175     "NumuNumubarSkew": {
176         "vData": 0.0,
177         "vModel": 0.0,
178         "fixed": false,
179         "prior": true,
180         "prior_mean": 0.0,
181         "prior_sigma": 0.1
182     },
183     "NueNuebarSkew": {
184         "vData": 0.0,
185         "vModel": 0.0,
186         "fixed": false,
187         "prior": true,
188         "prior_mean": 0.0,
189         "prior_sigma": 0.1
190     },
191     "NumuNueSkew": {
192         "vData": 0.0,

```

```

193         "vModel": 0.0,
194         "fixed": false,
195         "prior": true,
196         "prior_mean": 0.0,
197         "prior_sigma": 0.03
198     },
199     "NCscale": {
200         "vData": 1.0,
201         "vModel": 1.0,
202         "fixed": false,
203         "prior": true,
204         "prior_mean": 1.0,
205         "prior_sigma": 0.1
206     },
207     "TauNorm": {
208         "vData": 1.0,
209         "vModel": 1.0,
210         "fixed": false,
211         "prior": true,
212         "prior_mean": 1.0,
213         "prior_sigma": 0.2
214     },
215     "MuonNorm": {
216         "vData": 1.0,
217         "vModel": 1.0,
218         "fixed": false,
219         "prior": false,
220         "prior_mean": 1.0,
221         "prior_sigma": 0.05
222     },
223     "TrackNorm": {
224         "vData": 1.0,
225         "vModel": 1.0,
226         "fixed": false,
227         "prior": false,
228         "prior_mean": 1.0,
229         "prior_sigma": 0.1
230     },
231     "ShowerNorm": {
232         "vData": 1.0,
233         "vModel": 1.0,
234         "fixed": false,
235         "prior": false,
236         "prior_mean": 1.0,
237         "prior_sigma": 0.1
238     }
239 }
240 }

```

E Smearing Examples

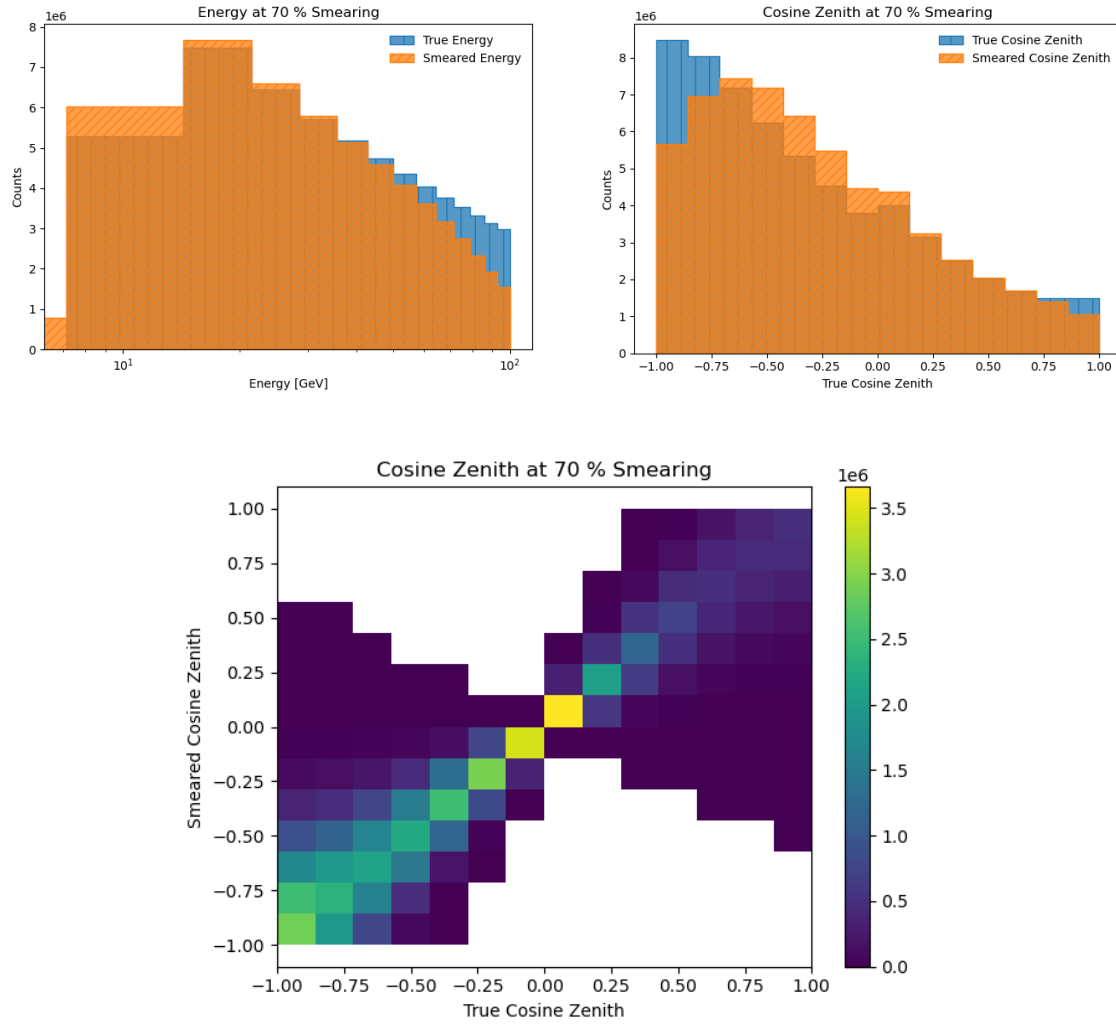


Figure 60: (Upper left) One dimensional histograms of the true and smeared energy, (upper right) one dimensional histograms of the true and smeared cosine zenith. (Lower) two dimensional histogram of true cosine zenith to smeared cosine zenith for smearing level of 70%.

