Event classification with the electronic detectors of the OPERA experiment using neural networks

Dissertation

zur Erlangung des Doktorgrades

des Department Physik

der Universität Hamburg

vorgelegt von

Martin C. Hierholzer

aus Münster

Hamburg

2012

Gutachter der Dissertation:	Prof. Dr. Caren Hagner Prof. Dr. Walter Schmidt-Parzefall
Gutachter der Disputation:	Prof. Dr. Caren Hagner JProf. Dr. Alessandro Mirizzi
Datum der Disputation:	30.01.2012
Vorsitzender des Prüfungsausschusses:	Prof. Dr. Dieter Horns
Vorsitzender des Promotionsausschusses:	Prof. Dr. Peter Hauschildt
Dekan der MIN-Fakultät:	Prof. Dr. Heinrich Graener

To my father.

Abstract

The OPERA experiment searches for $\nu_\mu \leftrightarrow \nu_\tau$ oscillations in appearance mode. It uses the emulsion cloud chamber (ECC) technique for a high spatial resolution combined with on-line components for event localisation and muon identification.

The analysis of events in an ECC detector takes considerable time, especially in case of ν_τ/ν_e candidate events. A ranking of events by a probability for being a ν_τ/ν_e event can speed up the analysis of the OPERA experiment. An algorithm for such an event ranking based on a classification-type neural network is presented in this thesis. Almost all candidate events can be found within the first 30% of the analysed events if the described ranking is applied. This event ranking is currently applied for testing purposes by the OPERA collaboration, a decision on a full application for the whole analysis is pending.

A similar neural network is used for discrimination between neutral and charged current events. This is used to observe neutrino oscillations in disappearance mode with the on-line components of the OPERA detector by measuring the energy dependence of the fraction of neutral current interactions. The confidence level of the observed oscillation effect is 87%. Assuming full mixing, the mass splitting has been determined to $|\Delta m^2_{32}| = 2.8^{+1.4}_{-1.7} \cdot 10^{-3}\,\mathrm{eV}^2$.

Zusammenfassung

Das OPERA-Experiment sucht nach $\nu_\mu \leftrightarrow \nu_\tau$-Oszillationen im *Apprearance*-Modus. Es verwendet die Emulsions-Nebelkammer-Technik (ECC) für eine hohe Ortsauflösung in Kombination mit Online-Komponenten für die Lokalisierung der Ereignisse und Identifikation der Myonen.

Die Analyse von Ereignissen im ECC-Detektor bringt einen erheblichen Zeitaufwand mit sich, insbesondere im Fall von ν_τ/ν_e-Kandidaten-Ereignissen. Das Aufstellen einer Rangfolge der Ereignisse nach ihrer Wahrscheinlichkeit, ein ν_τ/ν_e-Ereignis zu sein, kann die Analyse des OPERA-Experiments beschleunigen. In der vorliegenden Arbeit wird ein Algorithmus für eine solche Rangfolge basierend auf einem neuronalen Netz vom Klassifizierungs-Typ präsentiert. Wird die beschriebene Rangfolge angewendet, so werden fast alle Kandidaten-Ereignisse in den ersten 30% der analysierten Ereignisse gefunden. Der Algoritmus wird zur Zeit zu Testzwecken von der OPERA-Kollaboration eingesetzt, eine Entscheidung über die Anwendung auf die vollständige Analyse steht noch aus.

Ein ähnliches neuronales Netzwerk wird zur Unterscheidung von Neutralen- und Geladenen-Strom-Ereignissen verwendet. Dieses kann zur Beobachtung von Neutrino-Oszillationen im *Disappearance*-Modus mit den Online-Komponenten des OPERA-Detektors verwendet werden, indem der Anteil an Neutralen-Strom-Wechselwirkungen in Abhängigkeit der Energie bestimmt wird. Das Vertrauensnievau des beobachteten Oszillations-Effektes ist 87%. Unter der Annahme maximaler Mischung wird die Massendifferenz zu $|\Delta m^2_{32}| = 2.8^{+1.4}_{-1.7} \cdot 10^{-3}\,\mathrm{eV}^2$ bestimmt.

Contents

List of Figures

List of Tables

Preface

Today, the flavour oscillation of neutrinos is a very well established concept. Several decades after the first indications for oscillations of solar neutrinos by Cowan and Reines in 1956, these oscillations are confirmed by many other experiments. Neutrino oscillations have also been confirmed with other neutrino sources like atmospheric and reactor neutrinos and artificial neutrino beams. While in the solar sector the appearance of oscillated neutrinos has been confirmed and thus other models like a neutrino decay are excluded, only the disappearance of neutrinos from oscillation has been observed in the atmospheric sector so far. The appearance of ν_τ by $\nu_\mu \leftrightarrow \nu_\tau$ oscillations is yet to be confirmed.

This confirmation is the primary goal of the OPERA experiment. It searches for ν_τ in a long-baseline ν_μ beam. The detection of ν_τ involves the identification of τ leptons. Due to its short life time, this is only possible with a sub-millimetre track resolution inside the whole target mass. To achieve such high resolution, a hybrid detector design was chosen for OPERA. Sub-millimetre resolution is achieved with nuclear emulsions, while electronic detectors give an on-line information about the neutrino interactions.

The analysis of an event in OPERA involves developing and scanning the nuclear emulsion sheets, which takes considerable more time then the analysis of events in a pure on-line detector. The analysis of candidate events takes even additional time, as more information is required and thus more emulsion sheets need to be developed and scanned. Candidate events are also scanned multiple times in different scanning labs to increase the confidence in the result, which also requires more time.

The on-line components of OPERA are used to locate a neutrino interaction and to determine the emulsion sheets to be analysed. The information from the muon spectrometer can also help identifying muons and measuring their charge and momentum. Any classification to discriminate possible candidate from background neutrino interactions in the target before analysis of the emulsion sheets was

originally not foreseen. Even if it is preferable to analyse all events in the end for a better control of systematical effects, a ranking of events by the probability of the event to be a candidate event can considerably speed up the analysis.

In this thesis an algorithm for such a ranking is presented. It is based on a neural network, which uses only information available from the on-line components of the OPERA detector. A special technique is applied to reduce the influence of systematic errors of the Monte Carlo simulation used for the training of the neural network. A correctly set-up neural network outputs the a-posteriori probability for a given event belonging to the event class which the network has been trained to detect. Different versions of the network have been trained to detect different classes of events. Besides the detection of ν_τ events, a search for ν_e events with OPERA could be interesting as recently hints for a large Θ_{13} have been found by other experiments (see Section 1.3.1.3). Besides the networks to detect ν_τ and ν_e events, one combined network to detect both ν_τ and ν_e events has been trained. This network is working equally well as the independent networks. A ranking by the probability given by the network allows a discovery of almost all candidate events within the first 30% of the analysed events. Only ν_τ events involving a τ decaying into a μ have a non-negligible contribution in the remaining 70% of the data. Nevertheless the proposed ranking will speed up the discovery of such events as well.

A discrimination of neutral and charged current interactions is also possible with this neural network. This can be used to determine the fraction of neutral current interactions in energy dependence. For low energies an increase of the fraction of neutral current interactions from the disappearance of oscillated neutrinos is expected. This allows an observation of neutrino oscillations in disappearance mode with the on-line components of the OPERA detector. The significance of this observation is 87%. Assuming maximal mixing, the mass-splitting is determined to $|\Delta m^2_{32}| = 2.8^{+1.4}_{-1.7} \cdot 10^{-3}\,\mathrm{eV}^2$.

This thesis starts with a short review on the current status of neutrino physics in Chapter 1. The standard model of particle physics and the phenomenology of neutrino oscillations is shortly described. Then, results from current and recent neutrino experiments are shown. Finally, the different experimental techniques are shortly depicted at the example of several current and future neutrino experiments.

Chapter 2 gives a more detailed picture of the OPERA experiment. The hybrid detector is described in its components, with special emphasis on the emulsion cloud chamber technique used to detect ν_τ interactions. The first results, in particular the first ν_τ candidate event, are shortly described as well.

The relevant neutrino cross sections for this analysis are calculated in Chapter 3. Cross sections for charged current deep inelastic scattering are calculated in leading order QCD. Relations to other cross sections are calculated using a Monte Carlo simulation. Using additionally information on the beam composition, the relative abundances of the different event classes in the Monte Carlo simulation used for the training of the neural networks are calculated. In this chapter, a theoretical expectation for the NC ratio measured in Chapter 5 is provided as well.

In Chapter 4 neural networks are described mathematically. The correct layout of a classification Multi-Layer Perceptron is explained. The method used to reduce the effect of systematic errors is summarised as well.

The first four chapters build the foundation for the analysis presented in Chapter 5. First, the set-up and training of the neural network is described. The calculation of relevant detection efficiencies of the OPERA emulsion cloud chamber detector is shortly depicted. Finally, the results for the ranking of events and for the neutrino oscillation disappearance search by the neutral current ratio are presented.

A short conclusion completes the thesis at hand.

Chapter 1

Short review of neutrino physics

In this chapter the current status of neutrino physics is reviewed. Particular emphasis is given to oscillation physics. First, the fundamental model of neutrino oscillations is outlined. Next, the current experimental results are depicted. Finally, the most important experimental techniques are described at the example of several current and future neutrino experiments.

1.1 Neutrinos in the standard model of particle physics

In the standard model of particle physics there are twelve fermions and twelve vector bosons (see Figure 1.1. The fermions are divided into six quarks and six leptons in three generations. The quarks are charged with a fractional charge (u,c,t: $+\frac{2}{3}$; d,s,b: $-\frac{1}{3}$). Neutrinos have no electrical charge while the other leptons have a charge of -1. Neutrinos therefore can only interact via the weak force. All fermions have anti-particle counterparts differing in charge, lepton-number and chirality.

In the bosonic sector there are the photon, two charged W- and the neutral Z-boson and eight gluons. Photons are the exchange particles for the electro-magnetic interaction, while W- and Z-bosons are responsible for the weak and gluons for the strong interaction. In the standard model all particles but the neutrinos, the photons and the gluons have a mass.

Extensions to the standard model explore the integration of the gravitational force, e.g. by introducing the graviton as an additional boson. Also the particle masses are not explained by the current standard model. Mechanisms to generate

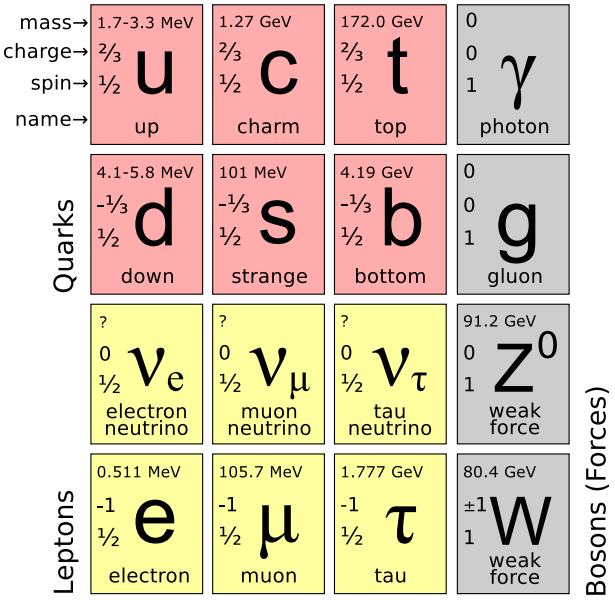

Figure 1.1: The particles in the standard model. Numbers taken from [42].

particle masses like the Higgs mechanism are tested by current experiments like ATLAS and CMS at the LHC.

The weak interaction only couples to left-handed particles and right-handed anti-particles and experimentally available neutrinos are always spin-polarized. Thus, it still remains unclear whether neutrinos are Majorana or Dirac particles, i.e. whether they are their own anti-particles or not. This is being tested in double beta decay experiments (see Section 1.3.4.1). The presence of a neutrino-less double beta decay would indicate that neutrinos are Majorana particles, but also requires neutrinos to have a non-zero mass – which therefore would present an extension to the standard model.

The number of light neutrino flavours (i.e. with masses less than $45\,\mathrm{GeV}$ and a weak isospin identical to the known neutrinos) has been determined to 3 from the decay width of the Z^0 boson [29].

Direct measurements of the neutrino masses – especially in tritium decay experiments – did not yet reveal any evidence for a non-zero mass of neutrinos. The upper limit for the effective ν_e mass from tritium decay has been determined to $2\,\mathrm{eV}$ [42].

1.2 Phenomenology of neutrino mixing

A first hint for neutrino oscillations was the deficit of solar neutrinos measured in the Homestake experiment by R. Davis in the 1970's. Davis measured the rate of solar neutrinos by the neutrino-induced inverse β-decay of $^{37}\mathrm{Cl}$ into $^{37}\mathrm{Ar}$ and counting the $^{37}\mathrm{Ar}$ abundance radio-chemically. The measured flux was about a third of the expectation from the standard solar model. Ever since, many experiments have confirmed neutrino oscillations for solar, atmospheric and accelerator neutrinos. The description in this section follows [42, 50, 86].

1.2.1 Neutrino oscillations in vacuum

Neutrinos undergo flavour oscillations, if their mass eigenstates $|\nu_i\rangle$ ($i = 1, 2, 3$) do not match their weak eigenstates $|\nu_l\rangle$ ($l = \mathrm{e}, \mu, \tau$) – this implies that at least two mass eigenstates have a non-zero rest mass. Weak and mass eigenstates are related via the PMNS matrix[1] U:

$$|\nu_l\rangle = \sum_i U_{li}^* |\nu_i\rangle . \tag{1.1}$$

Unlike its analogue in the quark sector – the CKM matrix – the PMNS matrix is not dominated by its diagonal elements, two of the mixing angles are large.

Neutrinos are produced as weak eigenstates. According to (1.1), weak eigenstates are a coherent superposition of mass eigenstates, which can be propagated through time (assuming $|\nu\rangle$ being a plane wave):

$$|\nu_l(t)\rangle = \sum_i e^{-iE_i t} U_{li}^* |\nu_i\rangle , \tag{1.2}$$

with $E_i = \sqrt{m_i^2 + p^2} \approx p + \frac{m_i^2}{2p}$ the energy of $|\nu_i\rangle$, m_i its mass and $p \gg m_i$ its momentum. As neutrinos are also observed as weak eigenstates only, (1.1) and (1.2) lead to the transition probability to the flavour l' after time t:

$$P_{ll'} = |\langle \nu_{l'} | \nu_l(t) \rangle|^2 = \left| \sum_i \sum_{i'} U_{li}^* U_{l'i'} \langle \nu_{i'} | \nu_i(t) \rangle \right|^2 . \tag{1.3}$$

Under the assumption of CP conservation, this leads to:

$$P_{ll'} = \delta_{ll'} - 4 \sum_i \sum_{i'} \Re \left[U_{li} U_{l'i}^* U_{li'} U_{l'i'} \right] \sin^2 x_{ii'} , \tag{1.4}$$

[1] Pontecorvo-Maki-Nakagawa-Sakata matrix. Pontecorvo predicted (a different kind of) neutrino oscillations, and Maki, Nakagawa and Sakata later introduced this matrix to explain flavour oscillations.

with

$$x_{ii'} = \frac{\Delta m^2_{ii'} L}{4E} = \pi \frac{L}{L_0}, \tag{1.5}$$

with the mass splitting $\Delta m^2_{ii'} = m^2_i - m^2_{i'}$ and the oscillation length $L_0 = \frac{4\pi E}{\Delta m^2_{ii'}}$, or in convenient units:

$$x_{ii'} = 1.27 \frac{\Delta m_{ii'}}{\mathrm{eV}^2} \frac{L/E}{\mathrm{m/MeV}}. \tag{1.6}$$

In case of three neutrino types, U can be parametrised as:

$$\begin{aligned} U &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \\ &\quad \cdot \begin{bmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (1.7) \\ &= \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}c_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{bmatrix} \cdot \\ &\quad \cdot \begin{bmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad (1.8) \end{aligned}$$

with $s_{ii'} = \sin\Theta_{ii'}$ and $c_{ii'} = \cos\Theta_{ii'}$. Due to unitarity, this matrix has six parameters: three mixing angles $\Theta_{ii'}$, one CP-violating phase δ and two additional Majorana phases α_k. Values and constraints for these parameters are given in Section 1.3.1.

For most experiments the two flavour approximation is sufficient, as $\Theta_{13} \ll \Theta_{23}, \Theta_{12}$ and $\Delta m^2_{21} \ll \Delta m^2_{32}$. The mixing matrix is then reduced to a two-dimensional rotation:

$$U = \begin{bmatrix} \cos\Theta & \sin\Theta \\ -\sin\Theta & \cos\Theta \end{bmatrix}, \tag{1.9}$$

which has only one parameter Θ. There also remains only one mass splitting $\Delta m^2 = m^2_2 - m^2_1$. The probability for two-flavour transitions simplifies to:

$$P_{ll'} = \delta_{ll'} - \frac{1}{2}\sin^2 2\Theta \left(1 - \cos 2\pi \frac{L}{L_0}\right), \tag{1.10}$$

with the oscillation length $L_0 = \frac{4\pi E}{\Delta m^2}$.

1.2.2 Matter effects

When neutrinos travel through matter, they forward-scatter coherently on the electrons, protons and neutrons present. Incoherent scattering processes – as described in Chapter 3 – do not affect neutrino oscillations in common scenarios, as the mean free path for those processes is much larger than the radii of usual aggregations of dense matter like the sun or the earth. Coherent scattering occurs via neural or charged current. Cross sections of neutral current scattering are equal for all flavours, therefore neutral current scattering does not affect neutrino oscillations. On the other hand, only $\overset{(-)}{\nu}_e$ undergo charged current scattering on the electrons present in matter. The flavour-dependent interaction probability corresponds to an effective mass term for $\overset{(-)}{\nu}_e$ and thus affects the mass splittings and flavour conversion probabilities. Neutrino oscillations in matter can still be described by the same equations used for vacuum oscillations if modified oscillation parameters Θ_m and Δm^2_m depending on the electron density are introduced. This effect was first suggested by Wolfenstein [88].

1.3 Current and future neutrino experiments

In this section a short overview of the most important neutrino experiments will be given. First, a global picture of the current situation is drawn. Each experimental technology will then be introduced with exemplary experiments, while for other experiments only the important differences will be named.

1.3.1 Current picture

Currently, only four of the eight oscillation parameters have been determined: two of the three mixing angles Θ_i and the two mass-splittings $\Delta m^2_{ii'}$. The sign of Δm^2_{32} is still unknown, which leaves two possible mass hierarchies. For Θ_{13} only an upper bound exists, though there are first hints for a non-zero Θ_{13} (see Section 1.3.1.3). All three phases are completely unknown at this time. Best fit values and uncertainties from a global analysis are listed in Table 1.1.

1.3.1.1 Solar sector

Θ_{12} and Δm^2_{21} have been determined from solar and reactor experiments. Experiments with solar neutrinos were the first to discover hints for neutrino mixing in

Table 1.1: Neutrino oscillation parameters from a global analysis. The sign of Δm^2_{32} and all three phases are currently unknown. $|\Delta m^2_{32}|$ is taken from [10], all other values from [43].

Parameter	Value
$\sin^2(2\Theta_{12})$	$0.861^{+0.026}_{-0.022}$
$\sin^2(2\Theta_{23})$	$> 0.92,\ \mathrm{CL} = 90\%$
$\sin^2(2\Theta_{13})$	$< 0.15,\ \mathrm{CL} = 90\%$
Δm^2_{21}	$(7.59 \pm 0.21) \cdot 10^{-5}\,\mathrm{eV}^2$
$\lvert\Delta m^2_{32}\rvert$	$(2.32^{+0.12}_{-0.08}) \cdot 10^{-3}\,\mathrm{eV}^2$

the 1970's. The SNO charged current (CC) measurement combined with Super-Kamiokande ν_e elastic scattering (ES) result provided the first direct evidence for a flavour change in 2001, which was later confirmed by the SNO neutral current (NC) measurement [42]. The results from SNO's CC, NC and ES measurements together with the Super-Kamiokande data are shown in Figures 1.2 and 1.3.

Due to matter effects (see Section 1.2.2), the solar result cannot easily be transferred to the (vacuum) oscillation parameters, as multiple solutions are possible. In combination with the reactor ($\bar{\nu}_e$) data from KamLAND, the so-called large mixing angle (LMA) solution is preferred with $> 5\sigma$ CL [42].

1.3.1.2 Atmospheric sector

Oscillations of atmospheric neutrinos, $\nu_\mu \leftrightarrow \nu_\tau$, have mainly been measured with water-Čerenkov detectors like Super-Kamiokande. The same oscillation channel can be measured with accelerator neutrinos, which is currently done e.g. by MINOS and OPERA. The most precise measurements are from MINOS ($|\Delta m^2_{32}|$) and from Super-Kamiokande (Θ_{23}). The best fit for atmospheric data is shown in Figure 1.4. OPERA is the only experiment operating in appearance mode in the atmospheric sector. Currently, it has observed one single ν_τ event, confirming $\nu_\mu \leftrightarrow \nu_\tau$ oscillations with a statistical significance of 95% [35]. In the next chapter the OPERA experiment and its present results will be discussed in detail. MINOS has also measured the oscillation parameters for $\bar{\nu}_\mu \leftrightarrow \bar{\nu}_\tau$, which are consistent with the $\nu_\mu \leftrightarrow \nu_\tau$ result [53].

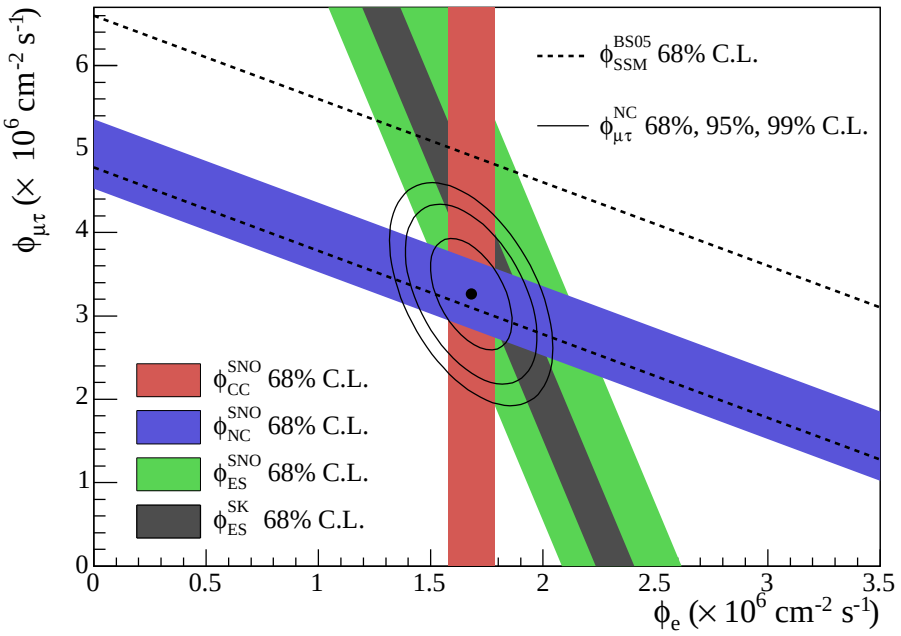

Figure 1.2: Combined result from SNO and Super-Kamiokande ^{8}B neutrino flux measurements for the different interaction channels (coloured bands) and the expectation from the standard solar model (SSM) for the total flux (dashed lines), shown as the flux of ν_μ and ν_τ vs. the flux of ν_e. The allowed region from all measurements (solid lines) clearly indicates a non-zero flux of ν_μ and ν_τ and therefore provides a strong evidence for neutrino oscillations. [14, 42]

1.3.1.3 Search for Θ_{13}

The CHOOZ reactor experiment [19] set the current world-best upper limit on Θ_{13} as shown in Table 1.1. A first direct hint for a non-zero Θ_{13} was delivered by the T2K experiment (see Sections 1.3.2.4 and 1.3.3.1) [55]. Θ_{13} has been found to be non-zero with a significance of 2.5σ. The allowed region is shown in Figure 1.5. The MINOS experiment (see Sections 1.3.2.4 and 1.3.3.3) also reports an excess of ν_e with a significance of 89% [9]. The Double Chooz experiment has as well measured a non-zero Θ_{13} [49]. A global analysis including solar data reveals a 3σ evidence for a non-zero Θ_{13} [45]. The continuing measurements by T2K, Double Chooz and other experiments will improve the significance of this result in near future.

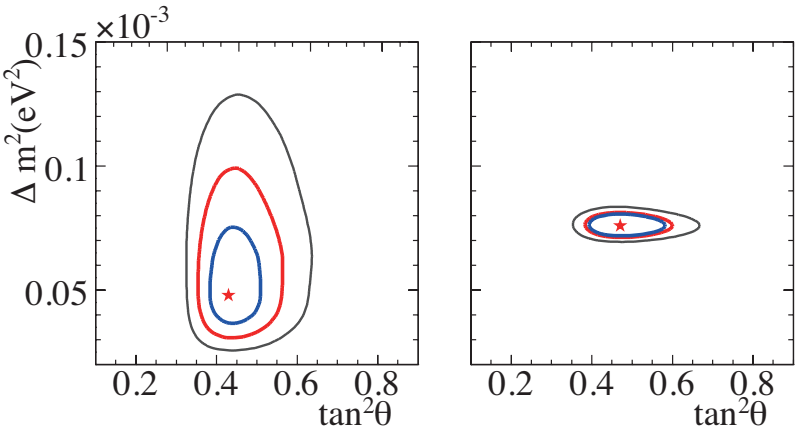

Figure 1.3: Allowed parameter region from global solar data without (left) and with (right) KamLAND, at 68%, 95%, and 99% CL as well as the best fit (star). [15, 42]

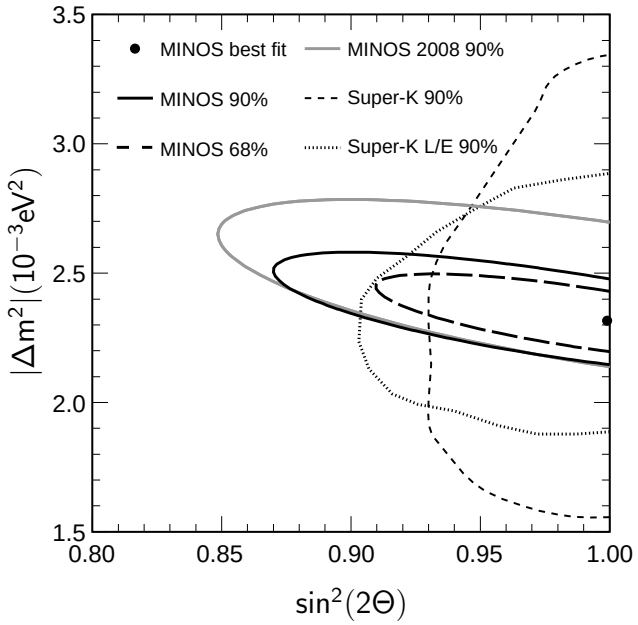

Figure 1.4: Allowed region of the Θ_{23}–Δm^2_{32} parameter space as published by the MINOS collaboration in 2011. Results from previous experiments are also included. [10]

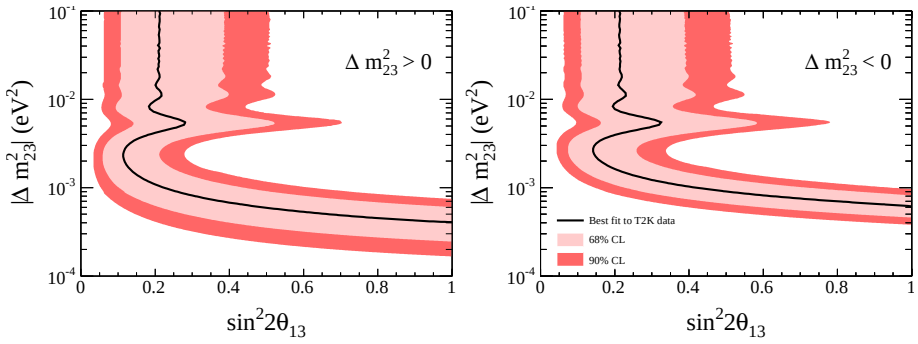

Figure 1.5: Allowed region in the Θ_{13}–Δm^2_{32} parameter space as published by the T2K collaboration. On the left the normal, on the right the inverted hierarchy scenario is shown. $\delta_{CP} = 0$ is assumed. [55]

1.3.1.4 Sterile neutrino flavours

LSND, a liquid-scintillator short-baseline experiment, has claimed the observation of an excess of $\bar{\nu}_e$ in 2001 [12]. As this result cannot be explained in a three neutrino-flavour scenario without contradicting existing results in the solar and the atmospheric sector, the LSND collaboration proposes a fourth, sterile neutrino flavour to explain all data. The LSND result was neither confirmed nor excluded by other experiments like KARMEN [20] until 2007, when MiniBooNE (see Section 1.3.3.1) published a result excluding an LSND-like result for ν_e [13]. Very recently, MiniBooNE published a preliminary result for $\bar{\nu}_e$ events favouring a LSND-like excess [90]. The situation remains therefore unclear, until MiniBooNE and similar experiments have gathered more data.