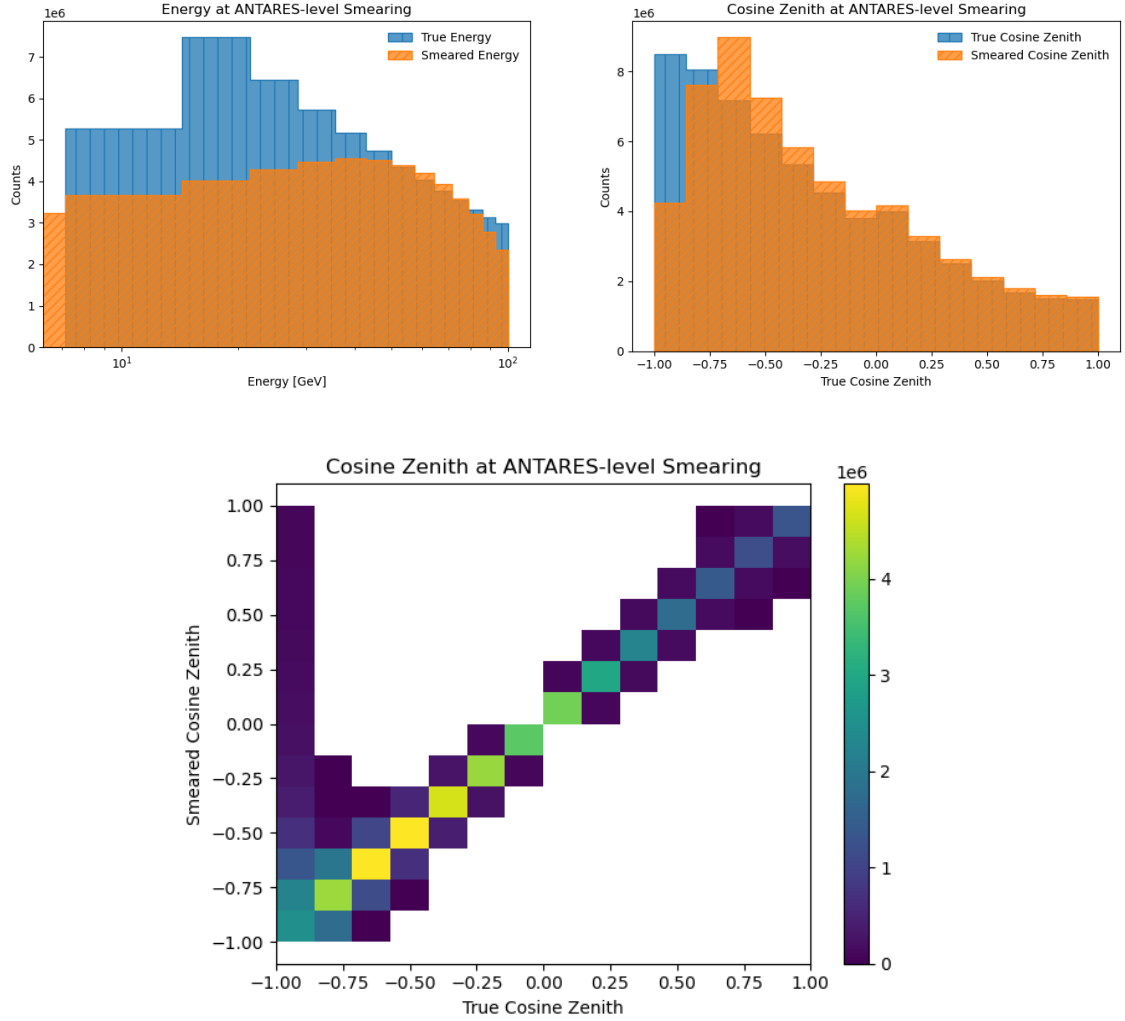


Figure 61: (Upper left) One dimensional histograms of the true and smeared energy, (upper right) one dimensional histograms of the true and smeared cosine zenith. (Lower) two dimensional histogram of true cosine zenith to smeared cosine zenith for the ANTARES smearing level.

List of Figures

1	The list of elementary particles according to the Standard Model. The colour code for the following particles corresponds to orange for quarks, blue for fermion, and green for bosons.	1
2	Neutrino cross-section for muon (anti)neutrinos. (Left) the antineutrino and (right) the neutrino cross-section as a function of energy up to 400 GeV [4].	2
4	τ branching ratios for each main decay. Taken from [7]	3
3	Illustrations of neutrino interactions for both NC and CC for deep-inelastic scattering. Taken from [5]	3
5	Measured and expected fluxes of natural and reactor neutrinos. Taken from [8]. . .	4
6	Air shower schematic. Taken from [14]	5
7	HHKM model prediction for the Frejus site averaged over an up-going direction as a function of energy. Taken from [7]	6
8	HHKM model prediction for the Frejus site for events at an energy 3.2 GeV as a function of the cosine of their zenith angle. Taken from [7].	7
9	Muon neutrino ν_μ survival and transition probabilities in vacuum in the three-flavor model as a function of the neutrino path over energy (L/E). (Left) influence of Δm_{31}^2 in the shorter ranges of (L/E), (right) effect of both mass square differences. Taken from [21]	9
10	A visual representation of the two alternative mass orderings. (Left) the NMO, where the atmospheric mass-squared is positive, and (right) the IMO, where the mass difference is negative, is presented. The mixture of the flavor eigenstates to the mass eigenstates is colorized with ν_e (red), ν_μ (blue) and ν_τ (green). Taken from [23]	10
11	Three-flavor oscillation parameters from our fit to global data as of March 2024. The recent results are presented with the left side representing the NMO and the right side the IMO. Furthermore, the lower part includes data from SK, whereas the upper part doesn't [25].	11
12	The 3σ confidence intervals of the absolute values of the PMNS matrix according to NuFit5.3 with/without the measurements of SK [26, 27]	12
13	Transition probability from $\nu_\mu \rightarrow \nu_\tau$ as a function of energy in GeV and the cosine of the direction in the case of NMO. Calculated with OscProb Software [35]. . . .	13
14	(Left) the polarization, (right) the schematic of the Cherenkov cone depending on the velocity of the particle compared to the speed of light in the medium. Taken from [45]	16
15	The attenuation length as a function of wavelength for the various sites of neutrino experiments	17
16	Illustration of the detector layout of ANTARES along with an enhancement of one storey created by Robert Lahmann.	18
17	Illustration of the ANTARES storey. Taken from [42]	19
18	3D visualization of both neutrino topologies of KM3NeT. (Left) visualization of a "track-like" event and (right) visualization of a "shower-like" event. Taken from [51].	20
19	Flowchart for the processing pipeline. The colour scheme for this flowchart is the following, light yellow for the input file, light red for the processing steps, light blue for the logic gate and light green for the final file.	24
20	Event rate in ANTARES as a function of time, for the entire operation of the detector.	25
21	(Left) the fraction of burst-like events and (right) the detector rates as a function of the data-taking month.	25
22	The atmospheric neutrino fluxes according to the Honda model. The logarithm of the flux is plotted as a function of the logarithm of energy to the cosine zenith direction.	28
23	The two transition, $\nu_\mu \rightarrow \nu_e$ and $\nu_\mu \rightarrow \nu_\tau$, and the survival probabilities for the ν_μ according to the PMNS parameters from [73]. Calculated with OscProb.	29

24	(Left) distribution of shower-like ν events and (right) the distribution of track-like ν events after applying the "low-energy true" cut for <i>AAFit ANN Energy Estimator</i> with a dashed line representing the "perfect" reconstruction.	32
25	(Left) distribution of shower-like ν events and (right) the distribution of track-like ν events after applying the "low-energy true" cut for <i>AAFit dEdX Energy Estimator</i> with a dashed line representing the "perfect" reconstruction.	33
26	(Left) direction distribution of shower-like ν events and (right) the distribution of track-like ν events.	33
27	Distribution of the AAFit quality parameter Λ , for both topologies.	34
28	(Left) Direction distributions using the BBFit "particle track" approximation for reconstructing track-like events and (right) the "bright point" approximation for shower-like event reconstruction.	36
29	(Left) Direction distributions using the BBFit "particle track" approximation for reconstructing shower-like events and (right) the "bright point" approximation for shower-like event reconstruction.	37
30	The energy reconstruction distributions for "shower-like" ν events, with (left) being the <i>NNFitShower</i> and (right) the <i>NNFitTrack</i> energy estimators.	38
31	The energy reconstruction distributions for "track-like" ν events, with (left) being the <i>NNFitShower</i> and (right) the <i>NNFitTrack</i> energy estimators.	38
32	The direction reconstruction distributions for "shower-like" ν events, with the (left) being the <i>NNFitShower</i> and on the (right) the <i>NNFitTrack</i> direction estimators lie.	39
33	The direction reconstruction distributions for "track-like" ν events, with the (left) being the <i>NNFitShower</i> and on the (right) the <i>NNFitTrack</i> direction estimators lie.	39
34	The distributions of the parameter error for both topologies, (upper left) error in the radial position, (upper right) error in the horizontal position, and (lower left) error in the zenith direction. In each distribution, a line marking a "possible" cut is marked.	40
35	Energy reconstruction performance for both shower-like and track-like events. (left) <i>NNFitShower</i> estimator for showers and (right) <i>NNFitTrack</i> estimator for tracks after the quality cuts are applied.	41
36	Direction reconstruction performance for both shower-like and track-like events. (left) <i>NNFitShower</i> estimator for showers and (right) <i>NNFitTrack</i> estimator for tracks after the quality cuts are applied.	41
37	Distribution of the cosine of the zenith direction and distribution of the error of the zenith in degrees. With black dashed lines, the proposed cuts in these two variables, are indicated.	42
38	Expected number of events for (left) "shower-like" and (right) "track-like" neutrinos ν	43
39	The expected number of NNFit reconstructed events, (left) shower-like events with <i>NNFitShower</i> and (right) track-like with <i>NNFitTrack</i> . for NNFit up-going events.	44
40	The expected number of reconstructed events in the "nnfit quality" scenario, (left) shower-like events with <i>NNFitShower</i> and (right) track-like with <i>NNFitTrack</i>	45
41	The interactions of each class of Swim are indicated within this flowchart. The lighter colour in the background point to User input. Taken from [78]	47