1.3.2 Neutrino sources

1.3.2.1 Solar neutrinos

Many early neutrino experiments used solar neutrinos, as they are available in large quantities for free and provided first hints for neutrino oscillations. They still play an important role in neutrino physics as they present a pure source of ν_e and are subject to matter effects due to the high density in the solar core. They also provide valuable insights into the processes in the solar centre – which are inaccessible by other observations.

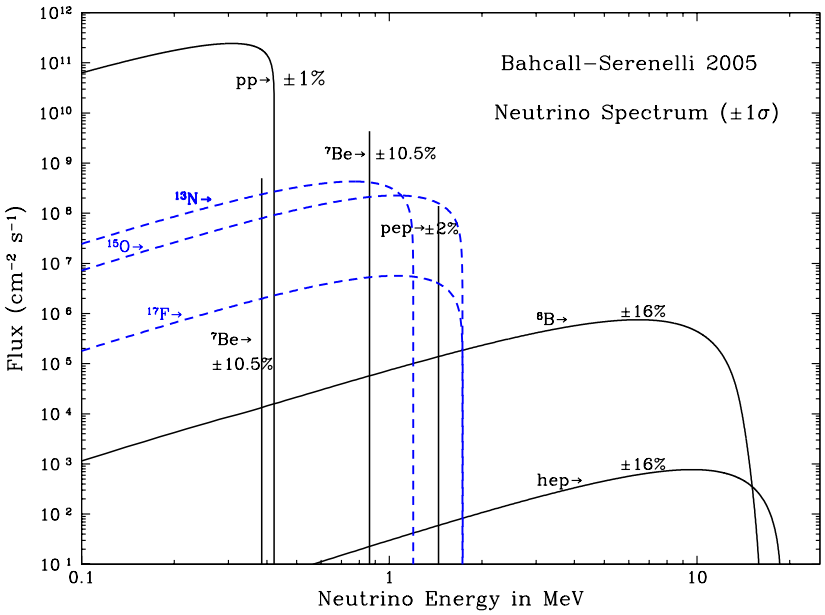

Figure 1.6: Spectrum of solar neutrinos, as predicted by the solar model BS05(OP) [28].

Like any other main sequence star, the sun fuses hydrogen into helium inside its core [28, 42]. For stars with $M < 1.3M_\odot$ the dominating fusion process is the proton–proton chain (or pp chain). Additionally, there is a small contribution from the CNO cycle, which dominates in stars with $M > 1.3M_\odot$. The sum of all processes result in the neutrino spectrum shown in Figure 1.6. The total flux of solar neutrinos is about $6.5 \cdot 10^{10}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ on earth. Neutrino detectors have to be optimized for low energies to detect solar neutrinos. So far, experiments have only detected neutrinos from ^{7}Be and ^{8}B decays and very recently neutrinos from the pep branch of the pp chain due to energy thresholds and background (see Section 1.3.3.2).

1.3.2.2 Atmospheric neutrinos

When high-energetic cosmic-ray particles hit nuclei of the Earth's atmosphere, a hadronic shower is produced. Eventually, the longer-lived hadrons $K^\pm$ and $\pi^\pm$ are produced, which themselves decay into μ and ν_μ. As the majority of μ also decays on their way to the Earth's surface, a ν_μ and ν_e each is produced as well.

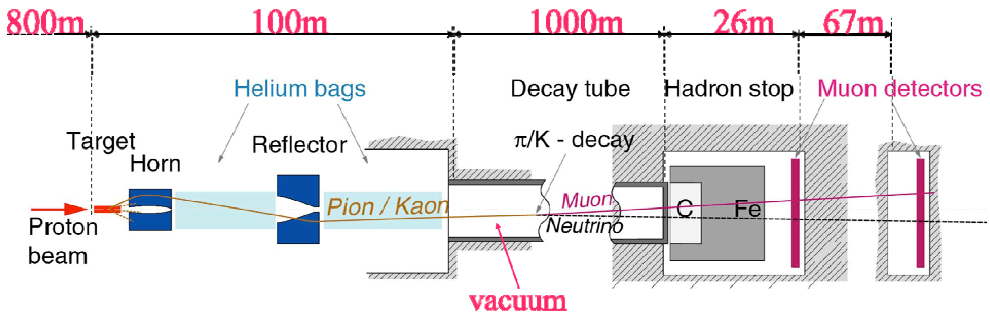

Figure 1.7: Sketch of a conventional neutrino beam facility. A proton beam incoming from the left hits the target. π/K are produced, charge selected and focussed by horn and reflector, and allowed to decay in the evacuated decay pipe into μ and ν_μ. The remaining hadrons are stopped and the beam is indirectly monitored by μ detectors. The specified dimensions are from the CNGS beam. [51]

Therefore in good approximation the expected ratio of atmospheric ν_μ to ν_e is 2:1. The energies are correlated to the energy of the primary cosmic ray particle, the spectral maximum is in the order of several GeV.

1.3.2.3 Reactor neutrinos

In nuclear fission reactors a constant flux of $\bar{\nu}_e$ is produced from the β^--decay of the neutron-rich fission products [40]. On average, six neutrinos are produced per fission. The continuous energy spectrum has a mean energy of 4 MeV. Early neutrino experiments (namely the first proof of the existence of neutrinos by Cowan and Reines [36]) used these neutrinos, as they were freely available. Also current and future experiments continue to use reactor neutrinos for low energy neutrino physics, e.g. measuring Θ_{13}. Typically, commercial power plants are used, due to their large thermal power. A good example is the Chooz plant in France with two blocks of a thermal power of 4.27 GW each, are used [40].

1.3.2.4 Conventional beams

Artificial neutrino sources feature many advantages. Experimental parameters like the baseline, the neutrino energy and the flavour can be chosen, the source can be pulsed (for background reduction) and the intensity can be higher. Conventional

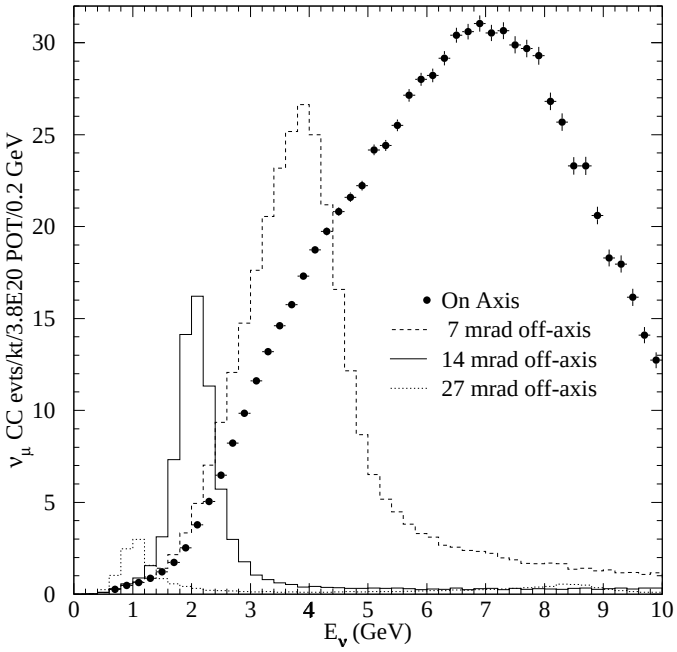

Figure 1.8: Spectrum of a neutrino beam depending on the off-axis angle, at the example of the NuMI beam. [24]

beams are inspired by atmospheric neutrino production. A beam of high-energetic protons hits a target (e.g. water or graphite) to produce K and π. To focus the π/K and select the charge special magnetic lenses called horn and reflector are used (see Figure 1.7). The particles then decay in an evacuated decay pipe into μ and ν_μ. The intensity and timing can be monitored by μ monitors.

As the ν_μ themselves cannot be focused, the beam still has a certain divergence. The neutrino spectrum observed in a detector depends on the angle relative to the beam axis. For the beam centre the spectrum is very broad. At larger angles, a narrower spectrum of lower-energetic neutrinos is selected (see Figure 1.8). Experiments which use this effect to select a more defined energy spectrum are called *off-axis*. Off-axis beams are often also called *super beams* due to the high intensity demands, but they still are conventional beams in terms of the neutrino production mechanism.

Currently, there are four neutrino beams operational: The CNGS beam at CERN in Europe delivering high-energetic neutrinos to OPERA and ICARUS, the NuMI beam at Fermilab in the USA creating a high-intensity beam for MINOS,

Table 1.2: Key parameters of current neutrino beams. [5, 13, 23, 26, 38, 39, 75, 78]

Beam	BNB	NuMI	CNGS	T2K
intensity $[10^{19}\ \frac{\text{p.o.t.}}{\text{a}}]$	18	21	4.5	~ 100
proton energy [GeV]	8	120	400	30
influence power at target [kW]	32	290	510	750
average neutrino energy [GeV]	0.7	$4 - 10^*$	18	0.6
base line [km]	0.5	735	730	295
off-axis angle	–	–	–	2.5°

* depending on configuration

MINERνA and (in future) NOνA, the T2K off-axis beam at J-PARC in Japan directed to Super-Kamiokande, and the low energy Booster Neutrino Beam-line (BNB) at FNAL in the USA currently running for MiniBooNE. For an overview on the important parameters of these beams (and references for further reading), see Table 1.2. The experiments are described in Section 1.3.3.

1.3.2.5 β-beams and neutrino factories

Currently, there are new approaches investigated to produce neutrino beams. A β-beam is an elongated storage ring for β-decaying ions at relativistic speeds. Ions decaying on the straight sections of the decay ring will produce a pure ν_e (or $\bar{\nu}_e$) beam, focussed by relativistic beaming. The European design study EURISOL [85] has considered several isotopes. ^{6}He can be produced in sufficient quantities and accelerated to a maximum $\gamma = 150$. ^{18}Ne can be accelerated up to $\gamma = 250$, but production in the required quantities seems not yet feasible. EURISOL uses existing infrastructure at CERN for the acceleration of the ions. Considering MEMPHYS in the Fréjus laboratory as a possible detector, $\gamma = 100$ would be an optimized choice for physics at a base line of 130 km. Typical energies of a β-beam are in the region of 100 MeV. Even lower energies can be selected by placing the detector off-axis, avoiding the need for a dedicated beam for low-energy neutrino experiments. In 2009, EURISOL published its finfluenceinal report [1]. Figure 1.9 shows a sketch of a possible β-beam facility.

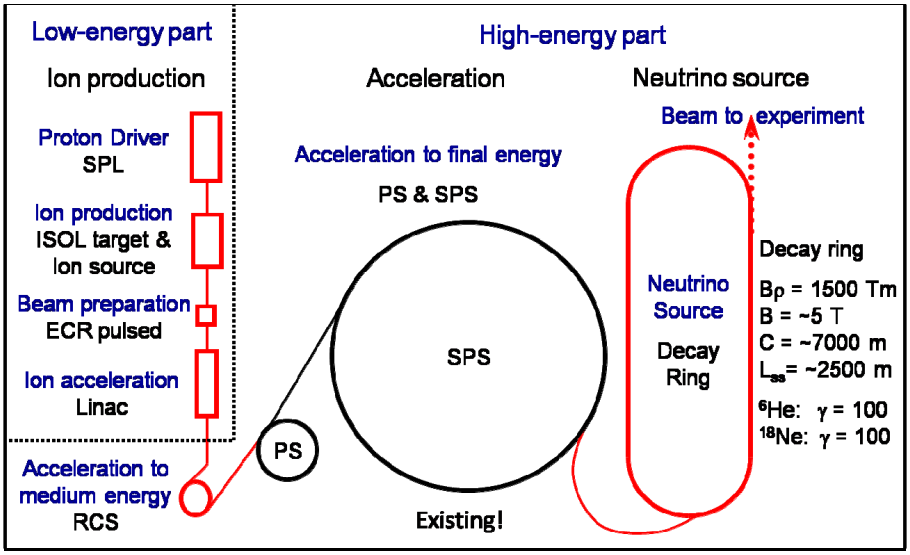

Figure 1.9: Possible layout of a β-beam at CERN. Ions are produced at an ECR (*electron cyclotron resonance*) source and pre-accelerated by a linear accelerator (Linac). They pass through a rapid cycling synchrotron (RCS) before being injected into the Proton Synchrotron (PS). The Super Proton Synchrotron (SPS) will accelerate the ions to their final energy before they are injected into the decay ring. The decay ring has two straight sections of 2500 m each on which the decaying ions contribute to the neutrino beam. [85]

Neutrino factories are muon storage rings which take advantage of the decaying muons to produce ν_μ and $\bar{\nu}_e$ (or $\bar{\nu}_\mu$ and ν_e in case of μ^+). As muon storage rings have not yet been built, neutrino factories might be realized even further in the future than β-beams. Currently, there are several design studies and R&D projects for components of a muon accelerator ongoing. In many aspects the technology is similar to those of muon colliders and thus both projects can benefit from each other, though the actual machines – if they should be built – will be very different. With help of a muon factory the so-called golden channel $\nu_e \leftrightarrow \nu_\mu$ could be measured very precisely by observing the appearance of differently charged muons in the detector w.r.t. the muons in the storage ring.

1.3.2.6 Supernova neutrinos

Core-collapse Supernovae present a natural source of neutrinos. At the end of the life-time of a large star ($M > 8M_\odot$), the iron core reaches the Chandrasekhar limit

and collapses to a proto-neutron star. The collapse itself produces a burst of ν_e, due to the conversion of p to n. However, the largest contribution to the neutrino output of a Supernova are the thermal neutrinos emitted by the cooling of the proto-neutron star, which carry away most of the energy of the explosion. They are equally distributed among the flavours. As the neutrinos immediately escape from inner regions of the exploding star, they are expected to reach Earth several hours before the corresponding photons.

So far, only the neutrinos of one Supernova have been observed. The Supernova SN1987A took place in a distance of $48\,\mathrm{kpc}$ from Earth in the Large Magellanic Cloud. Kamiokande-II, IMB and the Baksan observatory have detected in total 21.4 ± 1.2 neutrinos from the Supernova [83].

Many currently running neutrino detectors are ready to detect neutrinos from the next nearby Supernova. Some of them are connected to the SuperNova Early Warning System (SNEWS), which will give a warning to astronomers, before the Supernova will be visible in light [82].

Moreover, a diffuse neutrino background generated by the frequent Supernovae explosions at cosmological distances is also expected. While the faint signal is out of reach for current experiments, a detection could be feasible with next generation experiments.

1.3.3 Neutrino detectors

1.3.3.1 Water Čerenkov

Water Čerenkov detectors use the Čerenkov light emitted by charged secondary particles to detect neutrino interactions. Highly purified water is functioning as target and active material at the same time. The Čerenkov light is picked up by photo-multiplier tubes (PMT) mounted at the inner surface of the water tank. Different event classes can be discriminated by the shape of the Čerenkov light cone. E.g. ν_e CC events have a fuzzier shape than ν_μ CC events due to the electromagnetic shower produced by high-energy electrons in water. Water Čerenkov detectors can be scaled to very large target masses. The energy threshold is typically around $5\,\mathrm{MeV}$. As water can be purified to a very high level, there is almost no background from intrinsic radioactivity above the threshold.

Presently, the largest water Čerenkov detector is the **Super-Kamiokande** detector in Japan [4, 47]. A total of $50\,\mathrm{kt}$ of water splits into a $32\,\mathrm{kt}$ inner detector and

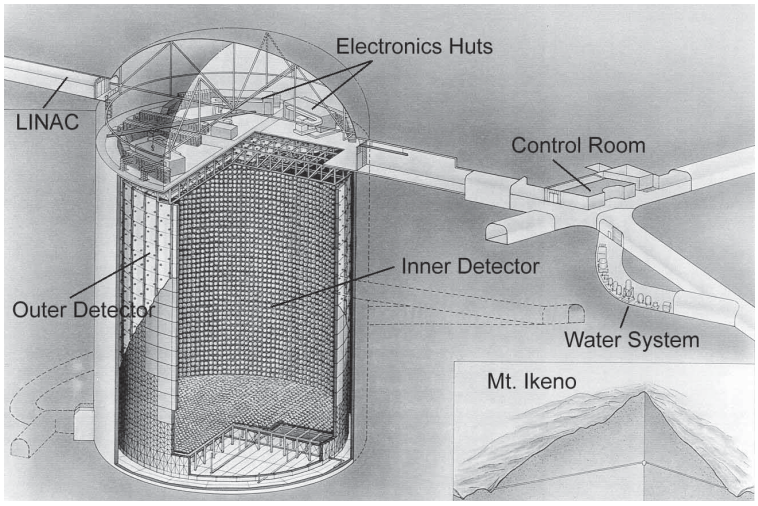

Figure 1.10: Sketch of the Super-Kamiokande detector. [84]

an outer veto detector. The inner detector is instrumented with 11,129 PMTs with a diameter of 50 cm, covering 40% of the surface. The fiducial volume for observing neutrino interactions is 22.5 kt. The outer detector is instrumented with 1,885 20 cm PMTs and detects incoming particles like cosmic ray muons. An electron linear accelerator is used for energy calibration, which can inject single electrons with energies between 4.4 MeV and 18.9 MeV at various positions inside the fiducial volume. An overview of the setup is shown in Figure 1.10.

The **HyperKamiokande** detector as a direct successor of the Super-Kamiokande experiment is planned [3]. It will consist of two identical cylindrical tanks with a fiducial mass of 0.56 Mt, using a design based on the Super-Kamiokande detector. It will also be used as a new far detector for the T2K beam, and therefore it will be located near the Super-Kamiokande detector.

Within the LAGUNA[2] design study for a European next-generation multi-purpose neutrino detector, a water Čerenkov solution called **MEMPHYS** is considered – besides other options described in the following subsections. MEMPHYS will be divided into three cylindrical modules, each similar to an up-scaled version of the Super-Kamiokande detector. The total fiducial volume will be around 0.5 Mt. The detector would be located in the Fréjus laboratory 130 km from CERN. Energies

[2] Large Apparatus studying Grand Unification and Neutrino Astrophysics

considered for a neutrino beam from CERN to MEMPHYS as far detector are between 0.2 and $0.4\,\mathrm{GeV}$.

The **MiniBooNE** experiment [13] uses also a Čerenkov detector, but with $800\,\mathrm{t}$ mineral oil instead of water. This way it can also use the scintillation light, e.g. to detect NC neutrino interactions. Another variation is the **SNO** experiment [56], which used $1\,\mathrm{kt}$ of ultra-pure heavy water. This allows the comparison of CC and NC channels, and therefore a comparison of the fluxes of ν_e and all neutrino flavours, as shown in Figure 1.2. In phase II of the SNO experiment, NaCl was added the the heavy water to increase the detection efficiency of NC events, in phase III the NaCl was removed and an array of proportional counters optimized for the detection of NC events was installed.

1.3.3.2 Liquid Scintillator

Neutrinos can also be detected by the scintillation light emitted from the secondary particles in an appropriate scintillating medium. Like water, liquid scintillator can function as target and detection material at the same time and allows large target masses. Analogous to water Čerenkov detectors the scintillation light is collected by PMTs. In most designs, the active scintillation volume is separated by a layer of buffer oil from the PMTs to shield against intrinsic radioactivity of the PMTs and the support structure. Typically, liquid scintillator detectors are sensitive to lower neutrino energies then water Čerenkov detectors but provide little to no directional information about the incident neutrino.

Borexino [17, 73] is a liquid scintillator detector designed mainly to detect solar neutrinos from the $^7\mathrm{Be}$ lines (see Figure 1.6). It is the first experiment to detect these neutrinos in real-time and with information on their energy. The detector layout is shown in Figure 1.11. The liquid scintillator used is pseudocumene (1,2,4-trimethylbenzene) doped with PPO (2,5-diphenyloxazole) as a wavelength shifter. Around the inner volume filled with scintillator is a layer of non-scintillating pseudocumene as buffer. 2,200 PMTs mounted on a steel sphere around the buffer layer pick up the scintillation light. On the outside of the steel sphere and on the floor, additional 200 PMTs are mounted to detect Čerenkov light from the outer veto detector filled with water. Borexino is running since 2007 and has successfully measured the spectrum of solar $^7\mathrm{Be}$ neutrinos, as well as detected geo-neutrinos and solar $^8\mathrm{B}$ neutrinos [77]. Very recently, it has also measured the flux of solar neutrinos from the pep reaction and presented a new constraint on solar CNO neutrinos [34].

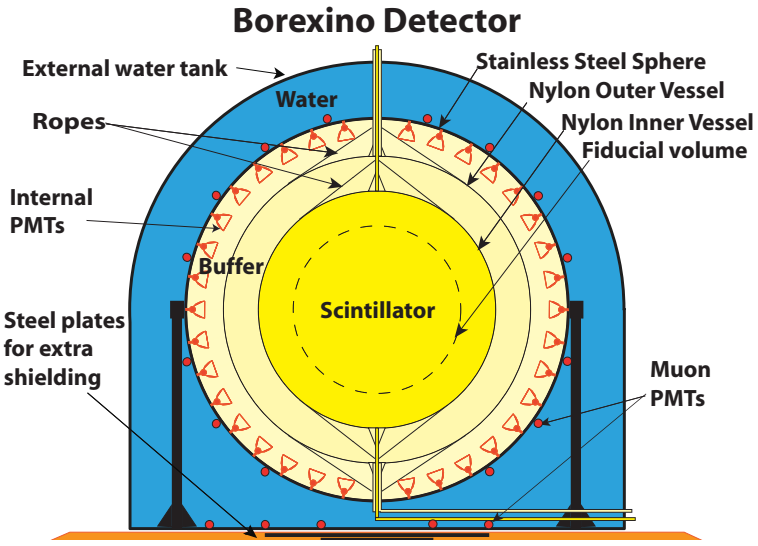

Figure 1.11: Layout of the Borexino detector. The detector has an onion-like structure with the following layers (from the inside): liquid scintillator (with a software cut to the 100 t fiducial volume), nylon sphere, buffer, steel sphere with PMTs, water buffer. [17]

KamLAND [48] is a 1 kt liquid scintillator experiment built to measure $\bar{\nu}_e$ from nuclear reactors in Japan. Its design is very similar to the one of Borexino. KamLAND is taking data since 2002 and has complemented measurements from the solar sector (cf. Section 1.3.1.1) and delivered the first direct evidence for geo-neutrinos.

Double Chooz [76] is the successor of the **CHOOZ experiment** which both measure $\bar{\nu}_e$ from the two Chooz reactor cores in France. Double Chooz uses 10 t of Gd-doped liquid scintillator as target. To fully contain the energy of γ emitted from neutron capture on Gd and positron annihilation close to the border of the target, a so-called gamma catcher consisting of pure liquid scintillator is surrounding the target. Similar to Borexino and KamLAND, the scintillating area is separated from the PMTs mounted on the steel vessel by a layer of buffer oil. Double Chooz has two veto detectors. The inner veto is surrounding the inner detector and consists of liquid scintillator. The outer veto is a layer of plastic scintillator strips on top of the detector. As CHOOZ has delivered the world-best limit on Θ_{13}, Double Chooz aims at positive detection or at least a more restrictive limit of $\sin^2(2\Theta_{13}) < 0.03$ (90% CL). To achieve a low systematic uncertainty, Double

Chooz will have a near and a far detector. The far detector is taking data since March 2011, the near detector is yet to be built.

Experiments similar to Double Chooz are **Daya-Bay** in China [52] and **Reno** in Korea. Daya-Bay uses six reactor cores as source and has two near detectors and one far detector. Reno has one near and one far detector and also uses six reactor cores. Both experiments are currently being set-up and already partially under commissioning.

The **SNO+** experiment [70] is a multi-purpose detector currently under construction. The existing SNO detector (see Section 1.3.3.1) will be modified to replace the heavy water with liquid scintillator. Its main physics goals are measuring the solar pep and CNO neutrino fluxes, geo-neutrino fluxes and the search for Supernova neutrino events.

As part of the LAGUNA design study also a liquid scintillator detector is being investigated: The envisaged target mass of the **LENA** [89] experiment is around $50\,\mathrm{kt}$. Outside a cylindrical steel tank, on which the the PMTs are mounted, is a water Čerenkov muon veto that is complemented with an additional veto detector on top of the detector. The physics potential reaches from low energy physics with Galactic Supernova and diffuse Supernova background neutrinos to the GeV range with long-baseline beams and atmospheric neutrinos.

1.3.3.3 Tracking Calorimeters

A tracking calorimeter is typically a segmented detector with interleaved passive and active components (*sandwich calorimeter*). The passive component provides a large target mass and is usually made from iron or lead. Between layers of the passive mass, active components with a good spatial resolution are placed. This allows a tracking of isolated particles like muons and a reconstruction of the energy deposited in particle showers. Often, the passive material is magnetised to allow a charge separation to discriminate between ν and $\bar{\nu}$.

The **MINOS** experiment [72] uses two magnetised iron calorimeters at the NuMI beam. The far detector (see Figure 1.12) has a mass of $5.4\,\mathrm{kt}$ and is located in a distance of $735\,\mathrm{km}$ of the beam source. The near detector has a mass of $0.98\,\mathrm{kt}$ and has been optimised for the higher neutrino interaction rate. The magnetic field strength is $1.42\,\mathrm{T}$ in the far and $1.28\,\mathrm{T}$ in the near detector. The active component is a plastic scintillator detector with a granularity of $4.1\,\mathrm{cm}$. The signal readout is

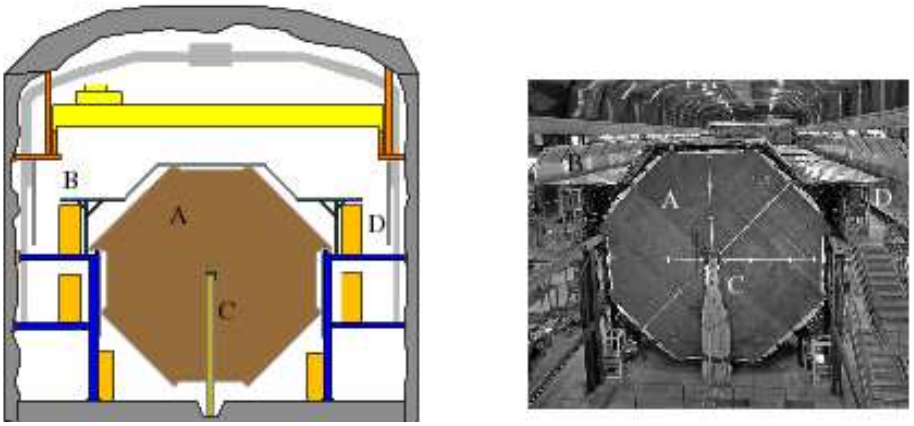

Figure 1.12: Sketch (left) and photo (right) of the MINOS far detector. 'A' is the most downstream iron plane, 'B' is the cosmic ray muon veto, 'C' is the magnet coil and 'D' is a part of the read out electronics. [72]

based on wavelength shifting fibres connected to PMTs. The MINOS experiment is taking data since 2005.

The **OPERA** experiment uses lead as target material, interleaved with plastic scintillator as online and nuclear emulsions as offline active components. Additionally, it has dedicated muon spectrometers. The detector is taking data since 2007. It is described in more detail in Chapter 2.

The **near detectors of the T2K experiment** [22] also consist of tracking calorimeters. The on-axis neutrino monitor is an array of iron–scintillator stacks. The off-axis near detector ND280 is highly segmented (see Figure 1.13). The upstream portion contains a fine-granularity tracking calorimeter consisting of scintillator planes and water target sections (P0D in the figure). It is specifically built to reconstruct NC events containing a π^0 and to measure neutrino cross sections on water. It is surrounded by an electromagnetic scintillator–lead calorimeter (P0D ECAL and Barrel ECAL). The downstream portion is a μ tracking system to reconstruct CC interactions. It consists of fine graded detectors (FGDs) used as additional neutrino targets and TPCs used for tracking and particle identification. The whole detector suite is placed inside a $0.2\,\mathrm{T}$ magnet formerly used in the UA1 experiment at CERN.