42	Under the assumption of perfect classification and reconstruction the theoretical all being "unrealistic" upper limit in the ANTARES sensitivity by freeing only the TauNorm parameter. (Left) the χ^2 profile is given and on the (right) the confidence level plot.	49
43	Flowchart of the analysis pipeline. The yellow refers to the files or variable definitions that act as input in the pipeline. The blue oval is a logic gate, where depending if I want to smear or not my data different paths are chosen. With red the various scripts used in the pipeline are defined and by green the files produced during the whole duration of the pipeline those being JSON files, smeared events in a ROOT format or the sensitivity plots.	52

44	Distributions of the smeared parameters for a chosen true value with a smearing level of 100% . (Right) the distribution of smeared energy values for events with true energy 50 GeV and (left) the distribution of smeared cosine zenith values for events with true cosine zenith of -0.5.	54
45	Distributions of the smeared cosine zenith for events from the horizon. (Right) distribution without the minimum threshold for an error, (left) distribution after the inclusion of the correction.	54
46	Resolutions of NNFit as a function of true energy and cosine zenith. (Upper left) direction resolution as a function of true energy, (lower left) direction resolution as a function of true direction with the hyperbolic approximation, (upper right) energy resolution as a function of energy with fit approximation and (lower right) energy resolution as a function of true direction.	55
47	Effective mass volume of the ANTARES for triggered and reconstructed events. . .	58
48	(left) Sensitivities for the "restricted" parameter scenario, (right) Sensitivities for the "realistic" scenario with cut selection criteria "Low energy up-going".	59
49	(left) Sensitivities for the "restricted" parameter scenario with cut selection criteria "NNFit up-going", (right) Sensitivities for the "realistic" scenario.	60
50	(left) Sensitivities for the "restricted" parameter scenario with cut selection criteria "realistic", (right) Sensitivities for the "nnfit quality" scenario.	61
51	(Left) sensitivities of smeared event and NNFit for the "restricted" scenario and (right) Sensitivities for the "realistic" scenario with the cut selection criteria of "nnfit up-going".	61
52	(Left) sensitivities of smeared event and NNFit for the "restricted" scenario and (right) Sensitivities for the "realistic" scenario with the cut selection criteria of "nnfit quality".	62
53	(Upper left) Potential algorithm improvements sensitivities with the cut selection criteria of "low energy up-going" and (upper right) zoomed version of the same plot in range (0, 0.25) of the TauNorm. (Lower left) Potential algorithm improvements sensitivities with the cut selection criteria of "nnfit up-going" and (lower right) zoomed version of the same plot in range (0, 0.25) of the TauNorm. These sensitivities are under the "restricted" scenario of systematic uncertainties.	63
54	(Upper left) Potential algorithm improvements sensitivities with the cut selection criteria of "low energy up-going" and (upper right) zoomed version of the same plot in range (0, 0.25) of the TauNorm. (Lower left) Potential algorithm improvements sensitivities with the cut selection criteria of "nnfit up-going" and (lower right) zoomed version of the same plot in range (0, 0.25) of the TauNorm. These sensitivities are under the "realistic" scenario of systematic uncertainties.	64
55	The two transition, $\nu_e \rightarrow \nu_\mu$ and $\nu_e \rightarrow \nu_\tau$, and the survival probabilities for the ν_e according to the, on the first column and for anti-particle probabilities on the second one. Calculated with OscProb.	71
56	The two transition, $\nu_\mu \rightarrow \nu_e$ and $\nu_\mu \rightarrow \nu_\tau$, and the survival probabilities for the ν_μ according to the, on the first column and for anti-particle probabilities on the second one. Calculated with OscProb.	72
57	True ν_τ distribution with (left) "shower-like", and (right) "track-like" events. . . .	73
58	NNFit reconstructed neutrinos with "nnfit upgoing" cut. (Left) "shower-like" ν_τ with NNFitShower, and (right) "track-like" ν_τ with NNFitTrack.	73
59	NNFit reconstructed neutrinos with "nnfit quality" cut. (Left) "shower-like" ν_τ with NNFitShower, and (right) "track-like" ν_τ with NNFitTrack.	74
60	(Upper left) One dimensional histograms of the true and smeared energy, (upper right) one dimensional histograms of the true and smeared cosine zenith. (Lower) two dimensional histogram of true cosine zenith to smeared cosine zenith for smearing level of 70%.	82

61	(Upper left) One dimensional histograms of the true and smeared energy, (upper right) one dimensional histograms of the true and smeared cosine zenith. (Lower) two dimensional histogram of true cosine zenith to smeared cosine zenith for the ANTARES smearing level.	83
----	--	----

List of Tables

1	Summary of the up-to-date results of the tau appearance measurements from the collaborations.	15
2	AAFit algorithm efficiency for low-energy up-going, well reconstructed by AAFit and reconstructed as up-going by both NNFit and AAFit.	34
3	Number of minimum hits required for a BBFit fit.	36
4	BBFit algorithm efficiency for low-energy up-going, and reconstructed as up-going by both NNFit and BBFit.	37
5	NNFit algorithm efficiency for low-energy up-going, NNFit reconstructed as up-going and after applying some selection cuts.	42
6	Expected number of events this low-energy , with percentages.	43
7	Expected number of events for NNFit reconstructed events, with percentages. . . .	44
8	Expected number of events for the "nnfit quality" scenario, with percentages. . . .	45
9	All systematic uncertainties and their treatment in the fit for the "restricted" scenario.	51
10	All systematic uncertainties and their treatment in the fit for the "realistic" scenario.	51
11	Parameterization of the energy resolution	56
12	Parameterization of the direction resolution	56
13	Each one of the 239 merged files of the <i>Showers</i> class. The first element contains the <i>Run ID</i> of the first run contained in this run and the second row contains the element of the last run.	74
14	Each one of the 29 merged files of the <i>Muons</i> class. The first element contains the <i>Run ID</i> of the first run contained in this run and the second row contains the element of the last run.	75
15	Each one of the 185 merged files of the <i>Taus</i> class. The first element contains the <i>Run ID</i> of the first run contained in this run and the second row contains the element of the last run.	76
16	Each one of the 82 merged files of the <i>NuMuCC</i> class. The first element contains the <i>Run ID</i> of the first run contained in this run and the second row contains the element of the last run.	77

Bibliography

- [1] Wolfgang Pauli. „Pauli exclusion principle“. In: *Naturwiss* 12 (1924), p. 741.
- [2] Joachim Wolf, Katrin Collaboration, et al. „The KATRIN neutrino mass experiment“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 623.1 (2010), pp. 442–444.
- [3] Max Aker et al. „Improved upper limit on the neutrino mass from a direct kinematic method by KATRIN“. In: *Physical review letters* 123.22 (2019), p. 221802.
- [4] Joseph A Formaggio and Geralyn P Zeller. „From eV to EeV: Neutrino cross sections across energy scales“. In: *Reviews of Modern Physics* 84.3 (2012), p. 1307.
- [5] Nicole Geisselbrecht. „Event Classification and Energy Reconstruction for ANTARES using Convolutional Neural Networks“. Master thesis. FAU, 2021.
- [6] Particle Data Group et al. „Review of Particle Physics“. In: *Progress of Theoretical and Experimental Physics* 2022.8 (Aug. 2022), p. 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptac097. eprint: <https://academic.oup.com/ptep/article-pdf/2022/8/083C01/49175539/ptac097.pdf>. URL: <https://doi.org/10.1093/ptep/ptac097>.
- [7] Steffen Hallmann. „Sensitivity to atmospheric tau-neutrino appearance and all-flavour search for neutrinos from the Fermi Bubbles with the deep-sea telescopes KM3NeT/ORCA and ANTARES“. PhD thesis. FAU, 2021.
- [8] Christian Spiering. „Towards high-energy neutrino astronomy: A historical review“. In: *The European Physical Journal H* 37.3 (July 2012), pp. 515–565. ISSN: 2102-6467. DOI: 10.1140/epjh/e2012-30014-2. URL: <http://dx.doi.org/10.1140/epjh/e2012-30014-2>.
- [9] Aleksandr Dmitrievič Dolgov. „Neutrinos in cosmology“. In: *Physics Reports* 370.4-5 (2002), pp. 333–535.
- [10] Hans Albrecht Bethe. „Possible explanation of the solar-neutrino puzzle“. In: *Solar Neutrinos*. CRC Press, 2018, pp. 324–327.
- [11] R. Aloisio et al. „Cosmogenic neutrinos and ultra-high energy cosmic ray models“. In: *Journal of Cosmology and Astroparticle Physics* 2015.10 (Oct. 2015), pp. 006–006. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2015/10/006. URL: <http://dx.doi.org/10.1088/1475-7516/2015/10/006>.
- [12] Kenneth Greisen. „End to the cosmic-ray spectrum?“ In: *Physical Review Letters* 16.17 (1966), p. 748.
- [13] Carlo Giunti and Chung W Kim. *Fundamentals of neutrino physics and astrophysics*. Oxford university press, 2007.
- [14] Francesco Blanco, P. Rocca, and Francesco Riggi. „Cosmic rays with portable Geiger counters: From sea level to airplane cruise altitudes“. In: *EUROPEAN JOURNAL OF PHYSICS Eur. J. Phys* 30 (July 2009), pp. 685–695. DOI: 10.1088/0143-0807/30/4/003.
- [15] Morihiro Honda et al. „Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data“. In: *Physical Review D—Particles, Fields, Gravitation, and Cosmology* 75.4 (2007), p. 043006.

- [16] Rikard Enberg, Mary Hall Reno, and Ina Sarcevic. „Prompt neutrino fluxes from atmospheric charm“. In: *Phys. Rev. D* 78 (4 Aug. 2008), p. 043005. DOI: 10.1103/PhysRevD.78.043005. URL: <https://link.aps.org/doi/10.1103/PhysRevD.78.043005>.
- [17] Takaaki Kajita. „Atmospheric neutrinos and discovery of neutrino oscillations“. In: *Proceedings of the Japan Academy, Series B* 86.4 (2010), pp. 303–321.
- [18] Yoshiyuki Fukuda et al. „Evidence for oscillation of atmospheric neutrinos“. In: *Physical review letters* 81.8 (1998), p. 1562.
- [19] Raymond Davis. „A review of the Homestake solar neutrino experiment“. In: *Progress in Particle and Nuclear Physics* 32 (1994), pp. 13–32.
- [20] Nobel Committee. *The Nobel Prize in Physics 2015*. <https://www.nobelprize.org/prizes/physics/2015/summary/>.
- [21] Victor Cattero. „Neutrino oscillations and invisible decay with the KM3NeT/ORCA detector“. PhD thesis. University of Valencia, 2024.
- [22] S. Gariazzo et al. „Neutrino masses and their ordering: global data, priors and models“. In: *Journal of Cosmology and Astroparticle Physics* 2018.03 (Mar. 2018), p. 011. DOI: 10.1088/1475-7516/2018/03/011. URL: <https://dx.doi.org/10.1088/1475-7516/2018/03/011>.
- [23] Sunny Vagnozzi. „Weigh them all! - Cosmological searches for the neutrino mass scale and mass ordering“. PhD thesis. June 2019.
- [24] F. Vannucci. „Interactions of neutrinos with matter“. In: *Progress in Particle and Nuclear Physics* 95 (2017), pp. 1–47. ISSN: 0146-6410. DOI: <https://doi.org/10.1016/j.ppnp.2017.03.003>. URL: <https://www.sciencedirect.com/science/article/pii/S0146641017300340>.
- [25] *NuFIT 5.2*. <https://www.nu-fit.org>. 2023.