NOvA [25] follows a slightly different approach. Passive and active material are the same. It will use $24\,\mathrm{kt}$ of liquid scintillator contained in a segmented rigid

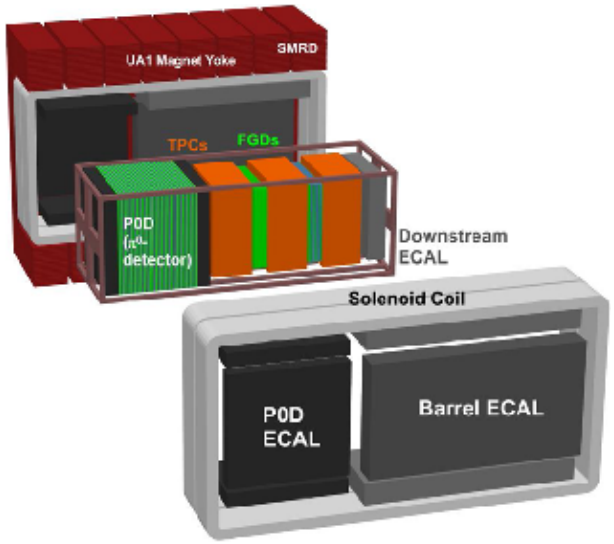

Figure 1.13: Exploded view of the T2K near detector ND280. [22]

PVC structure. The readout is performed via wavelength-shifting fibres connected to avalanche photo diodes. The granularity of the calorimeter is 0.15 radiation lengths. The detector will be located off-axis of the NuMI beam for a lower and narrower energy spectrum.

1.3.3.4 Liquid Argon TPC

A relatively new approach to detect neutrino interactions is a liquid argon time projection chamber (LAr TPC). Target and active materials are both liquid argon. An interacting neutrino will produce high energetic charged particles, which will ionise the argon atoms along their trajectories. A high voltage applied to two electrodes at opposite sides of the chamber will drift the electrons and ions to the anode and cathode, respectively. The scintillation light of the secondary particles can be used as a trigger for the drift time measurement. Thus, a full three-dimensional reconstruction of the neutrino interaction with a resolution in the order of $1\,\mathrm{mm}$ is possible.

The **ICARUS** experiment [81] is the largest currently operating LAr TPC. It is located in the LNGS underground laboratory and therefore can observe neutrino interactions from the CNGS beam. The detector has a total active mass of $476\,\mathrm{t}$ and consists of two equal modules. Each module houses two TPCs separated by a

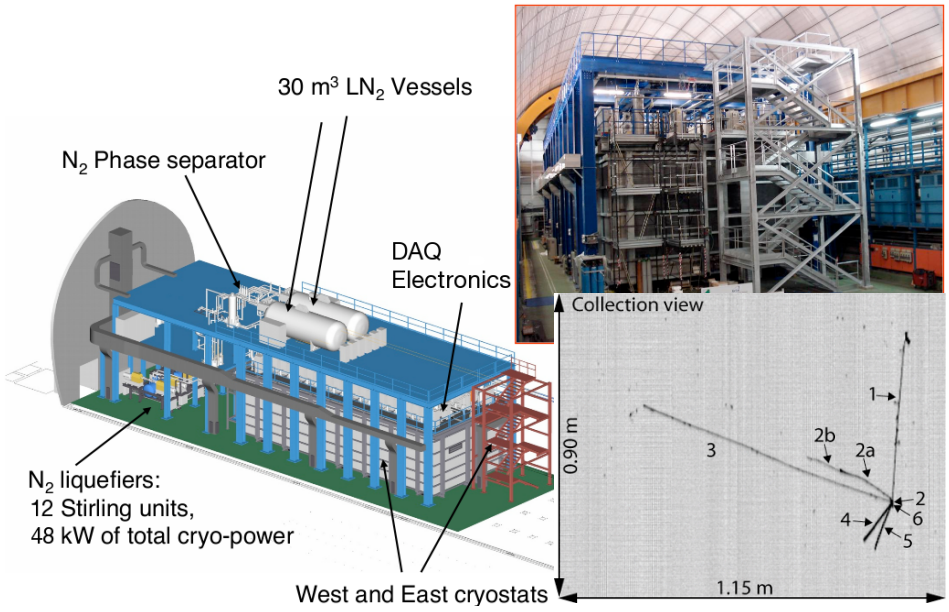

Figure 1.14: Schematic view (left) and photography (top right) of the ICARUS detector. On the bottom right, an example of a low-energy neutrino interaction is shown. Single particle tracks are resolved and the particles can be identified by dE/dx and range, e.g. the track labelled with 1 is a π and track 3 is a μ. [81]

common cathode. The set-up and an example of a neutrino event in ICARUS can be seen in Figure 1.14.

A LAr TPC option called **GLACIER** [27, 37] is also considered within the LAGUNA project. It has a cylindrical tank and the readout relies on Large Electron Multipliers (LEMs). The concept is assumed to scale to target masses of 100 kt. Current prototypes have target masses around 1 t.

MicroBooNE [57] is a LAr TPC with an active mass of 70 t. It is currently being set-up at the BNB beam with a base line of 100 m and will start taking data in 2014. Its purpose is to confirm or disprove the results of the MiniBooNE experiment. It will also allow a measurement of neutrino cross sections on argon.

1.3.4 Non-oscillation neutrino physics

Besides neutrino oscillation physics, there are other interesting fields in neutrino physics promising interesting new results. Some topics, like the observation of

solar neutrinos to improve the knowledge about the processes inside the sun, were already covered in the last sections. In this section, a short overview on double beta decay experiments, direct measurements of the neutrino mass and neutrino velocity experiments is given, to complete the picture.

1.3.4.1 Neutrinoless double beta decay

Double beta decays occur if, for a given isotope, the single beta decay is forbidden for kinematical reasons while the decay to the next but one isotope is energetically allowed. A neutrino-less double beta decay is possible if the ν emitted by the first decay can be absorbed as $\bar{\nu}$ in the second decay. This would prove that neutrinos have a non-zero Majorana mass. The currently best limit for this decay mode of ^{76}Ge has been set by the **Heidelberg-Moscow experiment** (HdM) to $T^{0\nu}_{1/2} = 1.9 \cdot 10^{25}\,\mathrm{a}$ at 90% C.L. [58]. A sub-group of the HdM collaboration has claimed the observation of a positive signal indicating a half-life of $T^{0\nu}_{1/2} = (2.23^{+0.44}_{-0.31}) \cdot 10^{25}\,\mathrm{a}$ [59, 60].

The succeeding **GERDA** experiment [71] is also using ^{76}Ge. Like in the HdM experiment, multiple Ge-semiconductor detectors act at the same time as the signal source. In the first phase, GERDA uses partially the same detectors as the HdM experiment. The detectors are directly submerged in liquid argon, which acts as cooling and shielding at the same time. The total mass in the first phase is $18\,\mathrm{kg}$. In a second phase, additional detectors will increase the total mass to about $40\,\mathrm{kg}$. The corresponding sensitivity reach is $T^{0\nu}_{1/2} = 1.5 \cdot 10^{26}\,\mathrm{a}$.

The **SNO+** experiment (see Section 1.3.3.2 and [70]) also aims at measuring the neutrino-less double beta decay. The liquid scintillator can be loaded with some $10\,\mathrm{kg}$ of ^{150}Nd. It will be able to improve the current limits for a neutrino-less double beta decay of ^{150}Nd by two orders of magnitude.

1.3.4.2 Neutrino mass

The **KATRIN** experiment [87] aims at measuring the effective mass of the ν_e by determining the spectral shape at the endpoint energy of the ^{3}H beta decay. The predecessor, the **Mainz experiment**, delivered the world-best upper limit for the ν_e mass of $m_{\nu_e} < 2.3\,\mathrm{eV}$ (95% C.L.) [62, 67]. Similar to the Mainz experiment, KATRIN will use a so-called MAC-E filter to measure the energy of the beta electrons. The MAC-E filter of the Mainz experiment will be used as a pre-spectrometer to reduce the rate of the electrons entering the main spectrometer.

The sensitivity expected after three years of data taking is $m_{\nu_e} < 0.2\,\mathrm{eV}$ (90% C.L.). The experiment is currently under construction.

1.3.4.3 Neutrino velocity anomaly

The OPERA experiment has measured the velocity of the ν_μ in the CNGS beam [7]. Precise measurements of both the baseline and the time of flight have been performed. The measured time is $\delta t = 57.8 \pm 7.8\,(\mathrm{stat.})^{+8.3}_{-6.9}\,(\mathrm{syst.})\,\mathrm{ns}$ shorter then the time a light signal would take for the same distance. The MINOS experiment performed a similar measurement already in 2007 [8]. Its result was compatible with neutrinos travelling at speed of light, while the best fit $\delta t = 126 \pm 32(\mathrm{stat.}) \pm 64(\mathrm{syst.})$ is also compatible with the OPERA result. Due to the difficult interpretation of such a result, the anomaly discovered by OPERA has to be confirmed or disproved by other neutrino beam experiments, in particular by a more precise measurement in MINOS.

Chapter 2

The OPERA experiment

The OPERA experiment was built to provide the first direct evidence of $\nu_\mu \longmapsto \nu_\tau$ oscillations by an appearance measurement. In this chapter, an overview of the OPERA experiment and its experimental techniques is given. In the first section the experimental reach is outlined. Section 2.2 describes the different components of the detector. In Section 2.3 the analysis procedure for neutrino events will be described. Finally, Section 2.4 shows first results of the experiment. The CNGS[1] beam used for the OPERA experiment has already been described in Section 1.3.2.4 of the previous chapter.

2.1 Physics performance

For a measurable quantity of oscillated neutrinos, the experiment has to be a *long-baseline experiment*. The CNGS beam used by OPERA has a mean neutrino energy of $E = 18\,\mathrm{GeV}$ and a baseline of $L = 730\,\mathrm{km}$, which is optimised for ν_τ-appearance[2]. The neutrinos are produced at CERN and the OPERA detector is located in the LNGS[3] underground laboratory in Italy.

[1] CERN Neutrinos to Gran Sasso

[2] The number of detected oscillated neutrinos does not depend strongly on the baseline L, as the oscillation probability increases with $\sin L$ while the number of neutrinos hitting the detector decreases with L^2 due to the beam divergence. A short L is not preferable as it increases the background of unoscillated neutrinos, while a too long L will actually decrease the number of oscillated neutrinos. As the oscillation probability also increases with $\sin \frac{1}{E}$, a low neutrino energy E is preferable. On the other hand, E must be above the τ production threshold.

[3] *Laboratori Nazionali del Gran Sasso*

Table 2.1: Expected number of observed ν_τ events for $22.5 \cdot 10^{19}$ p.o.t. as well as for the analysed sample of the data of the 2008 and 2009 runs as of July 2011. Full mixing and $\Delta m^2_{23} = 2.5 \cdot 10^{-3}\,\mathrm{eV}^2$ is assumed. [35]

Decay channel	Number of signal events expected for	
	$22.5 \cdot 10^{19}$ p.o.t.	the analysed sample
$\tau \to \mu$	1.79	0.39
$\tau \to e$	2.89	0.63
$\tau \to h$	2.25	0.49
$\tau \to 3h$	0.71	0.15
Total	7.63	1.65

Table 2.2: Expected number of background events for $22.5 \cdot 10^{19}$ p.o.t. as well as for the analysed sample of the data of the 2008 and 2009 runs as of July 2011. Specified errors are systematic. [35]

Decay channel	Number of background events expected for	
	$22.5 \cdot 10^{19}$ p.o.t.	the analysed sample
$\tau \to \mu$	0.09 ± 0.04	0.02 ± 0.01
$\tau \to e$	0.22 ± 0.05	0.05 ± 0.01
$\tau \to h$	0.24 ± 0.06	0.05 ± 0.01
$\tau \to 3h$	0.18 ± 0.04	0.04 ± 0.01
Total	0.73 ± 0.15	0.16 ± 0.03

At $E = 18\,\mathrm{GeV}$, the τ has a decay length in the order of $1\,\mathrm{mm}$. Therefore, a sub-millimetre track resolution is required for the OPERA experiment. As in previous experiments (CHORUS [80], DONuT [61]) this is achieved by the *emulsion cloud chamber* (ECC) technique described in the next section. In conjunction with the ECC technique, OPERA uses electronic detectors to locate neutrino interactions and gain complementary information like muon momenta and charges.

CNGS plans to deliver $22.5 \cdot 10^{19}\,\mathrm{p.o.t.}$ in total. In case of full mixing and $\Delta m^2_{23} = 2.5 \cdot 10^{-3}\,\mathrm{eV}^2$, 7.63 ν_τ events are expected (see Table 2.1). Due to the very

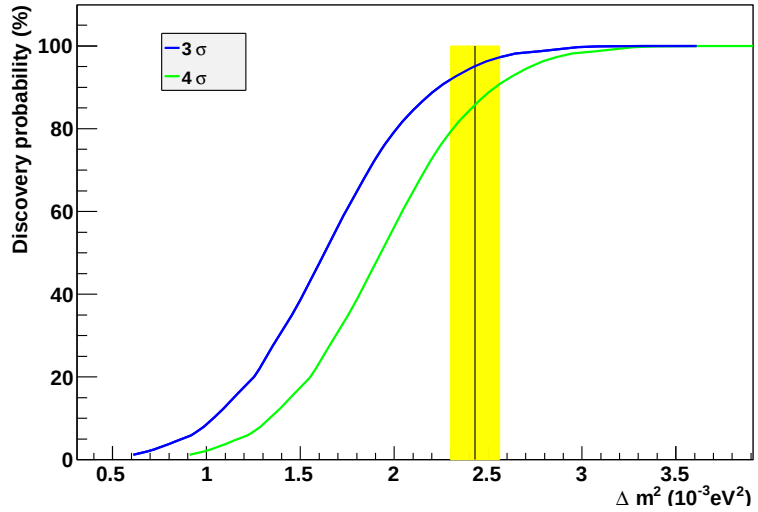

Figure 2.1: Discovery probability vs. Δm^2_{23} for OPERA, assuming full mixing. [68]

clean signature of ν_τ interactions in the ECC target, the expected background is only 0.73 ± 0.15 events (cf. Table 2.2). The discovery potential of OPERA depending on Δm^2_{23} and assuming full mixing is shown in Figure 2.1.

So far, OPERA has detected one ν_τ candidate event, which will be described in more detail in Section 2.4. Due to the low background, the single observed ν_τ event already presents evidence for a $\nu_\mu \rightarrow \nu_\tau$ oscillation with a significance of 95%. OPERA also has detected a number of ν_e candidates, which most likely come from the ν_e contamination of the CNGS beam (cf. Figure 3.5). The possibility to observe a ν_e appearance with OPERA is currently under reinvestigation, taking into account the recent hints for a large Θ_{13} (see Section 1.3.1.3).

2.2 Overview on the OPERA detector

A graphic overview on the OPERA detector is shown in Figure 2.2. The detector is divided into two identical *super modules* (SM). Each SM consists of a *target area* and a *muon spectrometer*. The ECC bricks (see Section 2.2.1) are housed in the target area. They are arranged in walls, which are interleaved with target tracker planes (see Section 2.2.2) to locate neutrino interactions. Downstream of each target area a muon spectrometer is located. On both sides of the OPERA detector is the *brick manipulation system*, which can extract single bricks from each target area. The total

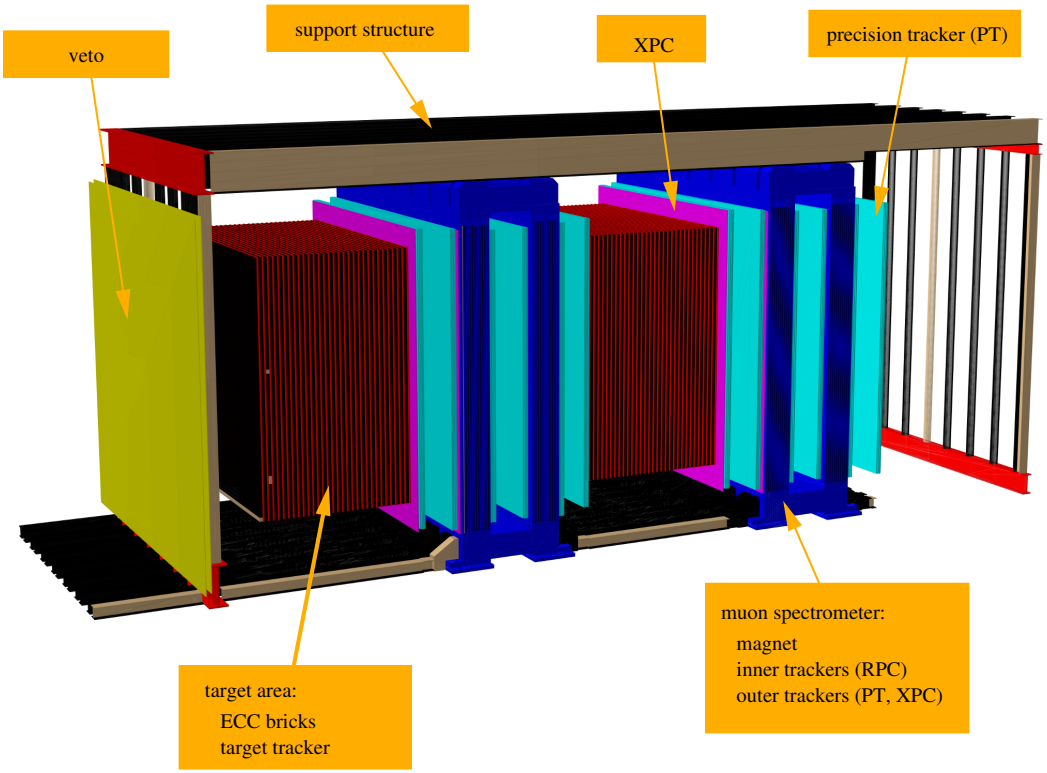

Figure 2.2: Schematic view of the OPERA detector. The two identical super modules each consist of a target area (red) and a muon spectrometer (blue). The muon spectrometer is complemented with the precision tracker (light blue) and the XPCs (purple). Upstream of the first target area is a veto detector (yellow). The brick manipulation system is not shown. The picture has been generated from the geometry information of the OPERA Monte Carlo software.

target mass of OPERA is about 1.25 kt, corresponding to a total number of 150,000 ECC bricks. In this section, the fundamental principle of the OPERA detector is described. For more information see [6].

2.2.1 ECC technique for ν_τ detection

The general signature of ν_τ interactions in the ECC target of OPERA is shown in Figure 2.3. To achieve the required track resolution, emulsion sheets are sandwiched with 1 mm thick lead plates. Each emulsion sheet has two 44 µm thick layers of nuclear emulsion on both sides of a 205 µm thick plastic base [21]. Each 56 lead plates and 57 emulsion sheets are packed into a so-called *brick* with the dimensions

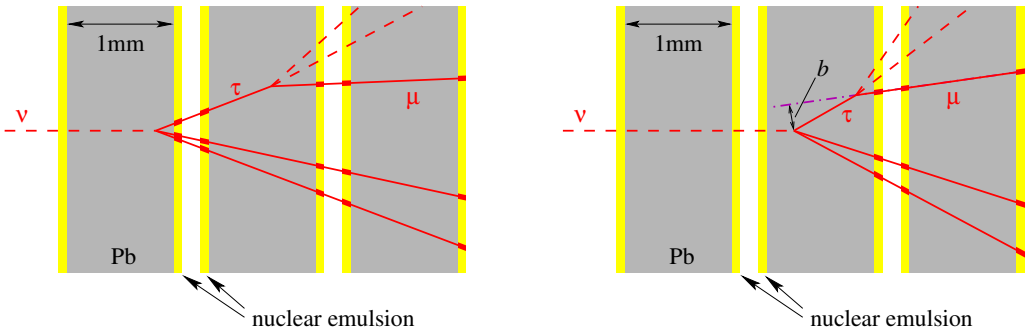

Figure 2.3: Schematic view of the ν_τ interaction signature in the OPERA ECC detector for the $\tau \to \mu$ decay channel. Two cases are to be discriminated, depending whether the τ decays in the next lead plate downstream of the interaction point (*long decay*, left image) or in the same lead plate (*short decay*, right image). In the first case, the signature is a kink in the leptonic track being directly visible. In the second case, the kink can only indirectly be observed by the impact parameter b.

$12.7 \times 10.2 \times 7.5\,\mathrm{cm}^3$, which corresponds to 10 radiation lengths in thickness. Each brick is equipped with a pair of emulsion sheets mounted on the downstream surface of the brick in a separated, light-tight compartment, called the *changeable sheet* (CS). The CS confirms the vertex prediction by the TT and delivers a more precise prediction of the vertex position inside the brick.

The OPERA emulsions have been specially designed for the requirements of the experiment. In contrast to previous experiments, they have been produced industrially by machines, therefore their homogeneity in thickness and sensitivity is unprecedented. Also a new technique called *refreshing* has been developed, which allows erasing tracks recorded before assembly of the bricks. Storing the emulsions at a relative humidity of 98% and a temperature of 27 °C erases more then 99% of the recorded tracks without affecting sensitivity [46].

After development, the emulsions need to be scanned. The thickness of the emulsions allows reconstruction of so-called *micro tracks* inside a single emulsion layer. Together with a matching micro track in the second emulsion layer of the same sheet a *base track* is formed.

As scanning the full area of the whole brick is not feasible, an event is reconstructed basically in three steps: First, to locate the neutrino interaction vertex, the tracks found in the CS are followed upstream, until they are no longer found in four consecutive sheets (*scan back*). Then on fifteen consecutive sheets an area of

Table 2.3: Selection criteria for ν_τ events and values for the first observed ν_τ candidate. [35]

Variable	Cut-off	Value for candidate
Missing p_T at primary vertex ($\mathrm{GeV/c}$)	< 1.0	$0.57^{+0.32}_{-0.17}$
Angle between parent track and primary hadronic shower in the transverse plane (rad)	$> \pi/2$	3.01 ± 0.03
Kink angle (mrad)	> 20	41 ± 2
Daughter momentum ($\mathrm{GeV/c}$)	> 2	12^{+6}_{-3}
Daughter p_T when γ-ray at the decay vertex ($\mathrm{GeV/c}$)	> 0.3	$0.47^{+0.24}_{-0.12}$
Decay length ($\mu\mathrm{m}$)	< 2 lead plates	1335 ± 35

$1\,\mathrm{cm}^2$ is scanned around the presumed neutrino vertex, five sheets upstream and ten sheets downstream (*volume scan*, *total scan*, TS). Any tracks originating from or near the vertex that have not yet been scanned are then followed downstream (*scan forth*).

The momentum of particles not hitting the spectrometer can be measured inside an ECC brick with *multiple Coulomb scattering* (MCS), as the scattering angles depend on the particle momentum. Monte Carlo studies and measurements at a π test beam have shown that the resolution can be better than 30% for momenta up to $8\,\mathrm{GeV/c^2}$. A comparison with μ hitting the spectrometer has confirmed this result. For more details see [2].

The decay search procedure searches for event topologies similar to those shown in Figure 2.3. In particular, it searches for a secondary vertex at which the τ decayed. In the majority of cases, the τ decays into an electron or into one or more hadrons. Therefore, the secondary vertex can be located as the origin of either a hadronic or an electromagnetic shower. In case of a decay into a μ, the track of the μ is followed until it stops if no clear μ identification is possible with the spectrometer. In case of a long decay topology (cf. Figure 2.3), the secondary vertex can be reconstructed by connecting with the track of the mother particle. In case of a short decay topology, only the impact parameter can be determined. In Table 2.3 the selection criteria for ν_τ events are shown. For comparison, also the values for the ν_τ candidate event are shown (cf. Section 2.4).

A background to the detection of ν_τ events can be the production of charmed mesons by beam ν_μ. D mesons have a similar lifetime and mass similar to the τ lepton [42]. D mesons can be produced by ν_μ CC DIS and thus can be discriminated from ν_τ interactions by the μ originating from the primary vertex. Only if the μ is not correctly identified, D mesons may contribute to the background (cf. Table 2.2). The cuts in Table 2.3 are chosen to select a clean sample of ν_τ events without a significant background contribution from charmed mesons. Additionally, the analysis of the decay products can sometimes identify D mesons, e.g. by observing a μ^+.

2.2.2 Electronic detectors

To locate neutrino interactions in the target, the target tracker (TT) is integrated in the target area. The TT is a plastic scintillator tracking detector with a granularity of $2.6 \times 2.6\,\mathrm{cm}^2$. Each scintillator strip is read out with a wavelength shifting fibre connected to multi-channel photo-multiplier tubes. Despite the TT not being specifically designed for this task, it can be used as a calorimeter for electromagnetic or hadronic showers with a low precision. For a higher precision, a smaller granularity would be required, which stands in contrast to the required dimensions of the ECC bricks.

The spectrometer downstream of each target area consists of a magnet with $1.52\,\mathrm{T}$ field intensity and inner and outer tracking detectors. The inner tracking detectors are planes of *resistive plate chambers* (RPC) and are located in between the slaps of the magnet iron yoke. The outer detectors are each two planes of drift tubes upstream and downstream of the magnet and in between the magnet arms. Additionally, there are layers of resistive plate chambers with tilted read-out strips (XPC) between the target area and the spectrometer, to disentangle ambiguities in multi-track events. The inner detector allows a direct measurement of the track curvature and provides the connection between the track segments. The RPC and XPC also function as a trigger for the drift time measurement of the outer tracker, which allows a more precise measurement of the bending angle and is therefore called the *precision tracker* (PT).

Upstream of the first target area is a glass RPC detector acting as a veto to tag neutrino interactions in the surrounding rock or other material upstream of OPERA (e.g. the Borexino detector). Figure 2.4 shows a visualisation of two neutrino events in the detector.

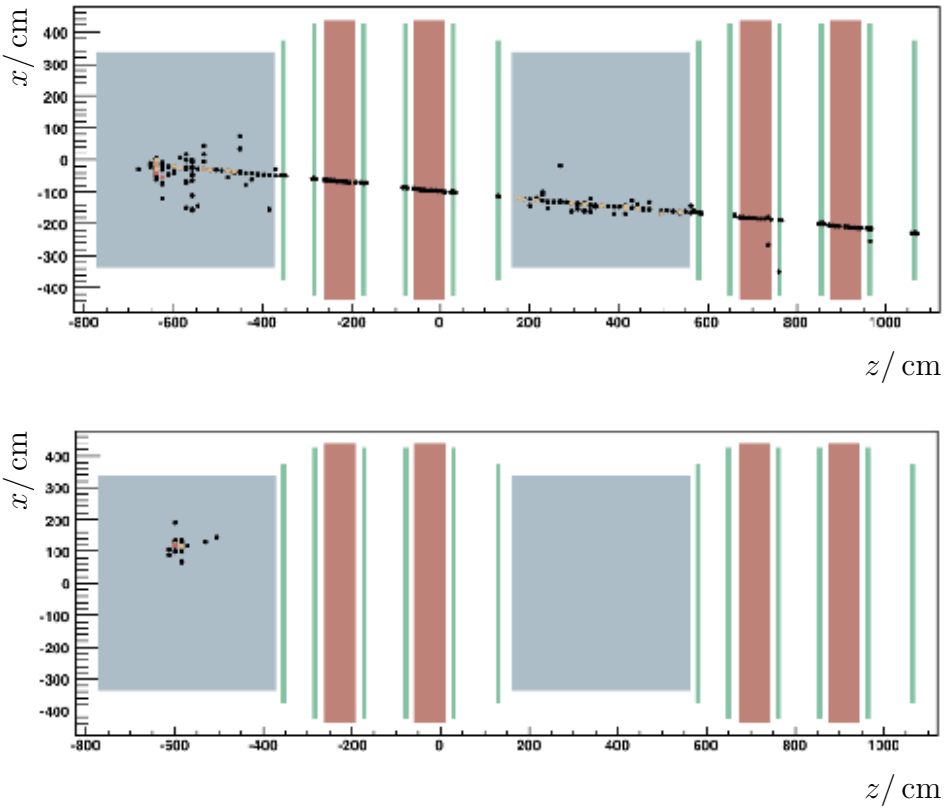

Figure 2.4: Visualisation of a charged current (top) and a neutral current (bottom) event in the OPERA electronic detector. The target area with the TT is shown in grey, the magnet arms with the RPC in red, and the PT in green. Detector hits are shown as dots. The colour of the dots in the TT encodes the deposited energy (black = low, brown = medium, red = high). The bending of the particle trajectory in the magnetic field is below the resolution of this display.

2.3 Analysis procedure for neutrino events

Neutrino interactions in the target are located by the TT. To reduce background induced by cosmic μ, only events within a time window of $\pm 20\,\mu s$ to the calculated arrival time of the neutrino bunches are considered. The events are also classified whether the interaction was inside one of the target regions – which is called a *contained* event – or the reaction took place in the magnet iron, the support structure, the surrounding rock, or some other material outside the target region. Only contained events are used for the oscillation search.

To increase the efficiency in finding the interaction brick, the three most probable bricks are determined. Starting with the most probable brick, the bricks are extracted from the target area by an automated *brick manipulation system* (BMS). Then the brick is exposed to X-rays, so that the CS and the most downstream film of the brick are marked for alignment between CS and the brick. The CS is then developed and analysed. If the CS confirms that the brick contains the neutrino interaction point, the brick will be further analysed. Otherwise, it gets a new CS attached and is put back into the target area. The next probable brick is then extracted.

Once the correct brick is found, it is exposed to the hard component of cosmic rays for alignment between the emulsion sheets. It will then be disassembled, and the emulsions are developed. The emulsions are sent to one of the scanning stations in Europe and Japan.