- [26] Maria Concepcion Gonzalez-Garcia, Michele Maltoni, and Thomas Schwetz. „NuFIT: Three-Flavour Global Analyses of Neutrino Oscillation Experiments“. In: *Universe* 7.12 (Nov. 2021), p. 459. ISSN: 2218-1997. DOI: 10.3390/universe7120459. URL: <http://dx.doi.org/10.3390/universe7120459>.
- [27] Ivan Esteban et al. „The fate of hints: updated global analysis of three-flavor neutrino oscillations“. In: *Journal of High Energy Physics* 2020.9 (Sept. 2020), p. 178. ISSN: 1029-8479. DOI: 10.1007/JHEP09(2020)178. URL: [https://doi.org/10.1007/JHEP09\(2020\)178](https://doi.org/10.1007/JHEP09(2020)178).
- [28] Stephen Parke and Mark Ross-Lonergan. „Unitarity and the three flavor neutrino mixing matrix“. In: *Phys. Rev. D* 93 (11 June 2016), p. 113009. DOI: 10.1103/PhysRevD.93.113009. URL: <https://link.aps.org/doi/10.1103/PhysRevD.93.113009>.
- [29] M. G. Aartsen et al. „Measurement of atmospheric tau neutrino appearance with IceCube DeepCore“. In: *Physical Review D* 99.3 (Feb. 2019). ISSN: 2470-0029. DOI: 10.1103/physrevd.99.032007. URL: <http://dx.doi.org/10.1103/PhysRevD.99.032007>.
- [30] N Agafonova et al. „Discovery of τ Neutrino Appearance in the CNGS Neutrino Beam with the OPERA Experiment“. In: *Physical review letters* 115.12 (2015), p. 121802.
- [31] N Agafonova et al. „Final Results of the OPERA Experiment on $\nu \tau$ Appearance in the CNGS Neutrino Beam“. In: *Physical review letters* 120.21 (2018), p. 211801.
- [32] K. Kodama et al. „Observation of tau neutrino interactions“. In: *Phys. Lett. B* 504 (2001), pp. 218–224. DOI: 10.1016/S0370-2693(01)00307-0. arXiv: [hep-ex/0012035](https://arxiv.org/abs/hep-ex/0012035).

- [33] N. Agafonova et al. „Discovery of τ Neutrino Appearance in the CNGS Neutrino Beam with the OPERA Experiment“. In: *Physical Review Letters* 115.12 (Sept. 2015). ISSN: 1079-7114. DOI: 10.1103/physrevlett.115.121802. URL: <http://dx.doi.org/10.1103/PhysRevLett.115.121802>.
- [34] K Kodama et al. „Observation of tau neutrino interactions“. In: *Physics Letters B* 504.3 (2001), pp. 218–224.
- [35] Joao Coelho et al. *joaoabcoelho/OscProb: v2.1.1*. Version v2.1.1. Feb. 2024. DOI: 10.5281/zenodo.10698607. URL: <https://doi.org/10.5281/zenodo.10698607>.
- [36] J Devan et al. „Measurements of the inclusive neutrino and antineutrino charged current cross sections in MINERvA using the low- ν flux method“. In: *Physical Review D* 94.11 (2016), p. 112007.
- [37] Teppei Katori and Marco Martini. „Neutrino–nucleus cross sections for oscillation experiments“. In: *Journal of Physics G: Nuclear and Particle Physics* 45.1 (2017), p. 013001.
- [38] Michele Gallinaro. „Lepton universality tests and searches for charged lepton flavor violation with the CMS experiment“. In: *PoS NuFact2021* (2021).
- [39] Maitrayee Mandal. „Tau Neutrino Appearance in the Flux of Atmospheric Neutrinos at the Super-Kamiokande Experiment“. In: *PoS ICHEP2022* (2022), p. 1074. DOI: 10.22323/1.414.1074.
- [40] S. Fukuda et al. „The Super-Kamiokande detector“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 501.2 (2003), pp. 418–462. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/S0168-9002\(03\)00425-X](https://doi.org/10.1016/S0168-9002(03)00425-X). URL: <https://www.sciencedirect.com/science/article/pii/S016890020300425X>.
- [41] Aya Ishihara. *The IceCube Upgrade – Design and Science Goals*. 2019. arXiv: 1908.09441 [astro-ph.HE]. URL: <https://arxiv.org/abs/1908.09441>.
- [42] ANTARES Collaboration. „ANTARES: The first undersea neutrino telescope“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 656.1 (Nov. 2011), pp. 11–38. ISSN: 0168-9002. DOI: 10.1016/j.nima.2011.06.103. URL: <http://dx.doi.org/10.1016/j.nima.2011.06.103>.
- [43] Silvia Adrián-Martínez et al. „Measurement of the atmospheric $\nu \mu$ energy spectrum from 100 GeV to 200 TeV with the ANTARES telescope“. In: *The European Physical Journal C* 73 (2013), pp. 1–12.
- [44] Silvia Adrián-Martínez et al. „Limits on dark matter annihilation in the sun using the ANTARES neutrino telescope“. In: *Physics Letters B* 759 (2016), pp. 69–74.
- [45] József Kónya and Noémi M. Nagy. „Chapter 5 - Interaction of Radiation With Matter“. In: *Nuclear and Radiochemistry (Second Edition)*. Ed. by József Kónya and Noémi M. Nagy. Second Edition. Elsevier, 2018, pp. 85–131. ISBN: 978-0-12-813643-0. DOI: <https://doi.org/10.1016/B978-0-12-813643-0.00005-6>. URL: <https://www.sciencedirect.com/science/article/pii/B9780128136430000056>.