The scanning stations each consist mainly of an optical microscope with an attached CMOS[4] camera. The emulsion is placed on a movable stage, so that the whole area can be scanned. In detail, the European and the Japanese stations work slightly different. The European system moves the stage to a position, that the desired field of view can be scanned. Then the camera scans the depth by moving in vertical direction while the stage is in rest. The Japanese system avoids stopping the stage, by moving also the optical system with the same constant speed as the stage, while scanning in depth. Both systems deliver a comparable resolution (about $0.3\,\mu\mathrm{m/pixel}$). The data is processed and analysed by a computer cluster in each scanning lab. All data is then shared throughout the collaboration by storing it in a central data base.

In case of a candidate event, more bricks are extracted and analysed to follow all tracks to their stopping points. This allows a good particle identification (e.g. separation of π from μ).

2.4 First results

In the sample of data collected in 2008 and 2009, OPERA has detected one ν_τ candidate [11]. The candidate event (as shown in Figure 2.5) is fulfilling all selection criteria shown in Table 2.3. The event is a muon-less neutrino interaction with seven tracks originating from the primary vertex. One of the tracks has a visible

[4]Complementary metal–oxide–semiconductor, a technique for the construction of integrated circuits, frequently also used for image sensors.

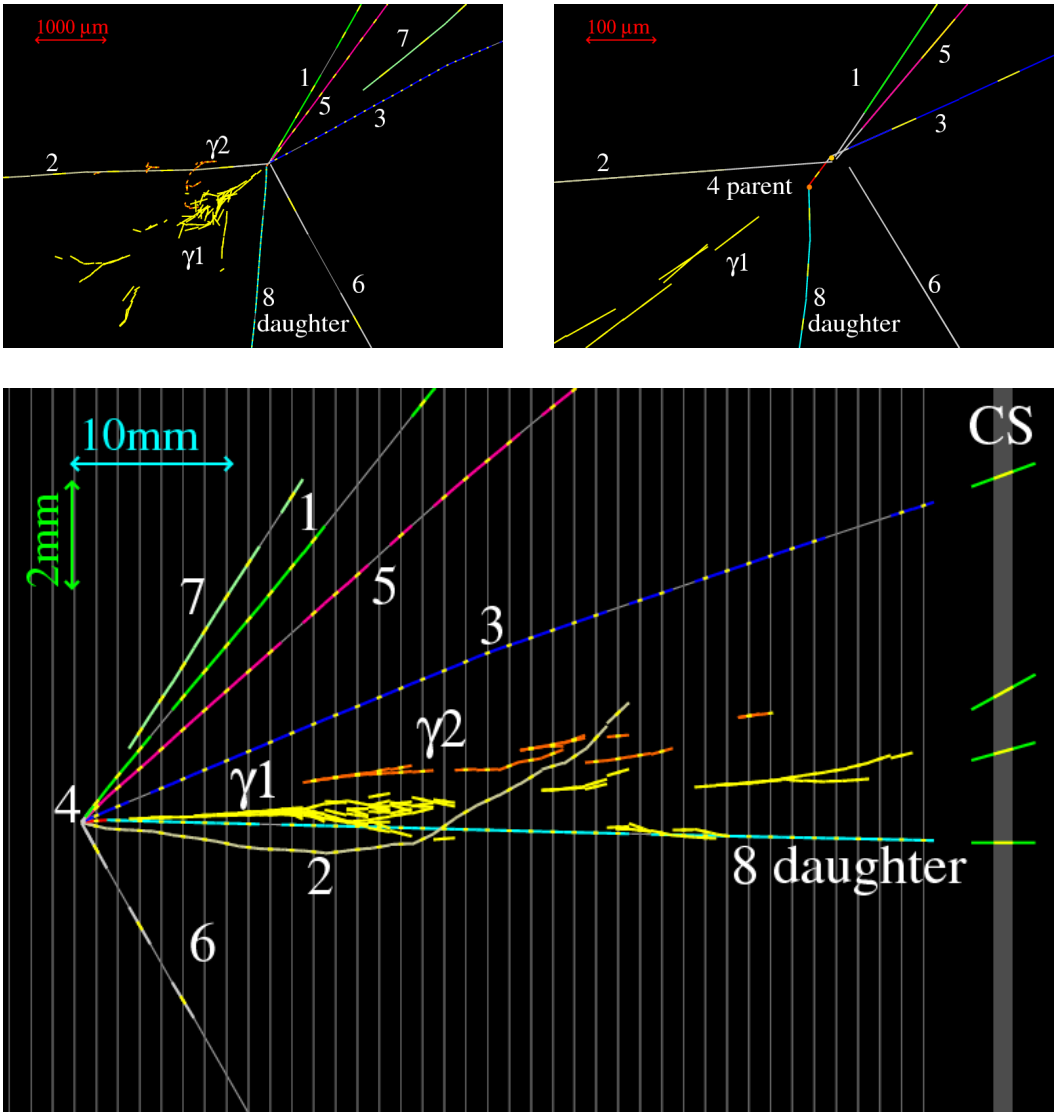

Figure 2.5: Visualisation of the ν_τ candidate event. Track 4 is the τ^-, track 8 is the daughter π^-. $\gamma 1$ and $\gamma 2$ are the electromagnetic showers caused by the two γ from the π^0 decay at the secondary vertex. The other tracks are described in [11], where the picture is taken from.

kink after $(1335 \pm 35)\,\mu m$ with an angle of $(41 \pm 2)\,mrad$. The daughter track has a high momentum ($(12^{+6}_{-3})\,GeV$) and generates a two-prong interaction seven walls downstream of the primary vertex. The impact parameter of the daughter track is $(55 \pm 4)\,\mu m$, while the tracks attached to the primary vertex match within $7\,\mu m$.

Additionally, there are two γ-rays, detected via their electromagnetic showers. The first shower is originating at a distance of $2.2\,mm$ from the secondary vertex. It is compatible with pointing to the secondary vertex with a probability of 32%, while the probability of pointing to the primary vertex is less then 10^{-3}. The second shower has a distance of about $13\,mm$ to the vertices and is compatible with pointing to both vertices, with a higher probability of pointing to the secondary vertex (82% vs. 10%). The total invariant mass of both γ-rays is $(120 \pm 20(stat.) \pm 35(syst.))\,MeV/c^2$, suggesting they originated from a π^0 decay.

Together with the invariant mass of the charged daughter particle (presumably a π^-) the total invariant mass of the decay products amounts to $(640^{+125}_{-80}(stat.)^{+100}_{-90}(syst.))\,MeV/c^2$ and is compatible with the mass of a $\rho(770)$. The presumed decay it therefore:

$$\tau \to \rho^- \nu_\tau \to \pi^0 \pi^- \nu_\tau.$$

The branching ratio of this decay is about 25% of all τ decays.

As discussed in Section 2.1, this single event presents evidence for a $\nu_\mu \to \nu_\tau$ oscillation with a confidence level of 95%. It is therefore important to observe more ν_τ events for providing statistically significant evidence.

Chapter 3

Cross sections for relevant neutrino interactions in OPERA

The analysis described in this thesis uses Monte Carlo data for the training of the neural network and for comparison with real data. The standard Monte Carlo event generator used for OPERA, NEGN, generates events separated into different classes depending on neutrino flavour, interaction type, scattering process, and – in case of ν_τ – the τ decay channel. As it does not provide information about the relative abundances of the classes, the relative interaction rates of the different event classes for the OPERA experiment are calculated in this chapter. Additionally, a theoretical expectation for the NC ratio is provided for comparison with the measured value in Chapter 5.

The different interaction types and scattering processes are divided into several categories. First, depending on the charge of the exchanged $W^{\pm}/Z^0$ boson, the interaction can be of *charged* or *neutral current* type (in future: CC or NC, respectively). Additionally, depending on the momentum transferred, the interactions are classified as *deep inelastic scattering* (DIS), *quasi elastic scattering* (QEL) or – if the process involves some intermediate state of the target nucleon – *resonant scattering* (RES).

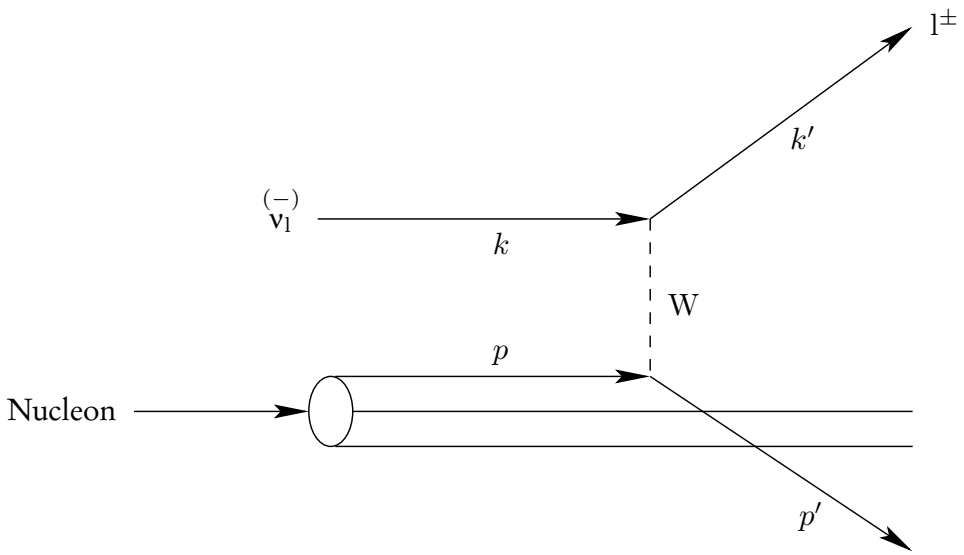

Figure 3.1: Feynman diagram of $\overset{(-)}{\nu_l}$ N $\longmapsto$ $l^{\pm}$X scattering. The incoming parton carries the momentum p and the incoming lepton the momentum k. After the scattering process the lepton carries the momentum k' and the parton p'. Fragmentation is not included in this picture.

3.1 Charged current deep inelastic scattering of neutrinos

3.1.1 Definition of kinematical variables

Cross sections of neutrino interactions are usually described by Lorentz-invariant kinematical variables as shown in Figure 3.1. The initial state four-momenta are k and p for the lepton and the parton, respectively. For the final state, the symbols are primed (k', p'). Thus the momentum transferred is $q = k - k' = p - p'$. The square of the four-momentum transferred is $Q^2 = -q^2 = -(k - k')^2$. The energy transferred is $\nu = E_\nu - E_l$, where E_ν is the energy of the incoming neutrino and E_l is the energy of the outgoing lepton. M is the rest mass of the nucleon, and m_l is the rest mass of the final state lepton. The neutrino mass is neglected in this calculation.

In the following, the Bjorken scaling variable $x = \frac{Q^2}{pq} = \frac{Q^2}{2M\nu}$ will be used, which can be interpreted as the contribution of the parton to the longitudinal momentum of the nucleon. Another important variable is the inelasticity $y = \frac{pq}{pk} = \frac{\nu}{E_\nu}$, describing the amount of energy transferred to the hadronic system relative to the available leptonic energy in the target rest frame. Using x and y, the four-momentum can be written as $Q^2 = 2xyME_\nu$.

3.1.2 Double-differential cross section

The double-differential cross section for charged current deep inelastic scattering can be calculated from the structure functions F_i. The calculations in this section follow [79, 86] and are executed in leading order approximation:

$$\begin{aligned}\frac{d^2\sigma}{dxdy} =& \frac{G_F^2 M E_\nu}{\pi}\left[y\left(xy+\frac{m_l^2}{2E_\nu M}\right)F_1\right.\\ &+\left(1-y-\frac{Mxy}{2E_\nu}-\frac{m_l^2}{4E_\nu^2}\right)F_2\\ &\pm\left(xy\left(1-\frac{y}{2}\right)-y\frac{m_l^2}{4ME_\nu}\right)F_3\\ &\left.+\left(xy\frac{m_l^2}{2ME_\nu}+\frac{m_l^4}{4M^2E_\nu^2}\right)F_4-\frac{m_l^2}{2ME_\nu}F_5\right],\end{aligned} \tag{3.1}$$

with the Fermi constant G_F. The sign of the F_3 coefficient corresponds to the neutrino (+) or the anti-neutrino (-) case.

In the parton model, the structure functions can be obtained from the parton distribution functions (*PDF*). As protons and neutrons are treated separately, non-isoscalar targets can be accounted for. If charm production is enabled (i.e. above the production threshold of charmed mesons), the structure functions for $\nu/\bar{\nu}$–proton scattering are

$$\begin{aligned}F_2^{(\nu p)} &= 2x\left[d+s+\overline{u}+\overline{c}\right],\\ xF_3^{(\nu p)} &= 2x\left[d+s-\overline{u}-\overline{c}\right],\\ F_2^{(\bar{\nu} p)} &= 2x\left[u+c+\overline{d}+\overline{s}\right],\\ xF_3^{(\bar{\nu} p)} &= 2x\left[u+c-\overline{d}-\overline{s}\right],\end{aligned} \tag{3.2}$$

where u,d,s,c are the PDFs of the respective quarks for the proton and $\overline{u}$,$\overline{d}$,$\overline{s}$,$\overline{c}$ are the PDFs of the anti-quarks.

Due to isospin symmetry, the PDFs for the neutrons are the same as those of the proton but with flipped isospin. Thus using the proton PDF the structure functions for $\nu/\bar{\nu}$-neutron scattering are

$$\begin{aligned}F_2^{(\nu n)} &= 2x\left[u+s+\overline{d}+\overline{c}\right],\\ xF_3^{(\nu n)} &= 2x\left[u+s-\overline{d}-\overline{c}\right],\\ F_2^{(\bar{\nu} n)} &= 2x\left[d+c+\overline{u}+\overline{s}\right],\\ xF_3^{(\bar{\nu} n)} &= 2x\left[d+c-\overline{u}-\overline{s}\right].\end{aligned} \tag{3.3}$$

Below the charm production threshold (or if charm production should be disabled for another reason) the structure functions have to be modified by the Cabibbo angle θ_c, so one obtains for the proton

$$\begin{aligned} F_2^{(\nu p)} &= 2x\left[d\cos^2\theta_c + s\sin^2\theta_c + \overline{u} + \overline{c}\right], \\ xF_3^{(\nu p)} &= 2x\left[d\cos^2\theta_c + s\sin^2\theta_c - \overline{u} - \overline{c}\right], \\ F_2^{(\overline{\nu} p)} &= 2x\left[u\cos^2\theta_c + c\sin^2\theta_c + \overline{d} + \overline{s}\right], \\ xF_3^{(\overline{\nu} p)} &= 2x\left[u\cos^2\theta_c + c\sin^2\theta_c - \overline{d} - \overline{s}\right] \end{aligned} \tag{3.4}$$

and for the neutron

$$\begin{aligned} F_2^{(\nu n)} &= 2x\left[u\cos^2\theta_c + s\sin^2\theta_c + \overline{d} + \overline{c}\right], \\ xF_3^{(\nu n)} &= 2x\left[u\cos^2\theta_c + s\sin^2\theta_c - \overline{d} - \overline{c}\right], \\ F_2^{(\overline{\nu} n)} &= 2x\left[d\cos^2\theta_c + c\sin^2\theta_c + \overline{u} + \overline{s}\right], \\ xF_3^{(\overline{\nu} n)} &= 2x\left[d\cos^2\theta_c + c\sin^2\theta_c - \overline{u} - \overline{s}\right]. \end{aligned} \tag{3.5}$$

To calculate the exclusive charm production cross section, the structure functions are modified in an analogue way:

$$\begin{aligned} F_2^{(\nu p)} &= 2x\left[d\sin^2\theta_c + s\cos^2\theta_c + \overline{u} + \overline{c}\right], \\ xF_3^{(\nu p)} &= 2x\left[d\sin^2\theta_c + s\cos^2\theta_c - \overline{u} - \overline{c}\right], \\ F_2^{(\overline{\nu} p)} &= 2x\left[u\sin^2\theta_c + c\cos^2\theta_c + \overline{d} + \overline{s}\right], \\ xF_3^{(\overline{\nu} p)} &= 2x\left[u\sin^2\theta_c + c\cos^2\theta_c - \overline{d} - \overline{s}\right] \end{aligned} \tag{3.6}$$

and for the neutron:

$$\begin{aligned} F_2^{(\nu n)} &= 2x\left[u\sin^2\theta_c + s\cos^2\theta_c + \overline{d} + \overline{c}\right], \\ xF_3^{(\nu n)} &= 2x\left[u\sin^2\theta_c + s\cos^2\theta_c - \overline{d} - \overline{c}\right], \\ F_2^{(\overline{\nu} n)} &= 2x\left[d\sin^2\theta_c + c\cos^2\theta_c + \overline{u} + \overline{s}\right], \\ xF_3^{(\overline{\nu} n)} &= 2x\left[d\sin^2\theta_c + c\cos^2\theta_c - \overline{u} - \overline{s}\right]. \end{aligned} \tag{3.7}$$

The remaining structure functions F_1, F_4 and F_5 can be obtained by the Callan-Gross relation:

$$2xF_1 = F_2$$

and the Albright-Jarlskog relations [16]:

$$\begin{aligned} F_4 &= 0 \\ xF_5 &= F_2 \end{aligned}$$

With these relations equation 3.1 simplifies to

$$\frac{d^2\sigma}{dxdy} = \frac{G_F^2 M E_\nu}{\pi}\left[\left(1-y+\frac{y^2}{2}+\frac{y m_l^2}{4xE_\nu M}-\frac{Mxy}{2E_\nu}-\frac{m_l^2}{4E_\nu^2}-\frac{m_l^2}{2xME_\nu}\right)F_2 \pm\left(xy\left(1-\frac{y}{2}\right)-y\frac{m_l^2}{4ME_\nu}\right)F_3\right]. \tag{3.8}$$

3.1.3 Correction of target and quark masses

The structure functions are only valid for vanishing masses of both the target nucleon and the final state quarks. To correct for the mass of the target nucleon (commonly referred to as the target mass), the Bjorken variable x has to be replaced in the structure functions with the Nachtmann variable [74] [64]

$$\xi^{\mathrm{TMC}} = \frac{2x}{1+\sqrt{1+\frac{4x^2M^2}{Q^2}}}. \tag{3.9}$$

For interaction channels in which only light quarks are produced (e.g. pion production), the quark mass can safely be neglected. On the other hand, the mass of the final state quark has to be considered when investigating charm production. Similar to the target mass correction, the quark mass m_q can be taken into account by replacing the Bjorken x with the slow rescaling variable [30]

$$\xi^{\mathrm{QMC}} = x\left(1+\frac{m_q^2}{Q^2}\right). \tag{3.10}$$

Equations 3.9 and 3.10 can be combined to correct both target and quark masses:

$$\xi = \frac{2x}{1+\sqrt{1+\frac{4x^2M^2}{Q^2}}}\left(1+\frac{m_q^2}{Q^2}\right). \tag{3.11}$$

This replacement is done in the structure functions F_i only, i.e. $F_i(x) \rightarrow F_i(\xi)$. Equation 3.8 remains otherwise unchanged.

3.1.4 Integration boundaries and cuts

The double-differential cross section has to be integrated over x and y. For this purpose the correct integration boundaries have to be obtained first. There are different constraints that have to be taken into account, both physical and technical.

For ν_μ interactions the mass of the final state lepton may be neglected, but numerical available PDFs are usually not defined for very small x. Hence, also for ν_μ interactions the lower limit in x is derived from the (final state) leptonic mass m_l.

The kinematically allowed region is [16][1]:

$$\begin{aligned} \frac{m_l^2}{2M\left(E_\nu - m_l\right)} \leq x \leq 1, \\ a - b \leq y \leq a + b, \end{aligned} \tag{3.12}$$

where

$$\begin{aligned} a &= \frac{1 - m_l^2\left(\frac{1}{2ME_\nu x} + \frac{1}{2E_\nu^2}\right)}{2\left(1 + \frac{Mx}{2E_\nu}\right)}, \\ b &= \frac{\sqrt{\left(1 - \frac{m_l^2}{2ME_\nu x}\right)^2 - \frac{m_l^2}{E_\nu^2}}}{2\left(1 + \frac{Mx}{2E_\nu}\right)}, \end{aligned} \tag{3.13}$$

with $m_l = 105.66\,\mathrm{MeV}$ for ν_μ and $m_l = 1776.82\,\mathrm{MeV}$ for ν_τ interactions [42].

Additionally, a cut on the invariant mass of the hadronic final state $W = \sqrt{M^2 - Q^2 + 2MyE_\nu}$ with $W \geq 1.4\,\mathrm{GeV}$ is applied, giving an additional x-dependent lower limit in y. This cut separates DIS from QEL and RES. Lower invariant masses require dedicated calculations. For charm production the cut on W is determined by the mass of the lightest charmed meson (i.e. the D^0 mass of $1864.83\,\mathrm{MeV}$ [42]), as no charm production is possible below this threshold.

For the numerical calculation of the cross sections the CT10 parton distributions are used [66]. These parton distributions are not only depending on x (or ξ) but also on Q, to account for scaling violating effects. Like any other numerical PDF set CT10 has a lower limit on Q: For $Q < 0.3\,\mathrm{GeV}$ the used CT10 PDF set is not defined. In leading order calculations of total cross sections, the PDFs are usually extrapolated for lower Q by setting a constant Q. As Equation (3.1) has a peak, which is shifted to the region $Q < 0.3\,\mathrm{GeV}$ if the quark mass is increased to the mass of the charm quark (cf. Equation (3.10)), this extrapolation creates an artificial peak in the resulting charm production cross section. Therefore, in these calculations the PDFs are set to 0 for $Q < 0.3\,\mathrm{GeV}$. However, leading order QCD is not applicable to small Q, thus for a more precise calculation of the cross section for small Q, next-to leading order calculations are required, also avoiding the need for

[1] with correction of a typographic error [63]

this extrapolation. Due to the low statistics present in the OPERA experiment this precision will not be required for this analysis.

3.1.5 Numerical calculation of cross sections

With a known target mass $M = 0.938\,\mathrm{GeV}$ [42] (the difference between the proton and the neutron mass shall be neglected) and a fixed neutrino energy E_ν, the double-differential cross sections can be plotted in the x-y plane. This has been performed for different energies, in Figure 3.2 for ν_μ, and in Figure 3.3 for ν_τ interactions.

The integration of the cross sections is done numerically using the NINTEGRATE function of Mathematica[2]. The inclusive CC ν_μ DIS, the exclusive charm production and the CC ν_τ DIS integrated cross sections are shown in Figure 3.4 as a function of E_ν. As expected, the charm production and the ν_τ cross sections are smaller than the inclusive CC ν_μ DIS cross section at lower energies.

To obtain the event ratio of the number of charm productions per number of CC ν_μ DIS events for the OPERA experiment, the calculated cross sections need to be integrated over the neutrino energy, weighted with the flux of the CNGS neutrino beam. As there is no near neutrino detector, the flux is known from Monte Carlo simulations [44], muon fluxes at the neutrino production site [23] and experience with similar experiments (in particular the WANF beam [41]) only. The neutrino spectrum used in this calculation is the same as the one used by the Monte Carlo generator NEGN, which is the standard event generator used for the OPERA experiment (see Figure 3.5).

Dividing these calculated cross sections by each other leads to the ratio of charm production from CC ν_μ DIS:

$$R^{(\mathrm{charm})} = \frac{\sigma^{(\mathrm{charm})}}{\sigma^{(\mathrm{CC}\ \nu_\mu\ \mathrm{DIS})}} = \frac{6.27 \cdot 10^{-39}\,\mathrm{cm}^2}{1.35 \cdot 10^{-37}\,\mathrm{cm}^2} = 0.047. \tag{3.14}$$

The charm production ratio compared to all events in the OPERA detector will be smaller due to other possible interaction types.

In a similar way, the ratio of ν_τ interactions can be calculated, here with an assumed two-flavour oscillation (see Section 1.2) with $\Delta m^2 = 2.4 \cdot 10^{-3}\,\mathrm{eV}^2$, $\sin 2\Theta = 1$ and $L = 732\,\mathrm{km}$:

$$R^{(\mathrm{tau})} = \frac{\sigma^{(\mathrm{CC}\ \nu_\tau\ \mathrm{DIS})}}{\sigma^{(\mathrm{CC}\ \nu_\mu\ \mathrm{DIS})}} = \frac{5.90 \cdot 10^{-40}\,\mathrm{cm}^2}{1.35 \cdot 10^{-37}\,\mathrm{cm}^2} = 0.0044. \tag{3.15}$$

[2] Mathematica is a computational software developed by Wolfram Research. The calculations in this thesis were performed using version 7.

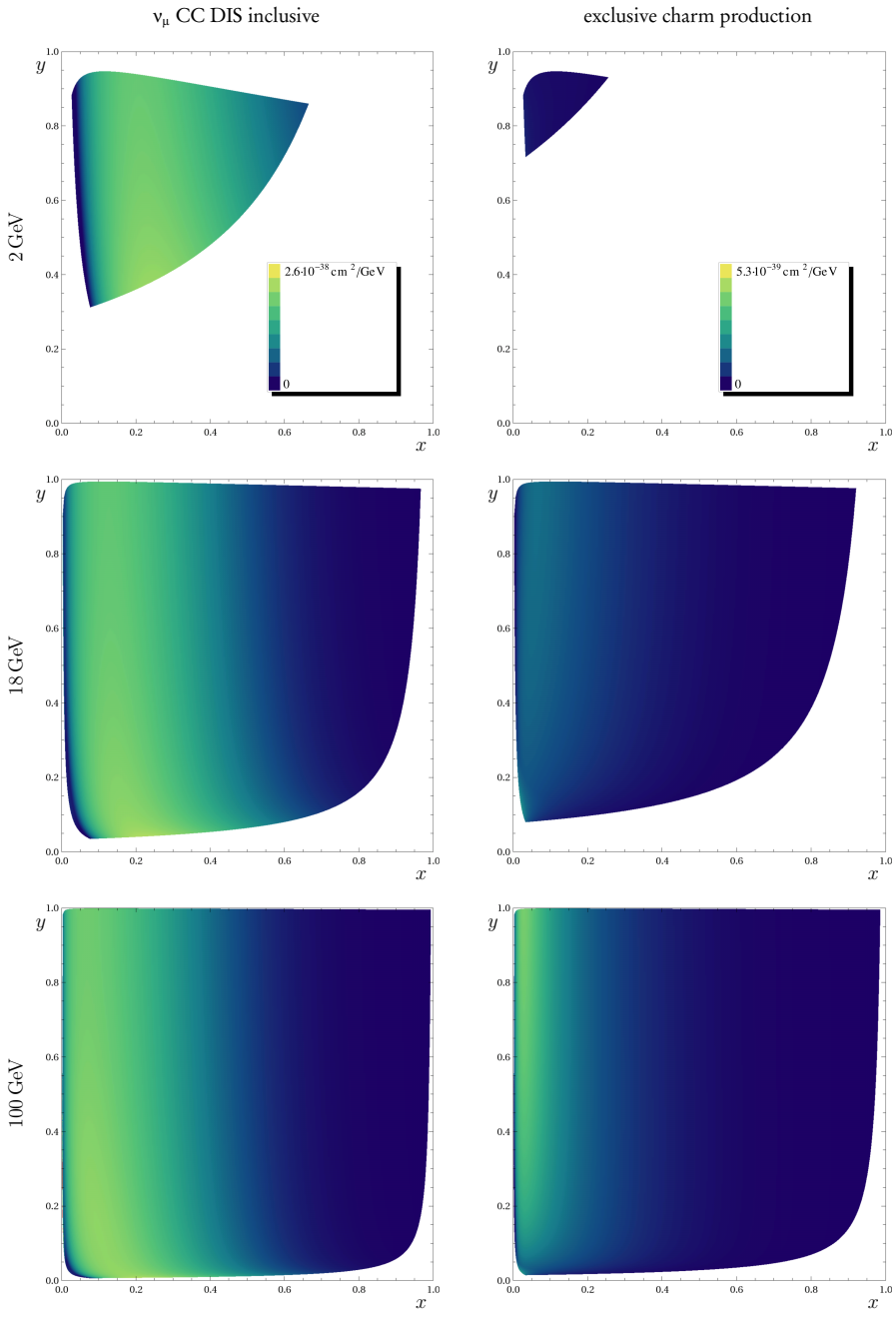

Figure 3.2: Double-differential cross sections for ν_μ CC DIS inclusive reactions (left) and charm production (right, note the different scale). The cross section is plotted for different neutrino energies: 2 GeV (top), 18 GeV (middle), and 100 GeV (bottom). 2 GeV is right above the charm production threshold, thus the cross section for charm production is very small. 18 GeV is the mean energy of the CNGS beam. As expected, the charm production is more concentrated at higher inelasticities y.

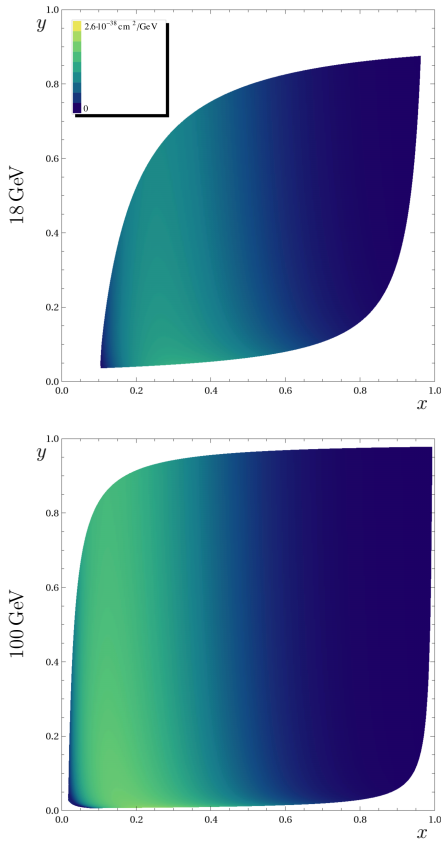

Figure 3.3: Double-differential cross sections for ν_τ CC DIS inclusive reactions, for neutrino energies of 18 GeV (top) and 100 GeV (bottom). 2 GeV is below the τ production threshold.

This ratio compares only DIS. Due to the larger leptonic mass in the final state, contributions from RES and QEL are larger for ν_τ than for ν_μ interactions. In the detector, DIS, RES, and QEL are not easily discriminated. Therefore, the observable ratio of ν_τ to ν_μ interactions will be larger (see next section).

3.2 Contributions from resonant and quasi-elastic scattering, and neutral current interactions

As the theoretical description of resonant and quasi-elastic scattering is rather complicated – and a semi-analytic calculation like in Section 3.1 would therefore go beyond the scope of this thesis – contributions from those processes are calculated relative to the respective DIS cross sections using the Monte Carlo neutrino event

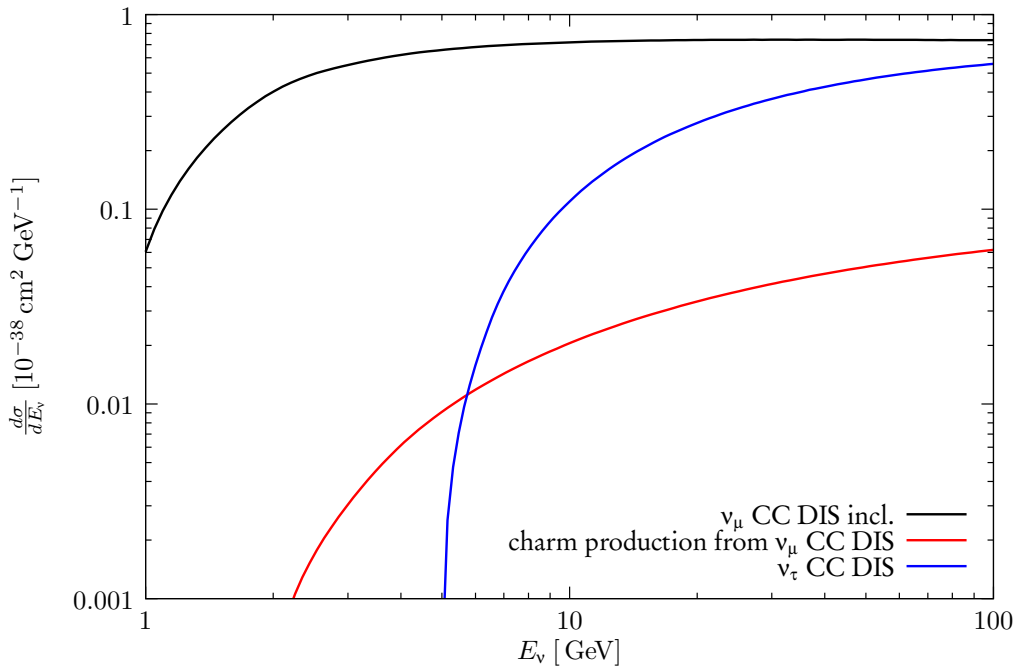

Figure 3.4: Single-differential cross sections for inclusive ν_μ CC DIS, for charm production from ν_μ and for ν_τ CC DIS vs. neutrino energy.

Table 3.1: NC/CC ratios of neutrino events in OPERA for different neutrino flavours. These values have been computed using the GENIE event generator and include all interaction types. For ν_τ, the oscillated ν_μ spectrum has been used with $\Delta m = 2.4 \cdot 10^{-3}\,\mathrm{eV}^2$. As the oscillation is more likely for energies below the τ production threshold, the NC/CC ratio is increased for ν_τ.

	ν_μ	$\overline{\nu}_\mu$	ν_e	ν_τ
NC/CC	0.296	0.379	0.296	1.00

Table 3.2: Relations of the DIS, RES, and QEL interaction rates for each neutrino flavour and interaction type in OPERA. Contributions from elastic scattering are neglected, as they can be neither observed with the OPERA detector nor simulated with NEGN. NC ν_τ events cannot be distinguished from NC ν_μ events and therefore are not simulated separately.

	DIS	RES	QEL
ν_μ CC	0.899	0.062	0.039
ν_μ NC	0.921	0.079	-
$\bar{\nu}_\mu$ CC	0.851	0.106	0.043
$\bar{\nu}_\mu$ NC	0.887	0.113	-
ν_e CC	0.929	0.043	0.027
ν_e NC	0.944	0.056	-
ν_τ CC	0.682	0.185	0.132

Table 3.3: Relations of the CC interaction rates of the different neutrino flavours in OPERA. $\bar{\nu}_e$ will be neglected in the analysis due to its small abundance. ν_τ rates are calculated separately. These numbers are from [44].

	ν_μ	$\bar{\nu}_\mu$	ν_e	$\bar{\nu}_e$
ν_i/ν_μ	1	0.024	0.0089	0.0006

Table 3.4: Branching ratio of the τ decay. Decay channels are combined to classes in the same way as the NEGN combines generated events to classes. Channels with more then three charged particles are added to the 3-prong channel, as NEGN cannot generate such events. Numbers are based on [42].

$\tau \to \mu\nu\nu$	$\tau \to e\nu\nu$	$\tau \to h\nu\nu$	$\tau \to hhh\nu\nu$
0.1471	0.1512	0.5489	0.1528

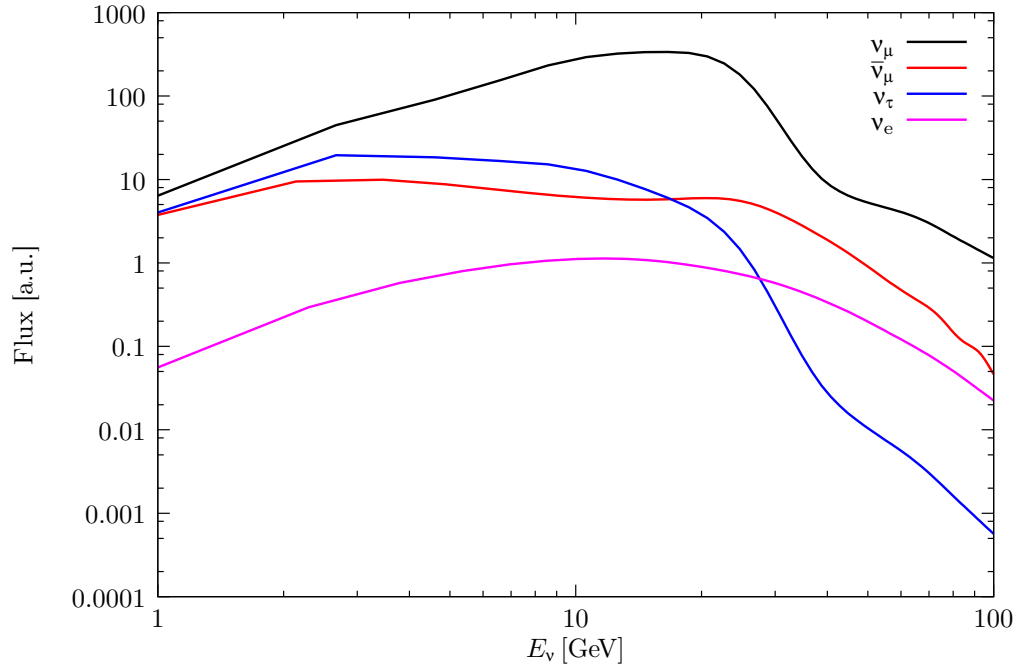

Figure 3.5: Neutrino spectrum of the CNGS beam for the different neutrino types. The mean energy is around 18 GeV for ν_μ. The spectra for ν_μ, $\bar{\nu}_\mu$, and ν_e are taken from the NEGN event generator, the ν_τ spectrum has been calculated from the ν_μ spectrum and two-flavour neutrino oscillations ($\Delta m^2 = 2.4 \cdot 10^{-3}\,\mathrm{eV}^2$, $\sin^2 2\Theta = 1$, $L = 732\,\mathrm{km}$). The spectra are available in 1 GeV bins only, causing kinks in this visualisation around 3 GeV. For the actual calculations, a different kind of interpolation is used.

generator *GENIE* [18]. GENIE can produce events for a given flux of a certain neutrino flavour with realistic contributions from the different interaction processes. The proportions are then extracted by simply counting events tagged with the respective interaction processes. This way, also the contributions from neutral current interactions can be determined, as shown in Table 3.1. The resulting ratio of the different scattering processes for each neutrino flavour and interaction type is shown in Table 3.2.

Combining these results with the neutrino flavour ratios from beam MC (Table 3.3) and the branching ratios of the τ decay (Table 3.4), realistic relative abundances of each NEGN event class can be calculated. The result of these calculations are shown in Table 3.5. These values will be used in Chapter 5 to weight the MC events in the analysis.

Table 3.5: Abundances of the event classes produced by NEGN, relative to ν_μ CC DIS without charm production. NC ν_τ events cannot be distinguished from NC ν_μ events and therefore are not simulated separately. τ decay channels are specified in brackets.

		ν_e	ν_μ	$\overline{\nu}_\mu$
CC	DIS	$9.2 \cdot 10^{-3}$	w/o charm: 1.000 charm only: 0.047	$22.7 \cdot 10^{-3}$
	RES	$0.43 \cdot 10^{-3}$	0.068	$2.8 \cdot 10^{-3}$
	QEL	$0.27 \cdot 10^{-3}$	0.043	$1.1 \cdot 10^{-3}$
NC	DIS	$2.8 \cdot 10^{-3}$	0.303	$9.0 \cdot 10^{-3}$
	RES	$0.16 \cdot 10^{-3}$	0.026	$1.1 \cdot 10^{-3}$

		ν_τ (e)	ν_τ (μ)	ν_τ (h)	ν_τ (hhh)
CC	DIS	$665 \cdot 10^{-6}$	$647 \cdot 10^{-6}$	$2.42 \cdot 10^{-3}$	$672 \cdot 10^{-6}$
	RES	$181 \cdot 10^{-6}$	$176 \cdot 10^{-6}$	$656 \cdot 10^{-6}$	$183 \cdot 10^{-6}$
	QEL	$129 \cdot 10^{-6}$	$126 \cdot 10^{-6}$	$469 \cdot 10^{-6}$	$130 \cdot 10^{-6}$
NC		-	-	-	-

Chapter 4

Artificial neural networks

The algorithm for the event classification in the OPERA detector presented in this thesis is based on a neural network. The goal is to design a neural network which approximates the probability for a given event belonging to a particular class of events. The general working principles of such classification-type neural network are described in this chapter. The particular features of the used network for the event classification in the OPERA detector are described in Chapter 5. Regression-type neural networks, which are not used in this thesis, are described in Appendix A for completeness.

4.1 Introduction

An artificial neural network (*ANN* or *NN*) is a function $\mathbb{R}^N \longmapsto \mathbb{R}^M$. The structure of this function is modelled on the networks of biological neurons found in the nervous system of human or animal life forms. Biological neural networks are capable of recognizing complex patterns while learning from experience. The fundamental building blocks of those networks – the neurons and their connections, the synapses – are relatively simple. The complexity is achieved by combining a large number of neurons and synapses.

This structure was first transferred to a mathematical model by Warren McCulloch and Walter Pitts in 1943. Ever since, many attempts have been made to mimic neural networks artificially. At first, conventional computers were incapable of simulating neural networks. Only at the beginning of the 1980s the idea was revived, as classical computers became fast enough. Many different structures and learning algorithms have been developed since then. In this chapter only one class

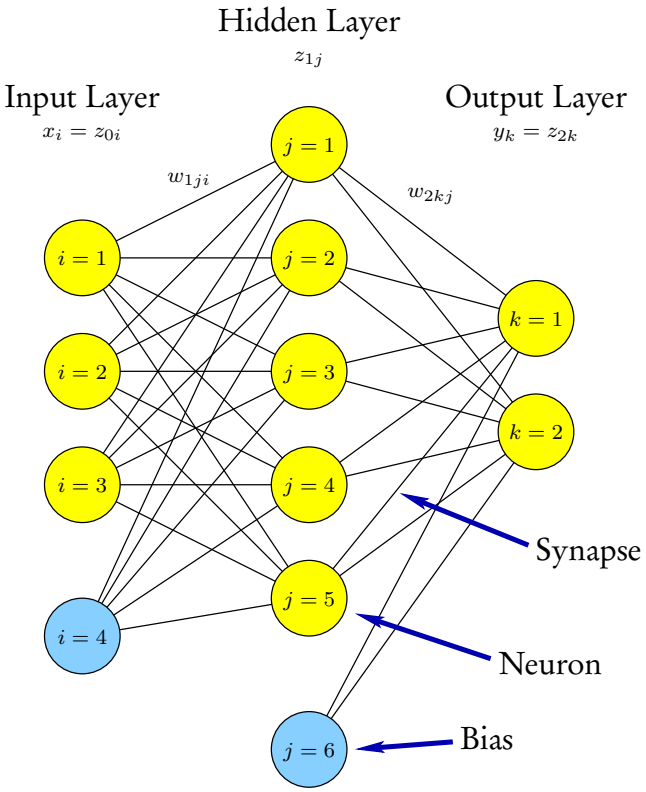

Figure 4.1: Sketch of a Multi-Layer Perceptron. In this example the network consists of three layers: the input, one hidden and the output layer. As in common terminology the input layer is not counted, this will usually be called a two layer network. Neurons are depicted as circles and synapses as lines. The indices on the neurons match the nomenclature used throughout this chapter. The first index always denominates the layer.

of artificial neural networks will be discussed: the Multi-Layer Perceptron. The description in this chapter follows [32].

4.2 The Multi-Layer Perceptron

The Multi-Layer Perceptron (*MLP*) is the network structure typically chosen for pattern recognition and classification problems. In an MLP the neurons are arranged in layers. Each neuron is connected with the neurons of the neighbouring layers. The classical MLP described in this chapter is a feed-forward network, which means that the network does not contain any loops. Hence the output of a neuron is always

connected to the next layer only, and never e.g. to the previous layer. In Figure 4.1, a typical structure of a small MLP is shown. In the following, $\mathbf{x}$ denominates the vector of input variables, $\mathbf{y}$ the vector of output variables and z_{lj} the output of the neuron j of the layer l. By this definition, $x_i = z_{0i}$ and $y_k = z_{Lk}$, with L the number of non-input layers (and thus $\mathbf{z}_L$ denominating the output vector).

Synapses connect two neurons by passing the output value of one neuron to the input of another while multiplying the value with an adjustable weight. This weight will be adjusted in the training process.

Neurons basically consist of two parts: an aggregation function, which combines the input values to a single value, and an *activation function* g, which determines the output value. The aggregation function is simply the sum of all input values. Therefore, the neuron j of the layer l with its connected input synapses can be described by:

$$z_{lj} = g\left(\sum_i w_{lji} z_{(l-1)i}\right). \tag{4.1}$$

One special neuron that always returns a constant, non-zero value (e.g. 1) is included in each but the output layer. This neuron gives a bias to each neuron of the successive layer. During training this bias is adjusted by changing the weight of the connecting synapse.[1]

In this thesis, classification-type networks are used. Such networks return an approximation of the a-posteriori probability of class membership $p(\mathcal{C}_1|\mathbf{x})$, which means the probability for an event belonging to an event class $\mathcal{C}_1$ under the condition that the input vector belonging to the event is $\mathbf{x}$. Two event classes are defined: the *signal* class $\mathcal{C}_1$ and the *background* class $\mathcal{C}_0$. The network has one output neuron, which shall approximate the a-posteriori probability. It is trained with a target value $t = 1$ for events of the class $\mathcal{C}_1$ and with $t = 0$ for events of the class $\mathcal{C}_0$. Extensions that allow more then two classes and an equal number of output values are not covered in this thesis.

For having the desired properties, the network needs to have the structure described in the following. A single hidden layer is sufficient [54]. The correct choice for the activation functions and the error function are described in the following sections.

[1] In some models this bias is instead introduced directly inside the neuron as part of the aggregation function, so each neuron gets an adjustable parameter – the bias. This will complicate the mathematical description of the MLP while representing a completely equivalent notation.

4.3 Activation functions

The input layer is not a true individual layer and not counted as such. Its only task is to distribute the values of the input variables to the synapses of the first hidden layer. Therefore, the input neurons have a linear activation function $g(x) = x$, so they simply pass the input value unchanged to the connected synapses.

The a-posteriori probability – which the classification-type neural network shall approximate – can be written as a logistic function $f(x) = \frac{1}{1+e^{-x}}$ using Bayes' theorem:

$$p(\mathcal{C}_1|\mathbf{x}) = \frac{p(\mathbf{x}|\mathcal{C}_1)p(\mathcal{C}_1)}{p(\mathbf{x}|\mathcal{C}_1)p(\mathcal{C}_1) + p(\mathbf{x}|\mathcal{C}_0)p(\mathcal{C}_0)} \tag{4.2}$$

$$= \frac{1}{1+e^{-a}}, \tag{4.3}$$

with

$$a = \ln \frac{p(\mathbf{x}|\mathcal{C}_1)p(\mathcal{C}_1)}{p(\mathbf{x}|\mathcal{C}_0)p(\mathcal{C}_0)} \tag{4.4}$$

and $p(\mathcal{C}_t)$ being the a-priori probability for the class $\mathcal{C}_t$. Therefore, the output neuron should have a logistic activation function. As the neurons in the hidden layer will represent a-posteriori probabilities of abstract sub-problems, the same argument is valid for their activation functions. Thus, the logistic function is used as the activation function of all neurons in the hidden and output layers.

4.4 Error function

In order to determine the optimal set of weights $\mathbf{w}$, an error function needs to be defined which will be minimised during the training process. If the training set is provided in form of the input vectors $\mathbf{x}^{(n)}$ and output (*target*) values $t^{(n)}$, where n is the event number (running from 1 to the number of events), a likelihood function can be defined under the assumption of errors being independent of $\mathbf{x}^{(n)}$:

$$\mathcal{L} = \prod_n p(\mathcal{C}_t^{(n)}|\mathbf{x}^{(n)}). \tag{4.5}$$

The classification-type MLP shall approximate the a-posteriori probability $p(\mathcal{C}_1|\mathbf{x})$ of the input vector $\mathbf{x}$ belonging to the class $\mathcal{C}_1$:

$$y \stackrel{!}{=} p(\mathcal{C}_1|\mathbf{x}) = 1 - p(\mathcal{C}_0|\mathbf{x}) \tag{4.6}$$

This leads to:

$$p(\mathcal{C}_t|\mathbf{x}) = \begin{cases} y & \text{for } t = 1 \\ 1 - y & \text{for } t = 0 \end{cases}, \tag{4.7}$$

which can be written in a compact form (as t can only have the values 0 or 1):

$$p(\mathcal{C}_t|\mathbf{x}) = y^t(1-y)^{1-t}. \tag{4.8}$$

This can be used with (4.5) to derive the likelihood function:

$$\mathcal{L} = \prod_n \left(y^{(n)}\right)^{t^{(n)}} \left(1 - y^{(n)}\right)^{1-t^{(n)}}. \tag{4.9}$$

It is more convenient to minimize the negative log-likelihood, which leads to the *cross-entropy error function*:

$$E = -\ln \mathcal{L} = -\sum_n \left[t^{(n)} \ln y^{(n)} + (1 - t^{(n)}) \ln(1 - y^{(n)})\right]. \tag{4.10}$$

This error function is depending on the weights $\mathbf{w}$ as it depends on the output value y of the network. E is minimised during the training process (see next section) by changing $\mathbf{w}$. Once the minimum has been found, the optimal $\mathbf{w}$ is known and the neural network can be used for the analysis.

4.5 Training the MLP

The purpose of the training process is to find the optimal $\mathbf{w}$ at which the error function is minimal. This involves solving a minimisation problem of a high dimension. The standard approach for this problem in this context is an iterative process and consists of an algorithm called back-propagation of errors, which is described in the next section, and an algorithm to change each weight (see Section 4.5.2).

4.5.1 Back-propagation of errors

The minimisation of the error function involves the calculation of its partial derivative for each weight $\frac{\partial E}{\partial w}$. For convenience, the error of a single event, $e^{(n)} := e$, with $E = \sum_n e^{(n)}$, is used first. Once $\frac{\partial e}{\partial w}$ is known, $\frac{\partial E}{\partial w}$ can be calculated as the sum over all events.

In the following, the two event classes are examined separately. If $t = 0$, e can be calculated from (4.10):

$$e = \ln(1 - y). \tag{4.11}$$

The output neuron with its logistic activation function, including all connected synapses, can be described as:

$$y = \frac{1}{1 - \exp\left(-\sum_j w_{Lj} z_{(L-1)j}\right)}. \tag{4.12}$$

This leads to:

$$\frac{\partial e}{\partial w_{Lj}} = z_{(L-1)j} \frac{1}{1 - e^{-\sum_j w_{Lj} z_{(L-1)j}}} = z_{(L-1)j}\, y. \tag{4.13}$$

In case of $t = 1$, one obtains from (4.10):

$$e = -\ln y, \tag{4.14}$$

which leads to:

$$\frac{\partial e}{\partial w_{Lj}} = z_{(L-1)j} \left(1 - \frac{1}{1 - \exp\left(-\sum_j w_{Lj} z_{(L-1)j}\right)}\right) = z_{(L-1)j}(1 - y). \tag{4.15}$$

Combining (4.13) and (4.15), and using $0 \leq y \leq 1$, gives:

$$\frac{\partial e}{\partial w_{Lkj}} = z_{(L-1)j} |y_k - t_k|. \tag{4.16}$$

For all other synapses the chain rule is used:

$$\frac{\partial e}{\partial w_{lji}} = \frac{\partial e}{\partial z_{lj}} \frac{\partial z_{lj}}{\partial w_{lji}}. \tag{4.17}$$

The second term can be directly calculated from (4.1):

$$\frac{\partial z_{lj}}{\partial w_{lji}} = z_{(l-1)i} \left.\frac{\partial g}{\partial z_{lj}}\right|_{z_{lj}}. \tag{4.18}$$

As only the product of w and z, but not w or z alone, appears in (4.1), the first term can be obtained based on the already known derivatives for the subsequent layer:

$$\frac{\partial e}{\partial z_{lj}} = \sum_k \frac{w_{(l+1)kj}}{z_{lj}} \frac{\partial e}{\partial w_{(l+1)kj}}. \tag{4.19}$$

Combining (4.17) through (4.19), the error of each synapse can be calculated from the derivatives of the error for the subsequent layer:

$$\frac{\partial e}{\partial w_{lji}} = z_{(l-1)i} \left.\frac{\partial g}{\partial z_{lj}}\right|_{z_{lj}} \sum_k \frac{w_{(l+1)kj}}{z_{lj}} \frac{\partial e}{\partial w_{(l+1)kj}}. \tag{4.20}$$

For this reason, the algorithm is called back-propagation of errors.

4.5.2 Updating the weights

To update the weights from the calculated derivatives of the error function, different algorithms can be applied.

The most basic one uses the per-event error $e^{(n)}$. The weights are updated for each $t^{(n)}$ according to the following update formula:

$$w \to w - \eta \frac{\partial e^{(n)}}{\partial w}, \tag{4.21}$$

with η being the *learning parameter*. η determines speed and precision of the learning process. Low η will lead to a higher precision with the trade-off of a slower convergence. The whole process is then repeated until the required precision is achieved. One iteration through all $t^{(n)}$ is called an *epoch*.

This basic update formula has several flaws with respect to speed and robustness against local minima. Therefore, a number of different update formulae exists which try to improve its performance. The first step towards more advanced algorithms is to make use of the overall error, so the update formula becomes:

$$w \to w - \eta \frac{\partial E}{\partial w}, \tag{4.22}$$

and each weight is only updated once per epoch. While being much faster than the original algorithm, it is very susceptible to local minima and stagnation due to flat regions in parameter space.

Further improvement can be achieved by picking a direction $\mathbf{d}$ in weight space and searching for a minimum along this direction using a one-dimensional parabolic fit. Several ways exist to choose the direction. This can be simply the negative gradient $\mathbf{d} = -\nabla E$, thus following the steepest decent. Also conjugate gradient algorithms can be used, as well as quasi-Newton methods that approximate the Hesse matrix. Usually quasi-Newton methods have the best performance for networks with a maximum of a couple of hundreds of weights, therefore such method is chosen for the training of the networks described in Chapter 5.

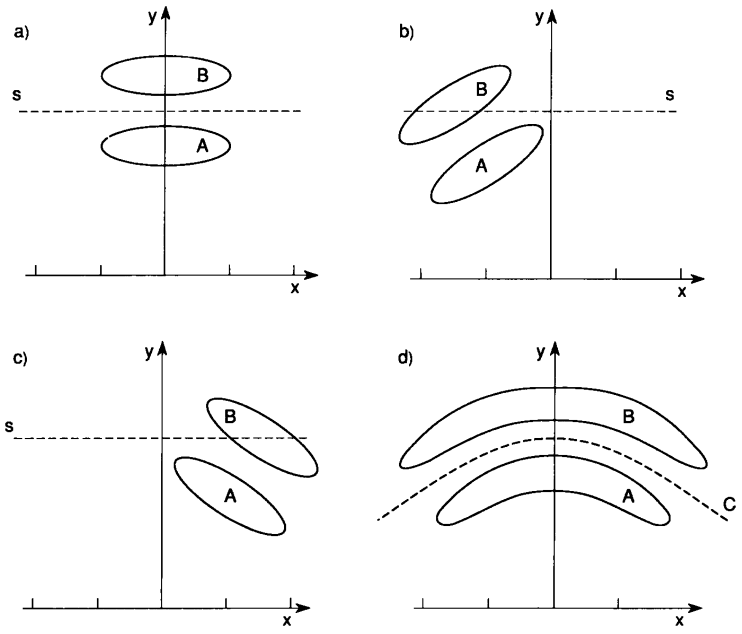

Figure 4.2: Illustration of the treatment of systematic errors in classification problems using the example of a simple two-class scenario. In a) the distribution of the two input variables x and y for the classes A and B is shown. Trained with this data, the neural network will implement a simple cut s to separate the two classes. If now a systematic error leads to a rotation in the x-y-plane, the distributions are shifted as shown in b) and c). Using the network previously trained with the data from a) the cut s now leads to a significant misidentification error for class B. If the network is instead trained with data smeared by the systematic error – as shown in d) – the network will implement a different cut c, which then provides an optimal separation of the classes even considering the systematic error. [69]

4.5.3 Treatment of systematic errors in classification problems

Propagating systematic errors through an MLP is not different from the propagation through any other function, though – due to the complexity of the MLP function – usually numerical methods are chosen. This can simply be done by modifying the input values by the amount of their errors. If the error of the input values is assumed to be uncorrelated, the variance of the output values can be obtained individually for each input value and is summed quadratically to obtain the total error of the output.

Especially in classification problems those errors can get quite big, if the network is trained only to work with the best estimate of the input values. The effect of the systematic error can hence be reduced by smearing the input values of the training data set with their systematic errors [69]. In Figure 4.2, the principle of this improvement is illustrated.

4.6 Implementation in ROOT

The data-analysis framework ROOT [33] includes an implementation of an MLP. Like the rest of the framework the MLP is implemented completely in an object-oriented manner. The main interface class *TMultiLayerPerceptron* is used to define and control the MLP. The structure of the network is specified as a string containing a list of input and output fields and the sizes of the hidden layers. Internally the network is being built from objects of the classes *TNeuron* and *TSynapse*.

For the user, this implementation has the advantage of being easily integrated into existing analysis software already using ROOT. On the other hand, this implementation is easy to maintain and to modify. Both advantages are traded off against a time penalty in the order of a factor of 2 compared to other, highly optimized implementations.

As the OPERA analysis framework is based on ROOT and as the implementation is relatively easy to understand, the *TMultiLayerPerceptron* implementation of ROOT is chosen for the analysis described in Chapter 5.