- [46] J.A. Aguilar et al. „Transmission of light in deep sea water at the site of the Antares neutrino telescope“. In: *Astroparticle Physics* 23.1 (Feb. 2005), pp. 131–155. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2004.11.006. URL: <http://dx.doi.org/10.1016/j.astropartphys.2004.11.006>.

- [47] Carolin Richardt et al. „Acoustic particle detection with the ANTARES detector“. In: *EURASIP Journal on Advances in Signal Processing* 2010 (Dec. 2010), p. 4. DOI: 10.1155/2010/970581.
- [48] J.A. Aguilar et al. „First results of the Instrumentation Line for the deep-sea ANTARES neutrino telescope“. In: *Astroparticle Physics* 26.4–5 (Nov. 2006), pp. 314–324. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2006.07.004. URL: <http://dx.doi.org/10.1016/j.astropartphys.2006.07.004>.
- [49] Hamamatsu. <https://www.hamamatsu.com/jp/en.html>.
- [50] Anatoli Fedynitch. „Cascade equations and hadronic interactions at very high energies“. PhD thesis. KIT, Karlsruhe, Dept. Phys., 2015.
- [51] KM3NeT Collaboration. *Neutrino topologies*. <https://www.km3net.org/pictures-and-videos/picture-galleries/>.
- [52] Steven H.D. Haddock, Mark A. Moline, and James F. Case. „Bioluminescence in the Sea“. In: *Annual Review of Marine Science* 2. Volume 2, 2010 (2010), pp. 443–493. ISSN: 1941-0611. DOI: <https://doi.org/10.1146/annurev-marine-120308-081028>. URL: <https://www.annualreviews.org/content/journals/10.1146/annurev-marine-120308-081028>.
- [53] Janik Hofestädt. „Measuring the neutrino mass hierarchy with the future KM3NeT/ORCA detector“. PhD thesis. FAU, 2017.
- [54] Jessica Craig et al. „Seasonal variation of deep-sea bioluminescence in the Ionian Sea“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 626-627 (2011), S115–S117. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/j.nima.2010.04.074>. URL: <https://www.sciencedirect.com/science/article/pii/S0168900210009289>.
- [55] A. Albert et al. „Monte Carlo simulations for the ANTARES underwater neutrino telescope“. In: *Journal of Cosmology and Astroparticle Physics* 2021.01 (Jan. 2021), p. 064. DOI: 10.1088/1475-7516/2021/01/064. URL: <https://dx.doi.org/10.1088/1475-7516/2021/01/064>.
- [56] G. De Bonis. „GENHEN release v7r1“. ANTARES-SOFT-2014-001. 2014.
- [57] I. Alikhanov. „The Glashow resonance in neutrino–photon scattering“. In: *Physics Letters B* 741 (2015), pp. 295–300. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2014.12.056>. URL: <https://www.sciencedirect.com/science/article/pii/S037026931400937X>.
- [58] A. Bueno and A. Gascón. „CORSIKA implementation of heavy quark production and propagation in extensive air showers“. In: *Computer Physics Communications* 185.2 (2014), pp. 638–650. ISSN: 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2013.09.025>. URL: <https://www.sciencedirect.com/science/article/pii/S0010465513003299>.
- [59] G. Carminati et al. „Atmospheric MUons from PArametric formulas: a fast GEnerator for neutrino telescopes (MUPAGE)“. In: *Computer Physics Communications* 179.12 (2008), pp. 915–923. ISSN: 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2008.07.014>. URL: <https://www.sciencedirect.com/science/article/pii/S001046550800266X>.
- [60] M. Ambrosio et al. „The MACRO detector at Gran Sasso“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 486.3 (2002), pp. 663–707. ISSN: 0168-9002. DOI: <https://doi.org/10.1016/>

- S0168-9002(01)02169-6. URL: <https://www.sciencedirect.com/science/article/pii/S0168900201021696>.
- [61] A GEANT. „detector description and simulation tool, 1993. Application Software Group“. In: *Computing and Networks Division, CERN, Geneva* (1993).
 - [62] J.A. Aguilar et al. „The data acquisition system for the ANTARES neutrino telescope“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 570.1 (Jan. 2007), pp. 107–116. ISSN: 0168-9002. DOI: 10.1016/j.nima.2006.09.098. URL: <http://dx.doi.org/10.1016/j.nima.2006.09.098>.
 - [63] G. Vannoni F. Schüssler K. Pazet. „Data Summary Trees for Reconstructed and Simulated Antares Events“. ANTARES Internal Notes.
 - [64] Rene Brun and Fons Rademakers. „ROOT — An object oriented data analysis framework“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997). New Computing Techniques in Physics Research V, pp. 81–86. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X). URL: <https://www.sciencedirect.com/science/article/pii/S016890029700048X>.
 - [65] J. Brunner. „The BBFit reconstruction algorithm.“ Antares Internal Note, ANTARES-SOFT-2009-012. 2009.
 - [66] C. Kopper T. Eberl. „The SeaTray software framework.“ Antares Internal Note, ANTARES-SOFT-2009-013. 2009.
 - [67] Y. Becherini et al. „Offline reconstruction and analysis software v1r6.“ Antares Internal Note, ANTARES-SOFT-2008-008. 2008.
 - [68] Juan Garcia-Mendez, Salva Ardid, and Miguel Ardid. „Dark matter search towards the Sun using Machine Learning reconstructions of single-line events in ANTARES“. In: *PoS ICRC2023* (2023), p. 1443. DOI: 10.22323/1.444.1443.
 - [69] KM3NeT Collaboration. *SWIM Software*. <https://git.km3net.de/oscillation/Swim>.
 - [70] M Honda et al. „Atmospheric neutrino flux calculation using the NRLMSISE-00 atmospheric model“. In: *Physical Review D* 92.2 (2015), p. 023004.
 - [71] Adam M Dziewonski and Don L Anderson. „Preliminary reference Earth model“. In: *Physics of the earth and planetary interiors* 25.4 (1981), pp. 297–356.
 - [72] JOACHIM KOPP. „EFFICIENT NUMERICAL DIAGONALIZATION OF HERMITIAN 3×3 MATRICES“. In: *International Journal of Modern Physics C* 19.03 (2008), pp. 523–548. DOI: 10.1142/S0129183108012303. eprint: <https://doi.org/10.1142/S0129183108012303>. URL: <https://doi.org/10.1142/S0129183108012303>.
 - [73] Ivan Esteban et al. „The fate of hints: updated global analysis of three-flavor neutrino oscillations“. In: *Journal of High Energy Physics* 2020.9 (2020), pp. 1–22.
 - [74] Aart Heijboer. „Track reconstruction and point source searches with ANTARES“. PhD thesis. Amsterdam U., 2004.
 - [75] Giulia Illuminati. „Searches for cosmic neutrino sources with ANTARES, KM3NeT and IceCube and time calibration of ANTARES“. PhD thesis. U. Valencia (main), 2020.
 - [76] Alfonso Lazo Pedrajas. „TRACK RECONSTRUCTION ALGORITHMS IN THE ANTARES AND KM3NET NEUTRINO TELESCOPES“. Master thesis. IFIC, 2021.

- [77] James Bergstra and Yoshua Bengio. „Random search for hyper-parameter optimization.“ In: *Journal of machine learning research* 13.2 (2012).
- [78] Simon Bourret. „Neutrino oscillations and earth tomography with KM3NeT-ORCA“. PhD thesis. APC, Paris, 2018.
- [79] M Hatlo et al. „Developments of mathematical software libraries for the LHC experiments“. In: *IEEE Transactions on Nuclear Science* 52.6 (2005), pp. 2818–2822.
- [80] O Lalakulich, K Gallmeister, and U Mosel. „Neutrino nucleus reactions within the GiBUU model“. In: *Journal of Physics: Conference Series*. Vol. 408. 1. IOP Publishing. 2013, p. 012053.
- [81] Sebastiano Aiello et al. „gSeaGen: The KM3NeT GENIE-based code for neutrino telescopes“. In: *Computer Physics Communications* 256 (2020), p. 107477.
- [82] Felipe Pezoa et al. „Foundations of JSON Schema“. In: *Proceedings of the 25th International Conference on World Wide Web. WWW '16*. Montréal, Québec, Canada: International World Wide Web Conferences Steering Committee, 2016, pp. 263–273. ISBN: 9781450341431. DOI: 10.1145/2872427.2883029. URL: <https://doi.org/10.1145/2872427.2883029>.
- [83] Chiara Filomena Lastoria. „Tau neutrino appearance measurement in KM3NeT/ORCA 6“. In: *PoS TAUP2023* (2024), p. 192. DOI: 10.22323/1.441.0192.
- [84] Luc Cerisy et al. „Particle identification in KM3NeT/ORCA“. In: *38th International Cosmic Ray Conference*. 2023, p. 1191.
- [85] S. Aiello et al. „Determining the neutrino mass ordering and oscillation parameters with KM3NeT/ORCA“. In: *The European Physical Journal C* 82.1 (Jan. 2022). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-021-09893-0. URL: <http://dx.doi.org/10.1140/epjc/s10052-021-09893-0>.

Acknowledgements

I would like to thank all the members from ECAP for the exceptional time I had during my master's thesis. From the lunches to the organized activities, it was a rewarding environment for students. The discussions during the Neutrino and KM3NeT meetings taught me a lot. I would like to extend my special thanks to

- PD Dr. Thomas Eberl for giving me the opportunity to conduct my thesis with him and also for his countless backing throughout my time at the Erlangen Centre for Astroparticle Physics (ECAP)
- Rodrigo Gracia-Ruiz, Nicole Geißelbrecht and Alba Domi for the fruitful discussions I had with them and always giving me a lending hand when I asked for their opinion
- Rodrigo Gracia-Ruiz, Elena Petsali and Dimitris Trigoniaris for proofreading my thesis

Lastly, I would like to give my gratitude to my loving parents, who have supported me in all my endeavors so far.

Eigenständigkeitserklärung

Hiermit versichere ich, Michail Chadolias (23078433), die vorgelegte Arbeit selbstständig und ohne unzulässige Hilfe Dritter sowie ohne die Hinzuziehung nicht offengelegter und insbesondere nicht zugelassener Hilfsmittel angefertigt zu haben. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen und wurde auch von keiner anderen Prüfungsbehörde bereits als Teil einer Prüfung angenommen.

Die Stellen der Arbeit, die anderen Quellen im Wortlaut oder dem Sinn nach entnommen wurden, sind durch Angaben der Herkunft kenntlich gemacht. Dies gilt auch für Zeichnungen, Skizzen, bildliche Darstellungen sowie für Quellen aus dem Internet.

Mir ist insbesondere bewusst, dass die Nutzung künstlicher Intelligenz verboten ist, sofern diese nicht ausdrücklich als Hilfsmittel von dem Prüfungsleiter bzw. der Prüfungsleiterin zugelassen wurde. Dies gilt insbesondere für Chatbots (insbesondere ChatGPT) bzw. allgemein solche Programme, die anstelle meiner Person die Aufgabenstellung der Prüfung bzw. Teile derselben bearbeiten könnten.

Ort, Datum

Unterschrift