Chapter 5

Analysis of neutrino interactions at OPERA

The full analysis of an event in the OPERA detector involves the development and the scanning of the emulsion sheets of one or more bricks (see Section 2.3) and therefore is very time consuming. Using only information from the electronic detectors allows a true on-line analysis. In this chapter, a method is presented to speed up the process of analysing the emulsion sheets by sorting the events by a likelihood for being a signal event. Additionally, the possibility of observing $\nu_\mu \rightarrow \nu_\tau$ oscillations in disappearance mode using only information from the electronic detectors is being investigated.

5.1 Set-up of a general-purpose MLP

The Multi-Layer Perceptron was mathematically described in Chapter 4. In this section, the set-up of such a neural network for the purpose of this analysis is characterised.

5.1.1 Structure of the MLP

To classify different neutrino interaction types, the kinematic topologies of the events need to be determined. The kinematic variables x and y cannot directly be measured. They can be derived from E_ν, ν and the scattering angle θ, but at least for neutral current interactions E_ν is unknown. The target area of the OPERA detector is not well suited for calorimetry of showers. Therefore, a direct analysis

of the event kinematics is not possible. The available information is instead used as input variables for the MLP.

For the separation of NC and CC events, it is necessary to discriminate $\mu^{\pm}$ from $\pi^{\pm}$. The OPERA reconstruction software already has an algorithm implemented for this purpose which is based on a cut on the track length multiplied by the mean density of the material crossed by the track. In this analysis, this algorithm is not directly applied, but the information used by the algorithm is used as additional input parameters of the MLP.

To further improve the MLP, additional information about the shower development in the target area is used. All input variables used for the MLP are:

- the momenta of the first μ^- and the first μ^+,
- the angle of first μ^- to the beam axis (which basically represents the scattering angle),
- the reconstructed hadronic energy and its uncertainty,
- the length of the longest track multiplied by the mean material density along its path,
- the length of the longest track in radiation lengths,
- the number of target tracker walls hit by the hadronic shower,
- the root mean square (RMS) of the hit distribution in the first, middle and last target tracker walls hit by the shower,
- the longitudinal and the transversal RMS of the distributions of all target tracker hits, both unweighted and weighted with the number of photoelectrons and
- the number of photoelectrons in the wall with the reconstructed interaction point, in all walls upstream and in the next wall downstream of the interaction point.

With its 18 input parameters, this MLP is close to the limit of what the implementation in the ROOT framework used in this analysis can handle on currently available computers. The MLP has one hidden layer, which needs to be significantly larger than the number of input parameters for processing all features of the data. For this analysis, 36 neurons in the hidden layer are used.

This general structure of the MLP is used for multiple purposes. It has one single output variable, which approximates the a-posteriori probability of membership of the signal class the network is trained to detect. The network has been trained in four different versions to distinguish each of the following event classes from all other events:

- all kinds of ν_τ interactions,
- all kinds of ν_e interactions,
- all kinds of both ν_τ and ν_e interactions and
- all kinds of neutral current interactions.

In the following, the each one of the above-mentioned event classes is referred to as the signal class $\mathcal{C}_1$, while the class of all other events is called the background class $\mathcal{C}_0$. The MLP is set up according to the rules for classification problems as stated in Chapter 4, i.e. all neurons in hidden and output layers have a logistic activation function and the cross-entropy error function is used as the error function during training.

To all data – both Monte Carlo (MC) and real data – one single cut is applied: Only events that are marked as *contained* inside the target area (i.e. the interaction point lies within the target area) by the package *OpCarac* [31] of the OPERA software are retained. As only very few MC events with the interaction point lying outside the target area (e.g. in the surrounding rock) are passing this cut, only events generated with a primary vertex inside the target area are used for training and testing. As the OPERA MC software uses Geant 3[1] for particle transportation, contributions from events with the neutrino vertex outside the detector and involving a neutral particle (e.g. a neutron or γ) entering the detector are generally underestimated. These events can mimic an NC-like neutrino interaction in the target area close to its border. These kinds of events are marked as *bordersoftnc* by *OpCarac*. While the sample of events marked as *bordersoftnc* contains also a fairly large fraction of neutrino interactions at the border of, but inside, the target – and therefore cannot be discarded in the neutrino analysis – it contains almost all of the described external NC-like events. Thus, for the training, events marked as *bordersoftnc* are included. On the other hand, for comparison with real data,

[1] Geant is a "detector description and simulation" tool developed at CERN. It simulates the transportation of particles through the passive and active material of an experimental set-up. Version 3 is no longer maintained but still in use by many experiments.

only events marked as *contained* as used, to eliminate any differences caused by the described effect.

In some cases not all information used as input variables is available, e.g. there will be no momentum information if no muon has been identified. In those cases the input variables are filled with a replacement value which does not occur normally, such as "0" for the momentum or "-1" for the angle to the beam axis.

For training and analysis, the data is first converted into a standard ROOT tree containing all the necessary information. This data format can then directly be processed by the MLP implementation of ROOT.

5.1.2 Treatment of systematical uncertainties

As described in Section 4.5.3, the effect of systematic uncertainties of the MC can be reduced by smearing the input variables of the neural network during the training. For this purpose, an estimation of the systematic uncertainty of each input variable is required.

The OPERA MC software does not provide any means for controlling systematic effects. This could e.g. be done by assigning each event a weight depending on the value of a certain model parameter (like the mass of the charm quark). If these weights are correctly processed by the simulation software, a tuning of these model parameters will be possible without having to re-run the full simulation chain. However, the implementation of such a mechanism in the OPERA software would go beyond the scope of this thesis.

Therefore, the systematic uncertainties are estimated by comparing the spectra of each input variable for MC and real data. The systematic uncertainty σ_n of the n-th input variable is obtained from the mean values ($\mu_n^{(\mathrm{MC})}$, $\mu_n^{(\mathrm{real})}$) and the RMS values ($\rho_n^{(\mathrm{MC})}$, $\rho_n^{(\mathrm{real})}$) of the variable:

$$\sigma_n = \frac{1}{2}\left[\left|\mu_n^{(\mathrm{MC})} - \mu_n^{(\mathrm{real})}\right| + \left|\rho_n^{(\mathrm{MC})} - \rho_n^{(\mathrm{real})}\right|\right] \tag{5.1}$$

Input variables that have only a physical meaning for positive numbers are smeared relatively, i.e. the obtained uncertainty is scaled by:

$$r = \frac{i_n}{\frac{1}{2}\left|\mu_n^{(\mathrm{MC})} - \mu_n^{(\mathrm{real})}\right|}, \tag{5.2}$$

where i_n is the value of the input variable. Variables which have a physical meaning for both positive and negative values, like angles, are smeared by the absolute

uncertainty. The actual smearing is performed by adding a random number with a Gaussian distribution with a width of $8\sigma_n$. The used spectra are shown in Appendix B.

5.1.3 Training with MC data

To train the MLP, a comprehensive set of Monte Carlo data with reasonably large statistics is required. The NEGN event generator generates events of neutrino interactions separately for different event classes. It does not provide any information about the relative abundances of these different classes. These relative abundances have been calculated in Chapter 3 and the result is given in Table 3.5. To avoid additional quantisation effects, each event is assigned a weight w, according to the calculated abundance a of its NEGN class and the number n of simulated events of that class: $w = \frac{a}{n}$.

In most cases throughout this analysis, the signal events have a much lower abundance than the background events. For an optimal training of the network and to make the network independent of the parameters to be observed in the analysis, the total abundances for the signal and the background classes, $a_{\mathcal{C}_0}$ and $a_{\mathcal{C}_1}$, respectively, are both normalized to 1, i.e. $a_{\mathcal{C}_0} = a_{\mathcal{C}_1} = 1$. Within the signal and the background classes, the proportions of the individual NEGN classes as given in Table 3.5 are used.

Per each of the 28 event classes, $2 \cdot 10^5$ events have been generated, resulting in a total statistics of $5.6 \cdot 10^6$ events. The training of each version of the network is done for several 1000 epochs, until no significant improvement is observed. Training takes a significant amount of CPU time, around 30 min per epoch on a 3.6 GHz AMD X6 CPU core.

The training process was monitored to detect over-training and other unwanted effects. Over-training will occur if the statistics of the training sample is too small. The network then learns the individual patterns of each event, instead of general features of the event classes. This can be observed by dividing the available MC data into two equal sets. One set is used for the training itself and therefore is called the *training sample*. The other set is used to test the neural network and therefore is called the *test sample*. After each epoch, the error (cf. Section 4.4) is calculated for both samples. An example for the errors in dependence of the epoch, with a visible, small over-training effect, can be seen in Figure 5.9. Additionally, thanks to a small modification of the ROOT software, the weights of all synapses are written

to a file after each epoch. This allows testing and using the network already during the training process. The training process is considered finished when no significant change of the error is observed and the result of the analysis is stable for several epochs.

In Appendix C results proving the integrity of the training process are shown. The learning curves exhibit no significant deviation between training and test sample – apart from the exceptions named in this chapter – which indicates the absence of over-training effects. It is also shown that the neural network output reflects the true a-posteriori probability – or at least a monotonic function of the probability as in the one exception also shown in this chapter.

Problems like locking to local minima of the error function or stagnation in flat regions were excluded by training the neural networks multiple times. Before each training procedure, all synapse weights were randomised, therefore the starting conditions were different each time. Comparable results have been achieved every time.

5.2 Obtaining the efficiencies for ν_τ and ν_e detection

The MC data used for training and testing of the neural networks does not contain any information about particle hits in the ECC. A different MC data set, which includes ECC information, was used to investigate the effect of the efficiencies from event localisation on the quality of the event ranking. Producing and storing this data takes considerably more computing resources and time, which are not easily available. Thus, this data set has a much lower statistics and therefore can only be used for an additional cross check.

For historical reasons, different software packages for analyses referring to the ECC and to the electronic detectors are in use by the OPERA collaboration. To allow a combined analysis of ECC and electronic detectors, a new software package called *OpEmuRec* is currently under development. This package does not yet allow a full decay search as described in Section 2.2.1, but studies regarding the location efficiency of an event can already be realised. The analysis steps described in Chapter 2 are reproduced until a presumed vertex has been found and the total scan volume has been defined. Thus, efficiencies from the actual volume scan, from scan forth and from kinematical cuts are not included. Additionally, the efficiency cuts cannot be very tight, as in the process of the analysis a recovery e.g. of a misidentified vertex

Table 5.1: Efficiencies for the detection of the different ν_e and ν_τ event classes with the ECC of the OPERA detector. The first two columns describe efficiencies of the electronic detector components: *OpCarac* gives the efficiency for the correct identification of the events as contained neutrino interactions, *BrickFinder* contains the efficiency to locate the correct brick. The last three columns contain the efficiencies of the ECC components of the detector: *CS* describes the efficiency of the changeable sheet, *SB* the efficiency of the scan back procedure and *TS* the geometric efficiency of the total scan volume being sufficiently large. Resonant scattering is not covered, as it is not included in the MC used for this analysis.

Event class	OpCarac	BrickFinder	CS	SB	TS
ν_e DIS	0.901	0.616	0.608	0.608	0.546
ν_e QEL	0.852	0.608	0.581	0.581	0.457
ν_τ μ DIS	0.959	0.680	0.450	0.450	0.417
ν_τ μ QEL	0.943	0.779	0.016	0.016	0.007
ν_τ e DIS	0.895	0.611	0.560	0.560	0.430
ν_τ e QEL	0.793	0.534	0.449	0.449	0.226
ν_τ 1h DIS	0.919	0.605	0.545	0.545	0.508
ν_τ 1h QEL	0.890	0.425	0.038	0.038	0.027
ν_τ 3h DIS	0.917	0.590	0.575	0.575	0.544
ν_τ 3h QEL	0.929	0.468	0.393	0.393	0.371

is possible in many cases. It is also important to note that the MC does currently not contain any background tracks.

The efficiency cuts applied are defined as following:

- the event must be tagged as *contained* or *bordersoftnc* (*OpCarac*),
- the brick with the highest probability must be the correct one (*BrickFinder*),
- the changeable sheet must contain at least one track (*CS*),
- in the scan back procedure, at least one track must have been found (*SB*) and
- the total scan volume must be sufficiently large to start a decay search, i.e. the vertex has to be at least three emulsion plates upstream of the border of the ECC brick (*TS*).

The last requirement (*TS*) is based on the true MC vertex to allow a later recovery in case the first attempt to locate the vertex failed.

The resulting efficiencies are listed for reference in Table 5.1. These numbers are not actually used for the further analysis. Instead, a sample of events surviving these efficiency cuts (thus being "efficient") is compiled and later on used to study the effect of the efficiencies on the output spectra of the neural networks.

5.3 Ranking of events

As described in Section 4.4, the output value y of the classification-type network will be an approximation of the a-posteriori probability $p(\mathcal{C}_1|\mathbf{x})$ for the event belonging to the signal class $\mathcal{C}_1$, i.e. $y \approx p(\mathcal{C}_1|\mathbf{x})$. As all events with the same probability $p(\mathcal{C}_1|\mathbf{x})$ result in the same output value y, even if the input vectors $\mathbf{x}$ are different, the following identity is true:

$$p(\mathcal{C}_1|y) = p(\mathcal{C}_1|\mathbf{x}). \tag{5.3}$$

Thus, the output value y is also an approximation of the a-posteriori probability $p(\mathcal{C}_1|y)$ for an event belonging to the class $\mathcal{C}_1$ under the condition that the event results in the output value y, which is the ideal choice for the purpose of an event ranking.

As the network has been trained with $a_{\mathcal{C}_0} = a_{\mathcal{C}_1}$, the network will not directly deliver the true probability. Instead, the true probability can be calculated from the output of the neural network and the ratio of the abundances $r = \frac{a_{\mathcal{C}_0}}{a_{\mathcal{C}_1}}$:

$$p_r(\mathcal{C}_1|y) = \frac{ry}{1 + ry - y}. \tag{5.4}$$

For a better display, this correction is not applied to the figures in this chapter. As Equation 5.4 is a monotonic function for $0 \leq y \leq 1$, this correction also does not affect the ranking order of events. Throughout this chapter, the uncorrected output of the neural network will therefore be used, which will be called the *likelihood*, despite the fact it is only an approximation of the true likelihood.

5.3.1 Ranking of ν_τ events

OPERA is primarily designed to detect ν_τ events. Ranking the events by the ν_τ likelihood can therefore speed up the main analysis of the OPERA experiment,

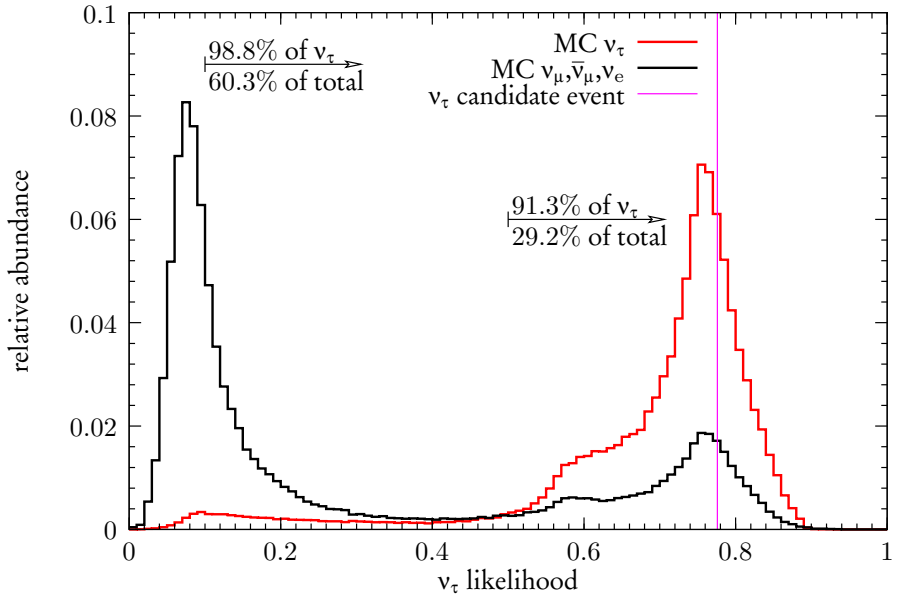

Figure 5.1: Output spectrum for the ν_τ network, shown separately for signal (ν_τ) and background ($\nu_\mu, \bar{\nu}_\mu, \nu_e$) classes.

especially since the analysis of candidate events takes considerably more time than of the majority of background events. Thus, this network has been trained with $\mathcal{C}_1$ containing all kind of ν_τ events and $\mathcal{C}_0$ containing ν_μ, $\bar{\nu}_\mu$ and ν_e events.

The output spectrum of the network is shown in Figure 5.1. The structure with two peaks indicates a good separation. The peak at low likelihoods has a very small contamination with ν_τ events. The peak at high likelihoods has a larger contamination with background events and therefore is shifted to slightly lower values around 0.75. About 30% of all events but more than 90% of the ν_τ events have a likelihood of larger than 0.5. The likelihood of the single ν_τ candidate discovered by OPERA at the time of writing is close to the maximum of the output spectrum for ν_τ events. As the fraction of ν_τ events in the real experiment is very small, the contamination with background events will be much larger than the actual number of signal events. Therefore, this network cannot be used to directly detect ν_τ events.

Figure 5.2 shows a comparison with real data of the run years 2008 to 2011. The showed MC data has been mixed realistically according to the ratios calculated in Chapter 3. There is no significant difference between MC and real data, which confirms the success of the treatment of systematic errors as described in Section 5.1.2.

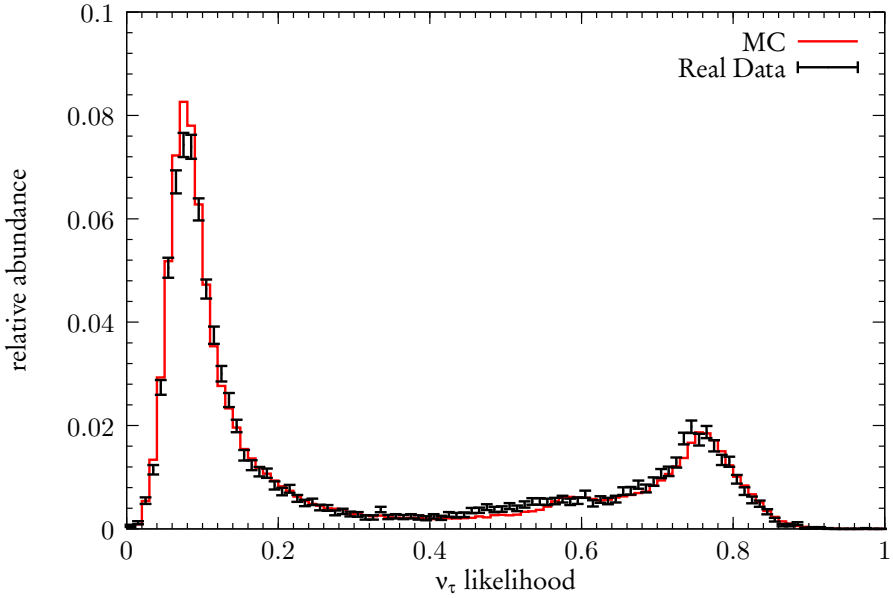

Figure 5.2: Comparison of the output of the ν_τ network for MC and real data.

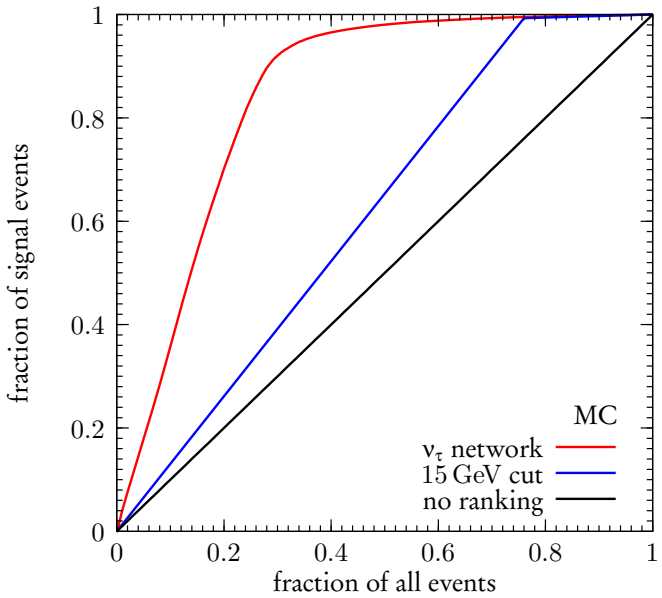

Figure 5.3: Speed increase for the ν_τ ranking. The black line ("no ranking") is for randomly distributed ν_τ events in the sample, the blue line ("15 GeV cut") displays the effect of the brick extraction strategy applied since July of 2011 to the data of the 2010 run. The red line ("ν_τ network") is obtained by sorting the events by the likelihood given by the neural network.

The fraction of signal events vs. the fraction of total events analysed selected by hypothetical cuts on the ν_τ likelihood – which will be called "speed increase" in future – is shown in Figure 5.3. For comparison, the brick extraction strategy used by the OPERA collaboration for the 2010 data as of July 2011 is also shown. This strategy involves prioritising all NC-like events and CC-like events with a reconstructed muon energy $E_\mu < 15\,\mathrm{GeV}$. A MC study performed as part of this analysis has shown that the prioritised data sample contains 76.3% of all events and 99.3% of the ν_τ events. A comparison with real data indicates a large systematic uncertainty on these numbers, as the prioritised sample of real data contains $(81.0 \pm 1.5\ (\mathrm{stat}))\%$ of all events. A selection of a prioritised data sample with the same efficiency of 99.3% based on the neural network shows an agreement between MC and real data within the statistical uncertainty (66.2% for MC and $(67.1 \pm 1.4\ (\mathrm{stat}))\%$ for real data).

5.3.1.1 Performance for the individual τ decay channels

In Figures 5.4 and 5.5 the output spectrum and the speed increase is shown separately for each τ decay channel. The hadronic and electronic decay channels allow the best separation and thus all events in these channels get a high likelihood. If ranking by the likelihood is applied, 99% of all events in these channels are located within the first $\sim 30\%$ of the analysed data. Due to the similarity of some events in the muonic decay channel with ν_μ CC events, these events get a low likelihood. Applying the event ranking by the likelihood will still improve the speed of the analysis of muonic ν_τ events, but not as much as for the other decay channels.

5.3.1.2 Effect of efficiencies

The comparison of "efficient" and all ν_τ events, according to the definitions given in Section 5.2, is shown in Figure 5.6. A small shift of the efficient events towards higher likelihoods, with respect to all events including inefficient events, is visible. A negative influence of the efficiency on the effectiveness of the event ranking can therefore be excluded.

5.3.1.3 Test of separate networks trained for NC- and CC-like events

To test whether a pre-selection of events by their NC likelihood can improve the separation, two individual networks have been trained to approximate the ν_τ

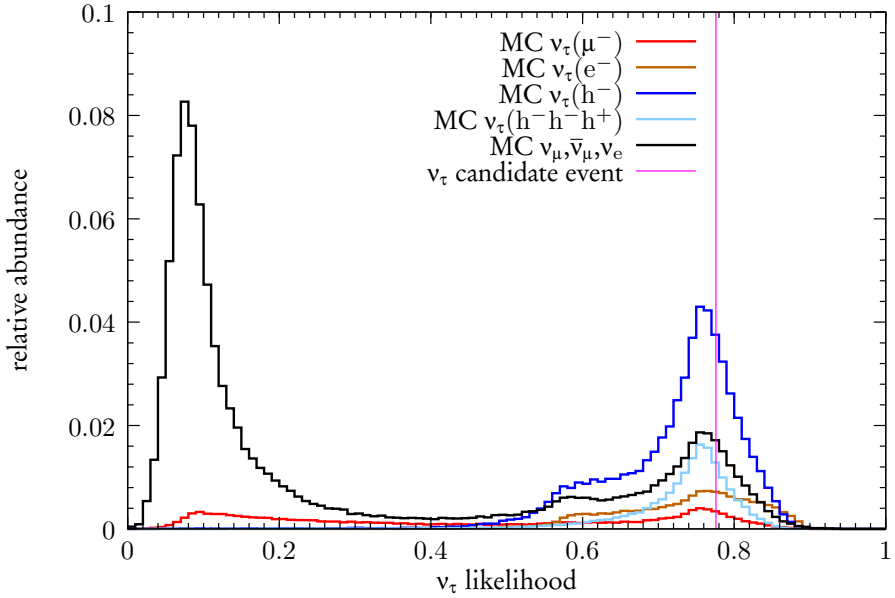

Figure 5.4: Output spectra of the ν_τ events for the different τ decay channels. The electronic, hadronic one-prong and three-prong channels only have significant contributions in the high-likelihood peak. Only the muonic decay channel has a contribution in the low-likelihood peak otherwise consisting of background events only.

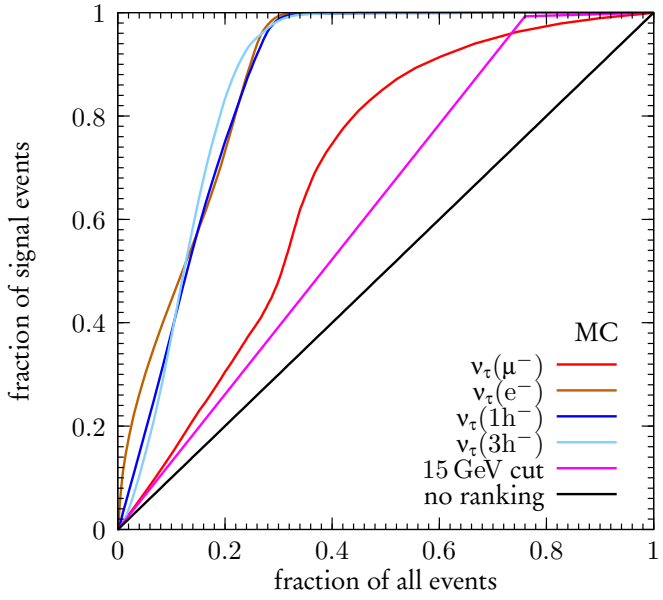

Figure 5.5: Speed increase for the ν_τ network, separated by the different τ decay channels.

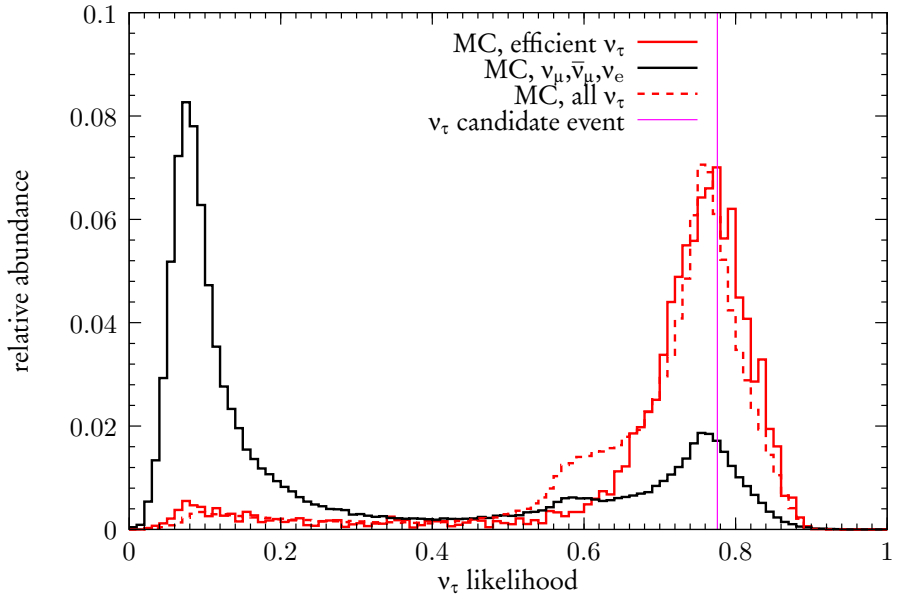

Figure 5.6: Comparison of the output spectra of the ν_τ network for efficient (solid red line) and all (dashed red line) ν_τ events.

likelihood for NC-like and for CC-like events. The neural network described later in Section 5.4.1 has been used to determine the NC likelihood l_{NC}. The event samples used for the training are produced with a slight overlap, to avoid a bad performance for events close to the cut: a cut on the NC-likelihood of $l_{\mathrm{NC}} > 0.2$ for the NC-like sample and a cut of $l_{\mathrm{NC}} < 0.8$ for the CC-like sample has been used.

The result is shown in Figure 5.7. For NC-like events, the network does not succeed in separating ν_τ from other types of events, regardless of a separate or common training. The separation for CC-like events is also not as good as the total separation shown in Figure 5.1. This allows the conclusion, that part of the decision is based on the NC-likelihood.

For a better comparison between the common and the separate training, the speed increases for both trainings are shown in Figure 5.8, which indicates a marginal better performance of the separate training. As shown in Figure 5.9, the separate training shows a slight over-training effect: The network is performing better for the training sample than for the test sample. This network is therefore not well suited for any real use. The over-training effect could be eliminated by using a substantially larger statistics during training. However, this is technically not feasible and there is no guarantee for an improvement of the performance.

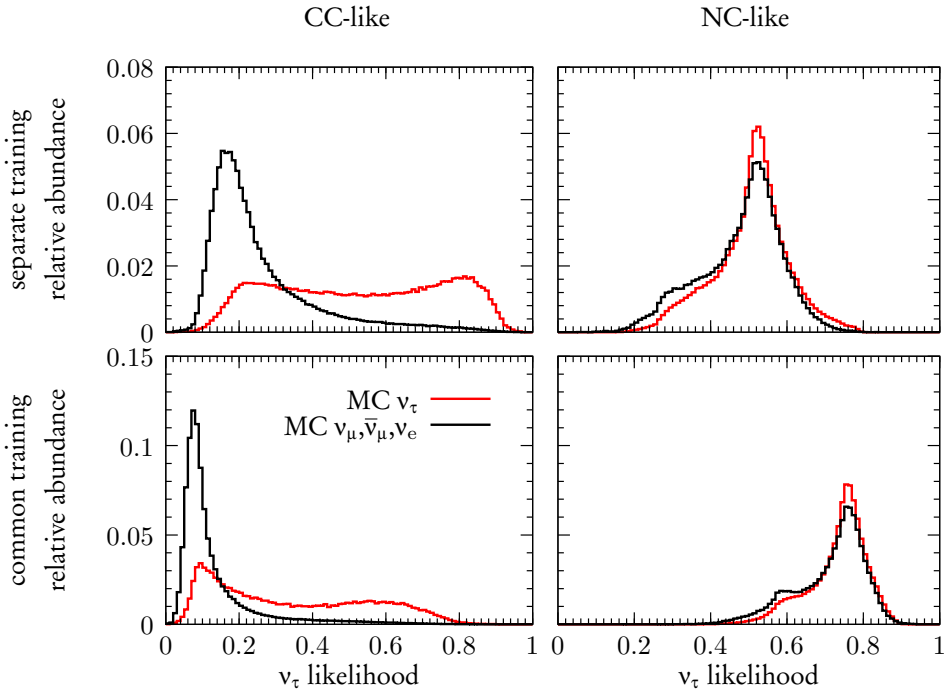

Figure 5.7: Comparison of separate training of NC- and CC-like events with the standard common training. The upper two plots show the separate and the lower two plots the common training. The left-hand plots are with CC-like and the right-hand plots with NC-like events. All histograms are individually normalised to 1.

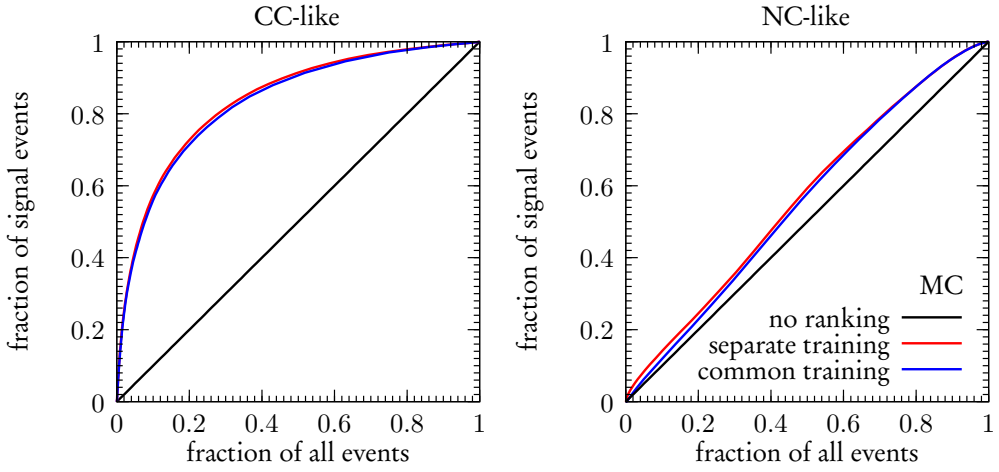

Figure 5.8: Speed increase for the ν_τ network, in comparison for the separate and common training. On the left, CC-like events, and on the right, NC-like events are shown.

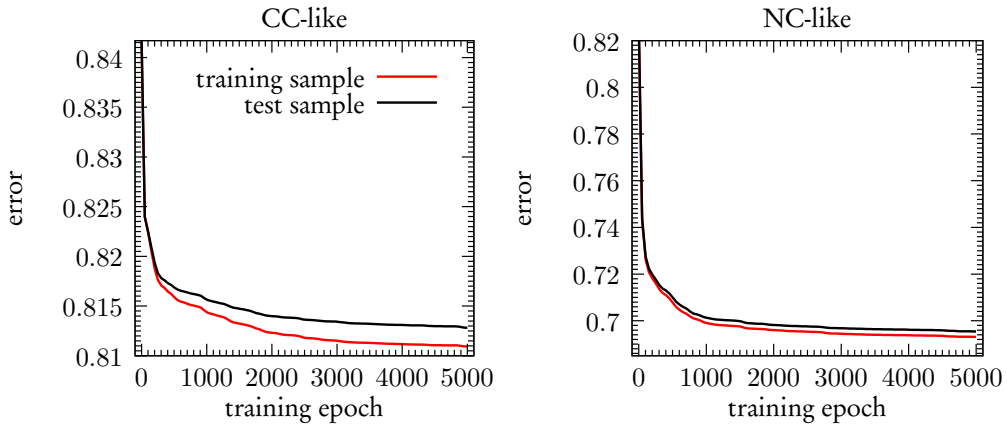

Figure 5.9: Cross-entropy error of the training and of the test sample over training epoch, for the separate training with CC-like (left) and NC-like (right) events. Ideally, the errors should be the equal for the training and for the test sample. The visible difference indicates a slight over-training effect.

5.3.2 Ranking of ν_e events

In a similar way as shown in Section 5.3.1 for ν_τ events, a ranking by the ν_e likelihood is possible. The neural network has been trained with all kinds of ν_e events as signal class $\mathcal{C}_1$ and all kinds of $\overset{(-)}{\nu}_\mu$,ν_τ events as background class $\mathcal{C}_0$.

The resulting output spectra for ν_e and for background events are shown in Figure 5.10. The ν_e spectrum is dominating the high-likelihood area and has no significant contributions below a likelihood of 0.35. The background spectrum shows a clear peak at a likelihood of 0, these events can therefore be excluded from a search for ν_e events. All candidate ν_e events that have been observed by OPERA at the time of writing have a likelihood of greater then 0.7 and their distribution follows the expected spectrum.

The comparison between MC and real data of the run years 2008 to 2011 shown in Figure 5.11 shows a good agreement. The MC data only includes ν_e from the beam contamination. An excess of ν_e from a $\nu_\mu \rightarrow \nu_e$ oscillation is not visible but can also not be excluded with the given statistics.

In Figure 5.12, the speed increase of a ν_e search is displayed. After analysing 38.2% of all events sorted by the ν_e likelihood, 99.9% of the ν_e events will be found. The comparison with "efficient" events in Figure 5.13 shows no impact of the location efficiencies on the ν_e ranking.

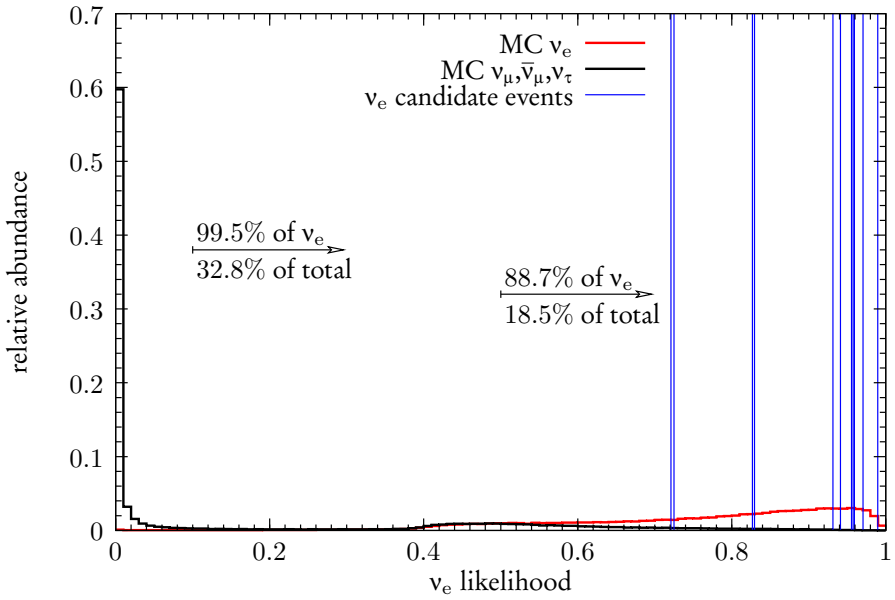

Figure 5.10: Output spectrum of the NN trained for ν_e ranking.

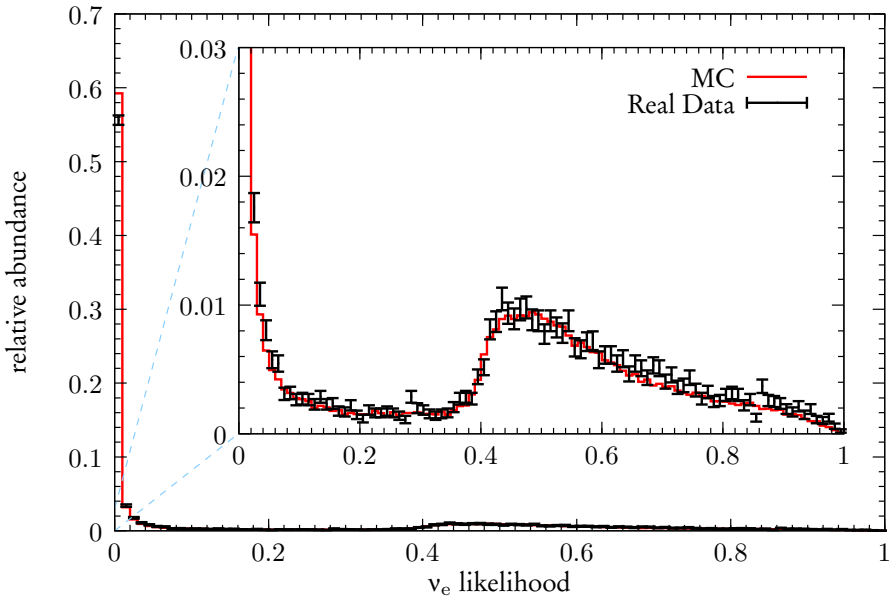

Figure 5.11: Comparison of the output spectra of the ν_e NN for MC and real data

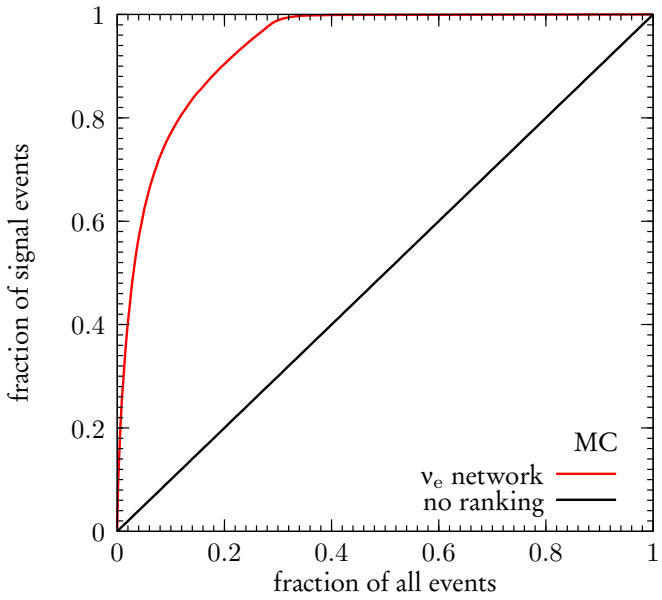

Figure 5.12: Speed increase for the ν_e ranking. The black line ("no ranking") is for randomly distributed ν_e events in the sample, the red line ("ν_τ network") is obtained by sorting the events by the likelihood given by the ν_e network.

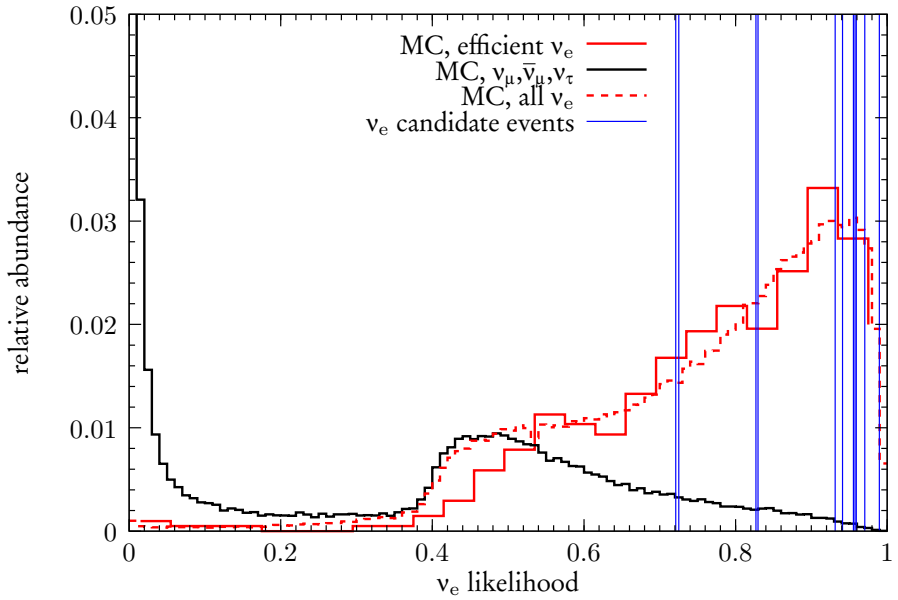

Figure 5.13: Comparison of the output spectra for the ν_e NN between all (dashed red line) and efficient ν_e (solid red line) events.

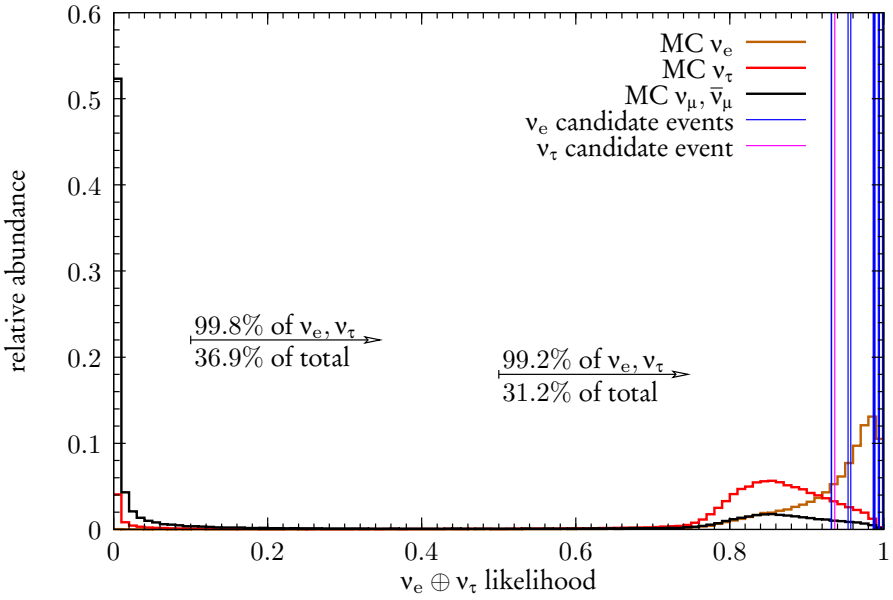

Figure 5.14: Output spectra of the NN trained for a combined ν_e and ν_τ ranking, shown separately for ν_e and ν_τ events. The spectra are slightly distorted as indicated by Figure 5.15, because the output of the neural network is not perfectly approximating the true likelihood. The ranking of events is not affected, as the mapping is monotonic.

5.3.3 Combined ranking of ν_τ and ν_e events

As OPERA is searching for ν_τ and ν_e events simultaneously, a combined ranking of both kinds of events seems reasonable. Thus, an additional network has been trained with both ν_τ and ν_e as signal class $\mathcal{C}_1$, realistically weighted according to Chapter 3. The background class $\mathcal{C}_0$ only consists of ν_μ and $\bar{\nu}_\mu$.

The resulting output spectra are shown in Figure 5.14. These spectra are slightly distorted, indicated by Figure 5.15. This distortion does not affect the event ranking, as the distortion function is monotonic. The probable cause for this distortion is the increased complexity of the signal class $\mathcal{C}_1$ with respect to the separate ν_τ and ν_e cases, while the size of the network remained unchanged. An increase of the size of the network, e.g. by adding additional hidden neurons, is not feasible, as the current size already is at the limits of the capabilities of the used neural network implementation. Furthermore, an increased size would require a larger statistics for the training sample, which is not easily available.

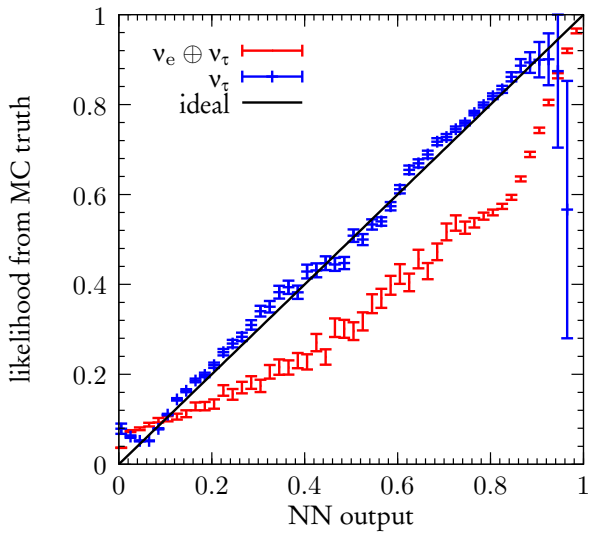

Figure 5.15: True likelihood over output of the neural network. Ideally, the two values should be equal. For the ν_τ network, no significant deviations from the ideal case are visible. The combined ν_τ and ν_e network shows a deviation, which distorts the spectra shown in Figure 5.14. As the mapping is still monotonic within the statistical uncertainties, ranking events by the NN output is not affected.

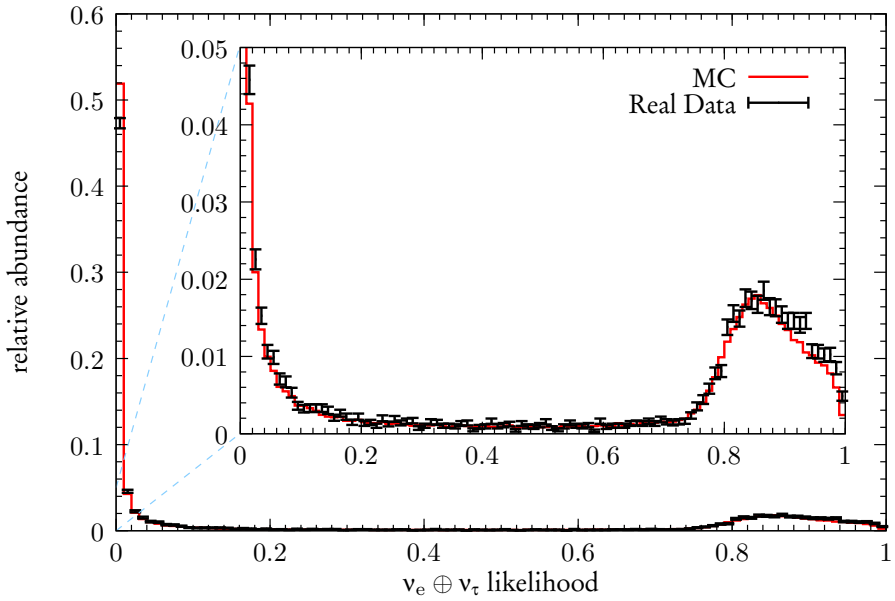

Figure 5.16: Comparison of the output spectra for MC and real data for the combined ν_τ and ν_e network.

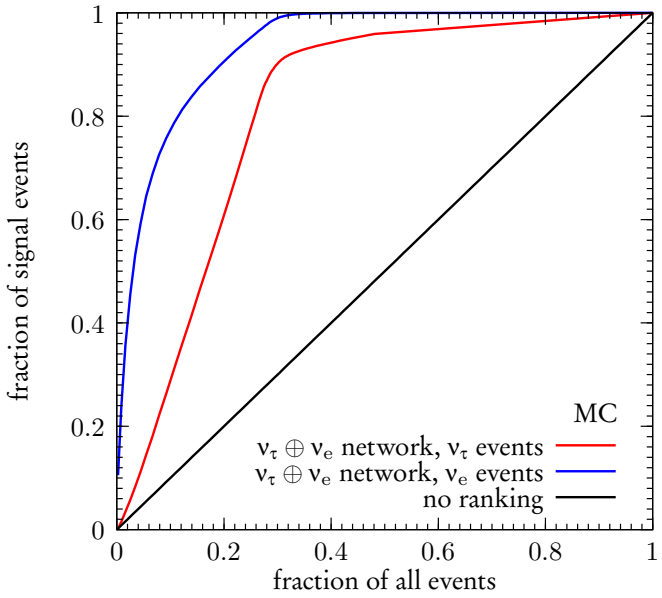

Figure 5.17: Speed increase for the combined ν_e and ν_τ ranking, shown separately for ν_e and ν_τ events. Both shapes match closely the development of the individual networks as shown in Figure 5.3 and Figure 5.12, respectively.

As shown in Figure 5.16, MC and real data of the run years 2008 to 2011 also match for the combined ν_τ and ν_e network. The speed increase of the search for ν_τ and ν_e events is shown in Figure 5.17. The combined ranking matches the performance of the individual rankings shown in the previous sections. The combined ranking can therefore be used for a simultaneous ν_τ and ν_e search.

5.4 Detection of neutrino oscillations using the NC fraction

$\nu_\mu \rightarrow \nu_\tau$ oscillations can in principle be observed at OPERA also in disappearance mode. As there is no CNGS neutrino near detector, the classical approach using the total neutrino flux is not possible due to the lack of a precise flux normalisation. Instead, the ratio of ν_μ DIS NC events per all ν_μ DIS events (*NC fraction*) can be measured in energy dependence. Below the τ production threshold, a ν_μ oscillated into a ν_τ can only interact via the NC. As the oscillation probability depends on the neutrino energy (see Section 1.2), variations of the NC fraction should be observable in case of neutrino oscillations. Under the assumption that the NC fraction is not energy-dependent if there are no neutrino oscillations, the oscillation parameters can

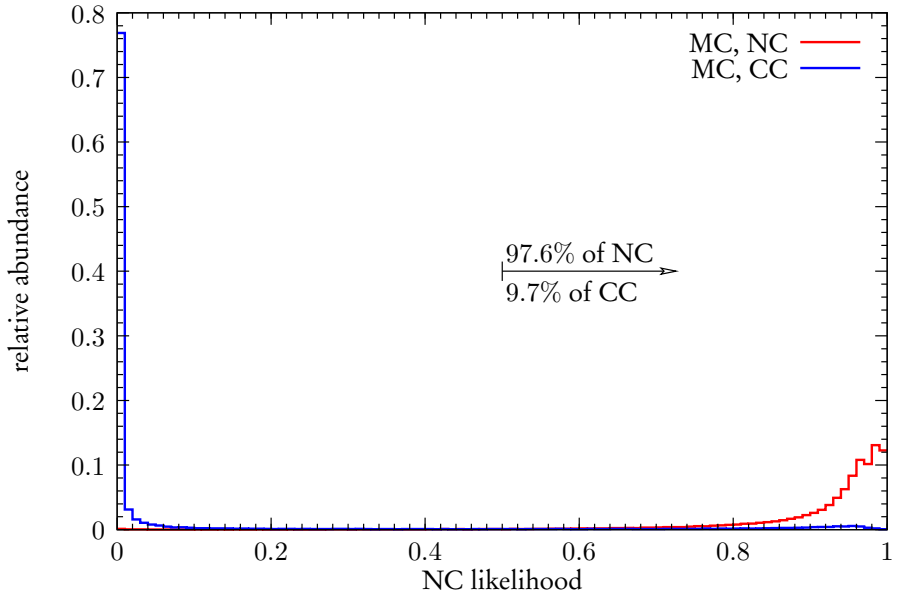

Figure 5.18: Output spectra of the NC network for true NC events and true CC events.

be determined using the energy dependence resulting from oscillations. It should be pointed out that this assumption has not yet been verified sufficiently. The corresponding measurement has to be conducted at a short baseline experiment and thus cannot be done as part of this thesis.

5.4.1 Separation of NC- and CC-like events

To separate NC and CC events, the neural network has also been trained with NC events as $\mathcal{C}_1$ and CC events as $\mathcal{C}_0$. The neural network output is then an approximation of the NC-likelihood which can be transformed into the probability for the event being a NC interaction with Equation 5.4 if the true NC/CC ratio is known. The output spectra of this network are shown in Figure 5.18 and the comparison with real data of the run years 2008 to 2011 is shown in Figure 5.19.

5.4.2 Determination of the NC fraction

The network described in the last section can be used to determine the NC fraction. To determine the NC fraction without any effects from neutrino oscillations, only events with a reconstructed visible energy of $E_{\text{vis}} > 15\,\text{GeV}$ are used. A MC study has shown that the influence of neutrino oscillations in this energy region

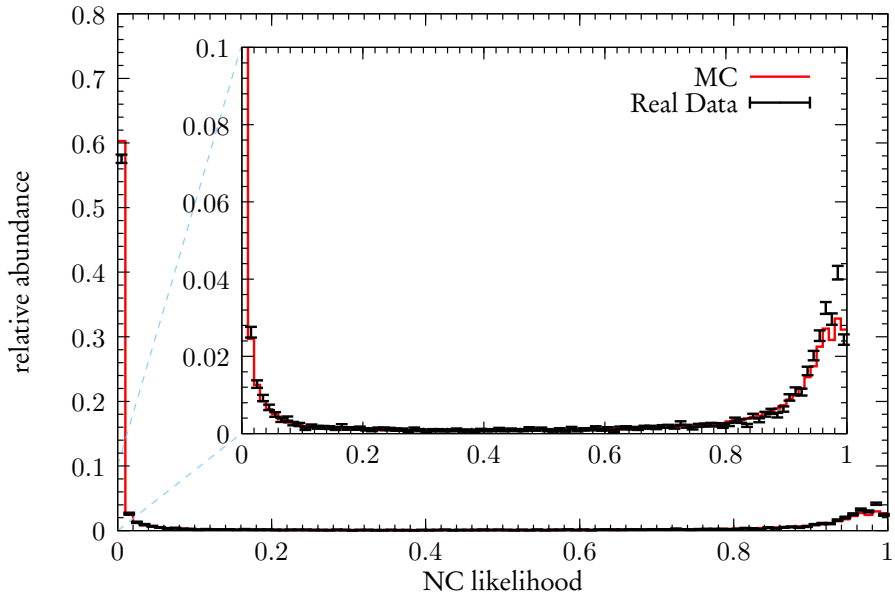

Figure 5.19: Comparison of the output spectra of the NC network for MC and real data.

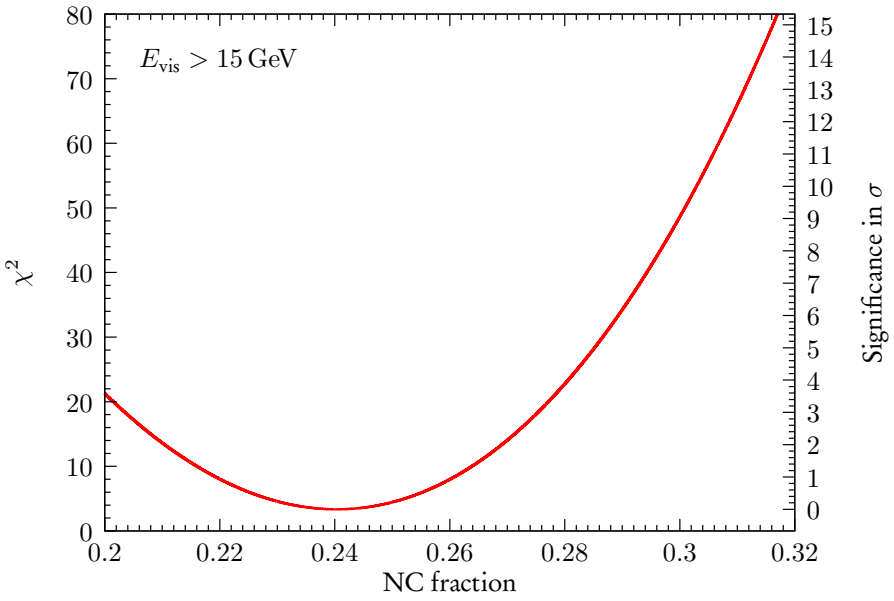

Figure 5.20: χ^2 over NC fraction for a visible energy of $E_{vis} > 15$ GeV. The number of degrees of freedom for this fit is 5.

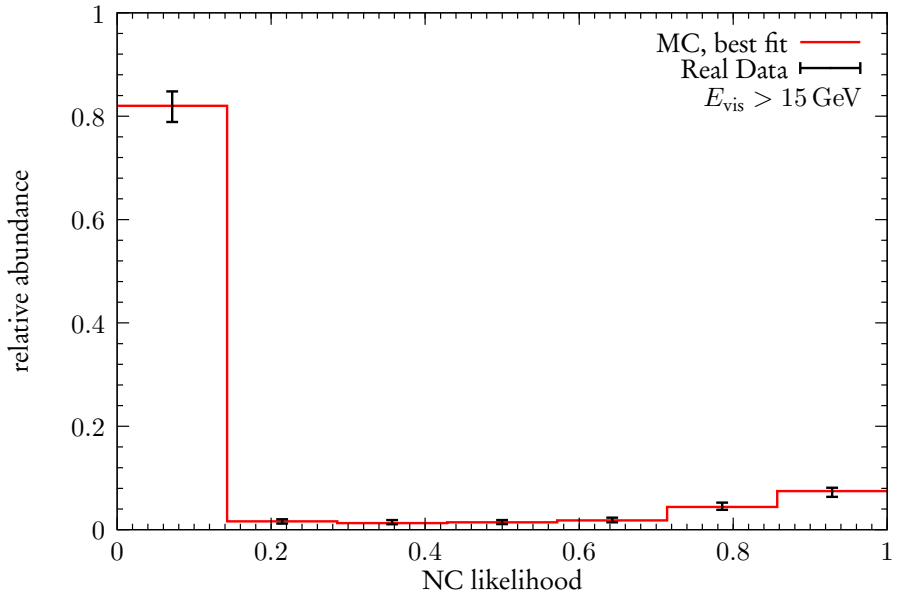

Figure 5.21: NC likelihood spectra for real data and MC best fit for $E_{vis} > 15\,\text{GeV}$. The error bars for real data are exaggerated by a factor of 3.

is negligible small (see Figure 5.22, which allows also a verification with real data). For the reconstruction of the visible energy, the algorithm described in [65] is used, which achieves a resolution of $< 30\%$. A sufficient energy resolution is important for the oscillation search described in the next section. The standard algorithm for the energy reconstruction in OPERA has been tested as well, which delivers a compatible but less significant result.

The determination of the NC fraction is done by varying the NC fraction and fitting the output spectrum of the NC network to real data using χ^2 minimisation: first, two histograms are filled with the NC-likelihood of ν_μ CC and ν_μ NC events, respectively. Both histograms additionally contain each half of the contributions from other neutrino flavours, so that a normalised mixture of both histograms always includes the correct amount of non-ν_μ events. The mixing ratio of the two histograms is equivalent with the NC fraction for ν_μ and is used as the free parameter in the χ^2 fit.

The resulting χ^2 distribution in dependence of the NC fraction is shown in Figure 5.20. Figure 5.21 shows the NC likelihood spectra for real data and the best fit MC data. Using this method, the NC fraction is determined as

$$\left.\frac{N_{\mathrm{NC}}}{N_{\mathrm{NC}}+N_{\mathrm{CC}}}\right|_{E_{\mathrm{vis}}>15\,\mathrm{GeV}} = 0.241^{+0.022}_{-0.023}. \tag{5.5}$$

The expected NC fraction, as calculated in Chapter 3 (see Table 3.1), is

$$\frac{N_{\mathrm{NC}}}{N_{\mathrm{NC}}+N_{\mathrm{CC}}} = \frac{0.296}{0.296+1} = 0.228, \tag{5.6}$$

which is consistent with the measured value within its uncertainty. As mentioned in Section 5.1.3, the training of the neural network was performed with equal amounts of NC and CC events and the expected NC fraction has not been introduced into the analysis at any point. Therefore, this result proves that this method for the determination of the NC fraction is producing reliable results.

5.4.3 NC fraction in energy dependence

To observe effects from neutrino oscillations, the NC fraction has been determined in energy bins with a width of 6 GeV. This bin width has been tuned with MC data to optimise the significance of the result. The method described in Section 5.4.2 has been applied to each energy bin. Again, the algorithm described in [65] is used for the energy reconstruction. The result for real data of the OPERA run years 2008 to 2011 in comparison with MC is shown in Figure 5.22. The MC data has been tuned to match the unoscillated NC fraction as measured in 5.4.2 for a more conservative estimation of the significance of this measurement.

For the observation of oscillations, the first energy bin is the most important one, as the mean oscillation probability is the largest in this energy bin. The χ^2 distribution for the fit of the first bin is shown in Figure 5.23, and the corresponding NC likelihood spectra for real data and MC best fit are shown in Figure 5.24. The NC fraction for $E_{\mathrm{vis}} < 6\,\mathrm{GeV}$ is determined as

$$\left.\frac{N_{\mathrm{NC}}}{N_{\mathrm{NC}}+N_{\mathrm{CC}}}\right|_{E_{\mathrm{vis}}<6\,\mathrm{GeV}} = 0.302^{+0.067}_{-0.049}. \tag{5.7}$$

Under the assumption of maximal mixing ($\sin 2\Theta_{23} = 1$), Figure 5.25 shows the NC fraction for the first energy bin for a given mass splitting as expected from a MC calculation. This can be used to determine the mass splitting from the measured NC fraction as

$$|\Delta m^2_{32}| = \left(2.8^{+1.3}_{-1.5}\ (\mathrm{lowE})\ ^{+0.6}_{-0.7}\ (\mathrm{highE})\ \right)\cdot 10^{-3}\,\mathrm{eV}^2, \tag{5.8}$$

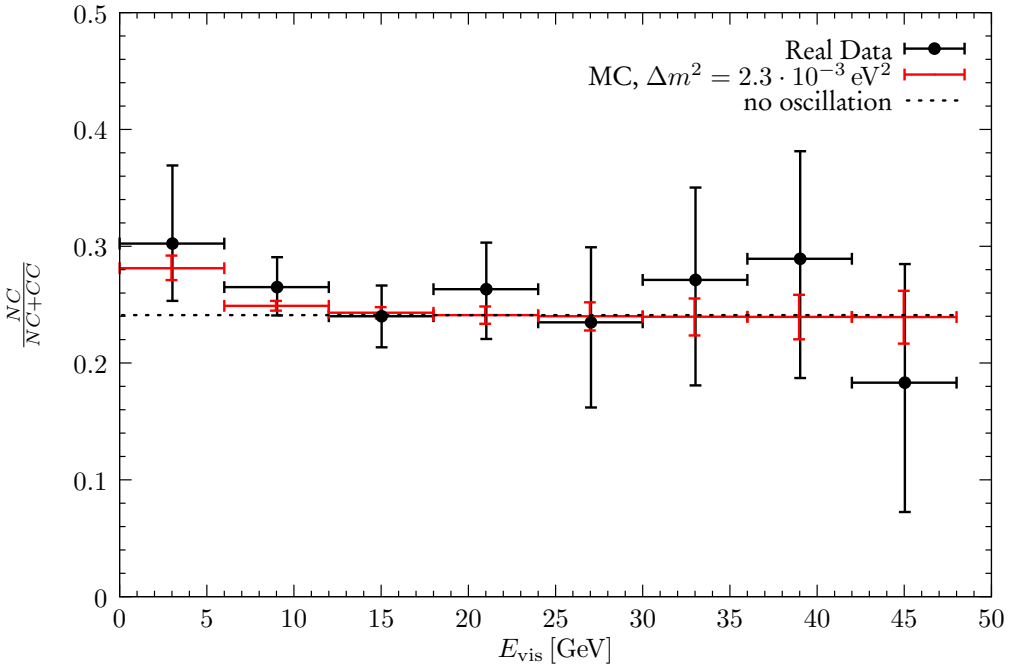

Figure 5.22: NC fraction vs. visible energy. The horizontal error bars indicate the energy bin width and do not include the resolution of the energy reconstruction. The dashed line ("no oscillation") represents the measured NC fraction in real data without oscillation effects (see Section 5.4.2). The MC data has been tuned to match this NC fraction.

with "lowE" denominating the error from the measurement in the low-energy bin $E_{\mathrm{vis}} < 6\,\mathrm{GeV}$ and "highE" denominating the error from the measurement of the unoscillated NC fraction at high energies (Section 5.4.2). As both measurements are statistically independent, the errors are added quadratically:

$$|\Delta m^2_{32}| = \left(2.8^{+1.4}_{-1.7}\right) \cdot 10^{-3}\,\mathrm{eV}^2. \tag{5.9}$$

Ambiguous solutions with higher mass splittings for the first energy bin can be excluded by the results for the higher-energy bins (see Figure 5.22).

The significance for an observation of neutrino oscillations with this result can be determined by assuming the null hypothesis

$$\left.\frac{N_{\mathrm{NC}}}{N_{\mathrm{NC}} + N_{\mathrm{CC}}}\right|_{E_{\mathrm{vis}}<6\,\mathrm{GeV}} = \left.\frac{N_{\mathrm{NC}}}{N_{\mathrm{NC}} + N_{\mathrm{CC}}}\right|_{E_{\mathrm{vis}}>15\,\mathrm{GeV}}. \tag{5.10}$$

The result is a combination of the two measurements of the NC fraction shown in Section 5.4.2 and in this section. Therefore, the significance α has been

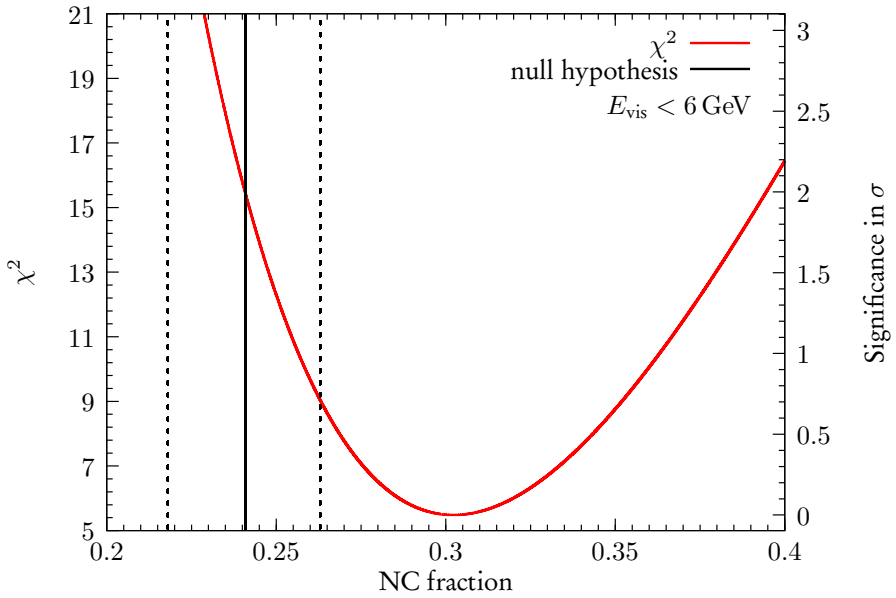

Figure 5.23: χ^2 over NC fraction for $E_{vis} < 6$ GeV. The number of degrees of freedom for this fit is 5. The vertical line ("null hypothesis") indicates the unoscillated NC fraction, as predicted in Chapter 3, the dashed lines indicate its 68% confidence interval.

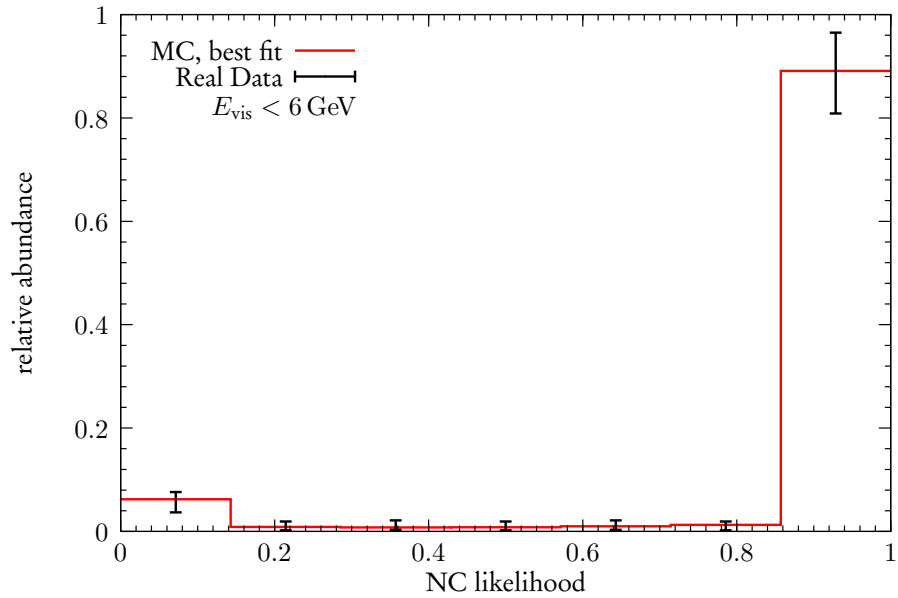

Figure 5.24: NC likelihood spectra for real data and MC best fit for $E_{vis} < 6$ GeV. The error bars for real data are exaggerated by a factor of 3.

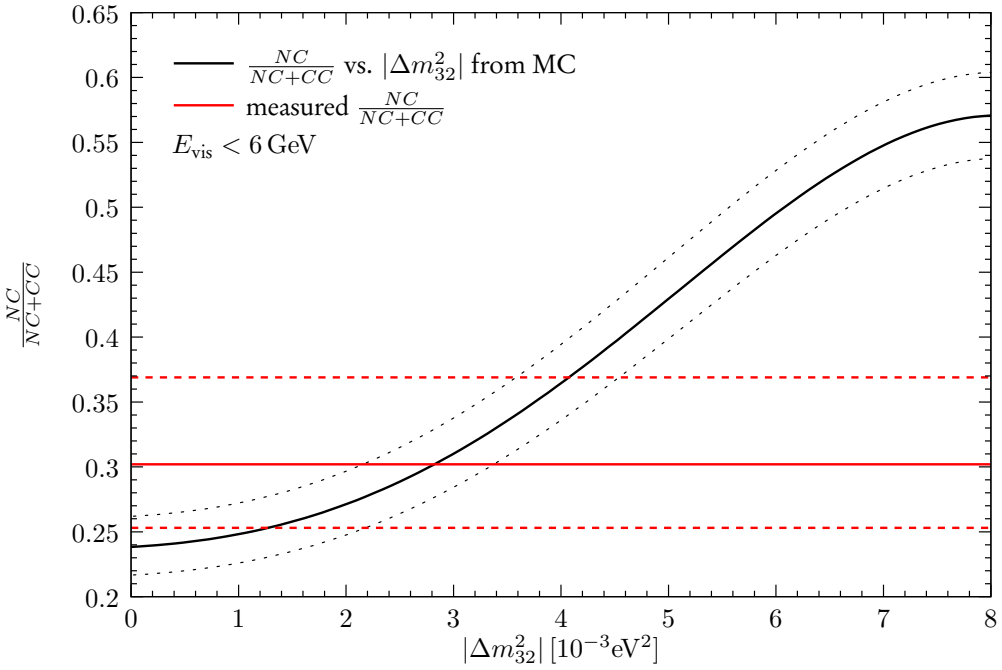

Figure 5.25: NC fraction vs. Δm^2_{23} for the lowest energy bin, assuming maximal mixing ($\sin 2\Theta_{23} = 1$). The horizontal line indicates the measured NC fraction for this energy bin. The intersection of the two curves gives the resulting mass-splitting. The dashed lines indicate the 68% confidence intervals.

determined by integrating the product of the probability for a certain unoscillated NC fraction $p(r_{\text{NC}})$ to be the true value multiplied by the significance $\alpha(r_{\text{NC}})$ of this measurement for the true NC fraction r_{NC}:

$$\alpha = \int_{-\infty}^{\infty} p(r_{\text{NC}})\alpha(r_{\text{NC}})dr_{NC}. \tag{5.11}$$

The resulting significance is 87% or 1.5σ. Systematic effects are partially included due to their treatment described in Section 5.1.2, which also reduces the separation power of the neural network slightly and thus increases the specified errors. Additional systematic effects, which cannot be controlled due to shortcomings of the MC software of OPERA, might come from the energy reconstruction. As the analysis is only sensitive to differences between MC and real data, those effects are probably smaller than the statistical uncertainty.

The result is compatible with the current global best fit $|\Delta m^2_{32}| = (2.32^{+0.12}_{-0.08}) \cdot 10^{-3}\,\text{eV}^2$ (see Section 1.3.1). It currently presents the best determination of $|\Delta m^2_{32}|$ that has been performed with the OPERA experiment.

Chapter 6

Conclusions

Neutrino oscillations are a very well established explanation for several phenomena observed by many neutrino experiments in the last decades. While other models like a neutrino decay are excluded in the solar sector by appearance measurements, a comparable observation in the atmospheric sector is still pending. The OPERA experiment is the only experiment currently running which can do an appearance measurement of ν_τ from a ν_μ source. The result of this experiment is therefore crucial to consolidate our current picture of neutrino physics.

The OPERA experiment has so far observed one ν_τ candidate event. For a significant observation of ν_τ appearance a larger statistics is required. The full analysis of neutrino interactions in the OPERA detector takes significant time and especially the analysis of candidate events take additional time. The typical delay between the occurrence of a neutrino interaction in the OPERA experiment and the completion of its analysis is in the order of one year.

Due to the recent hints for a large Θ_{13} it seems possible that the OPERA experiment can perform an appearance measurement of ν_e in the ν_μ beam as well. To be competitive, a fast analysis of ν_e events is important.

Initially there was no suppression of background neutrino interactions in the target (e.g. from ν_μ) with informations from the electronic detectors foreseen. For parts of the 2010 data a very basic algorithm for such suppression, involving prioritising NC-like and soft muon events, was implemented. This algorithm results only in a mild background suppression while introducing a large additional systematical uncertainty to efficiency estimations.

In this thesis, an improved algorithm to prioritise ν_τ and ν_e events over ν_μ background has been presented. This algorithm is based on a neural network, which returns the probability for the event being a signal event. This probability can be used to rank the neutrino events in OPERA. If this ranking is applied, almost all signal events can be found within the first 30% of the data sample. Only the $\tau \to \mu$ decay channel shows a non-negligible contribution in the remaining 70% of the sample, even though the described ranking algorithm will also present a significant speed increase for this channel. Due to the large delay between event recording and completion of the analysis, the speed increase for the analysis will not be significantly affected by the fact, that data taking of the OPERA experiment will still continue for at least one year. The application of this event ranking for the official analysis of the OPERA experiment is currently being discussed by the collaboration. The ranking is already used for testing purposes by several groups, which confirmed the robustness of the used algorithm.

The neural network can also be used to discriminate NC and CC events. This has been used to determine the fraction of NC events without effects from neutrino oscillations to $N_{\mathrm{NC}}/(N_{\mathrm{NC}} + N_{\mathrm{CC}}) = 0.241^{+0.022}_{-0.023}$, which is consistent with the theoretical prediction and therefore proves that the used algorithm is working as expected. Neutrino oscillations increase the NC ratio seen in OPERA for low energies. This expectation has been confirmed by the presented analysis. The confidence level of the measurement presenting a hint for neutrino oscillations is 87%. The determined mass splitting $|\Delta m^2_{32}| = 2.8^{+1.4}_{-1.7} \cdot 10^{-3}\,\mathrm{eV}^2$ (assuming maximal mixing) is compatible with the current global best fit. This is the best measurement of this oscillation parameter with the OPERA experiment so far.

Appendix A

Regression-type neural networks

A Multi-Layer Perceptron similar to the one described in Chapter 4 can be used for a model-independent fit of a continuous function. Any continuous function in $\mathbb{R}^N$ can be approximated to an arbitrary precision by a linear combination of the logistic function [54].

A two-layer MLP (i.e. an MLP with one single hidden layer) with a linear activation function for the output neurons ($g(x) = x$) can then be described as a linear combination of the activation function z_{1j} of the hidden layer:

$$y_k(\mathbf{x}) = \sum_j w_{2kj} z_{1j}(\mathbf{x}). \tag{A.1}$$

Hence, a two layer MLP with the logistic function as activation function for the hidden neurons and linear output neurons can be used for solving the regression problem.

The neural network can have multiple output neurons, in which case it will be approximating a multi-dimensional function. In the following, the vector of output values will be denominated as $\mathbf{y}$ and the vector of target values as $\mathbf{t}$. The (typically unknown) function to be approximated is called $\mathbf{h}(\mathbf{x})$. The remaining designations will be as in Chapter 4.

A.1 Error function for regression

The approach to find the correct error function for training the regression MLP is similar to the classification case described in Section 4.4. The analogue to Equation (4.5) is in this case:

$$\mathcal{L} = \prod_n p(\mathbf{t}^{(n)}|\mathbf{x}^{(n)}). \tag{A.2}$$

In the following, the errors of the components of the output vector are assumed to be uncorrelated, which leads to:

$$p(\mathbf{t}|\mathbf{x}) = \prod_k p(t_k|\mathbf{x}). \tag{A.3}$$

The correct choice of the error function for modelling the unknown function $\mathbf{h}(\mathbf{x})$ with the MLP depends on the type of error of the target variables t_k. Assuming Gaussian errors:

$$t_k = h_k(\mathbf{x}) + \varepsilon_k, \tag{A.4}$$

with ε_k being normally distributed:

$$p(\varepsilon_k) \sim \exp\left(-\frac{\varepsilon_k^2}{2\sigma^2}\right), \tag{A.5}$$

the probability density for the target variables can be obtained by replacing the unknown function $\mathbf{h}(\mathbf{x})$ with our model $\mathbf{y}(\mathbf{x};\mathbf{w})$, depending also on the set of weights $\mathbf{w}$:

$$p(t_k|\mathbf{x}) \sim \exp\left(-\frac{[y_k(\mathbf{x};\mathbf{w}) - t_k]^2}{2\sigma^2}\right). \tag{A.6}$$

As in the classification case, the negative log-likelihood is used as the error function:

$$E = -\ln \mathcal{L} = -\sum_n \ln p(\mathbf{t}^{(n)}|\mathbf{x}^{(n)}). \tag{A.7}$$

Together with (A.3) and omitting constant factors and addends, as they are not important for a minimization problem, this leads to the well-known *quadratic error function*:

$$E = \frac{1}{2}\sum_n \|\mathbf{y}(\mathbf{x}^{(n)};\mathbf{w}) - \mathbf{t}^{(n)}\|^2, \tag{A.8}$$

which is therefore the correct choice for the error function for the training of a regression-type neural network.

A.2 Back-propagation of errors

The back-propagation of errors for a regression-type network works in the same way as for a classification-type network. Here, the calculation for the synapses of the last layer L is straight forward: Equations (A.8) and (4.1) directly lead to (4.16). Hence for both application scenarios the same formulae to calculate $\frac{\partial e}{\partial w}$ for the output layer can be used.

Appendix B

Spectra of input variables

In this appendix the spectra of the input values used for the neural network are shown. The structure of the neural network was explained in Section 5.1.1. The spectra are compared between MC and real data. The differences between the spectra indicate a systematic error of the used MC. The method to reduce the influence of these errors described in Section 5.1.2 is based on the spectra shown here.

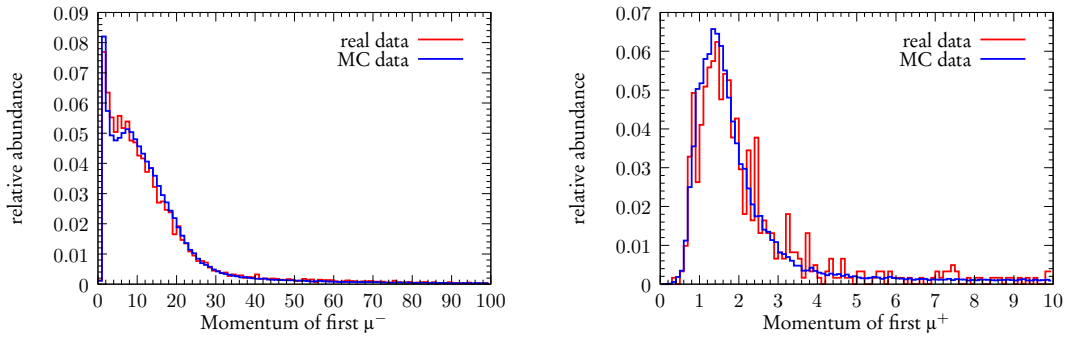

Figure B.1: Momenta of first μ^- and first μ^+.

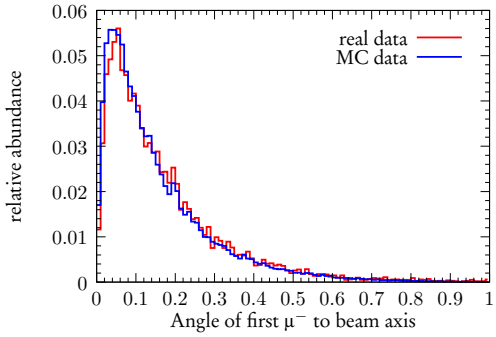

Figure B.2: Angle to beam axis of first μ^-.

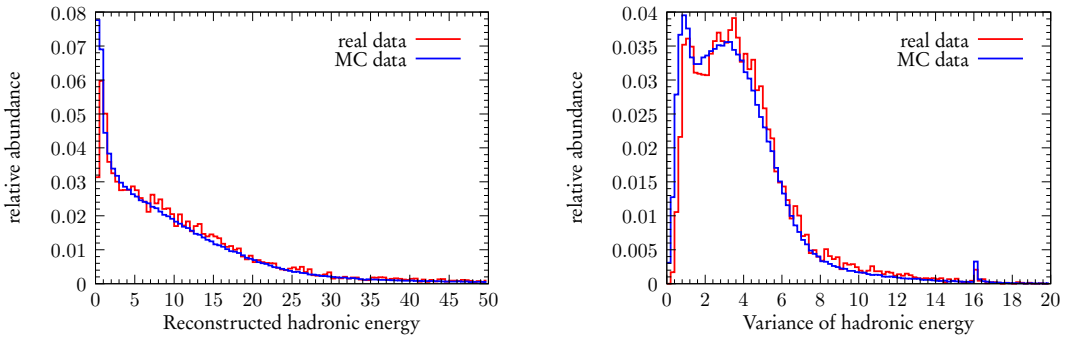

Figure B.3: Reconstructed hadronic energy and its variance.

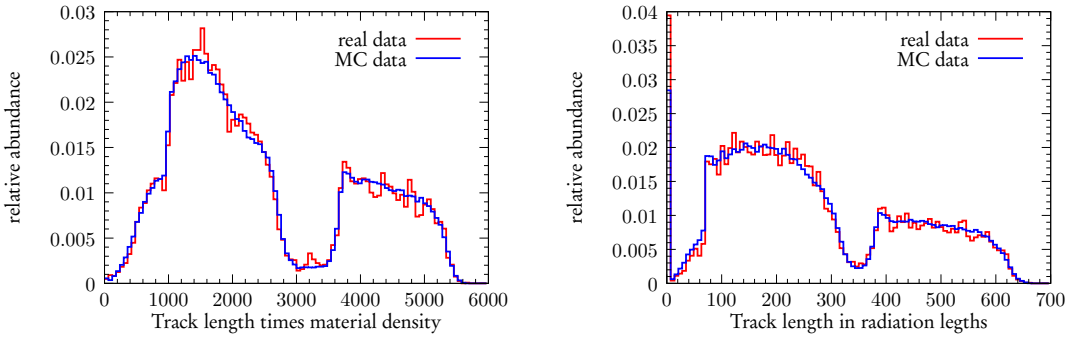

Figure B.4: Length of longest track multiplied with mean material density along its path and length of longest track in radiation lengths.

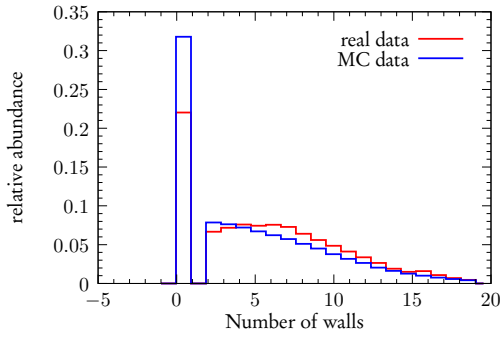

Figure B.5: Number of target tracker walls hit by the shower.

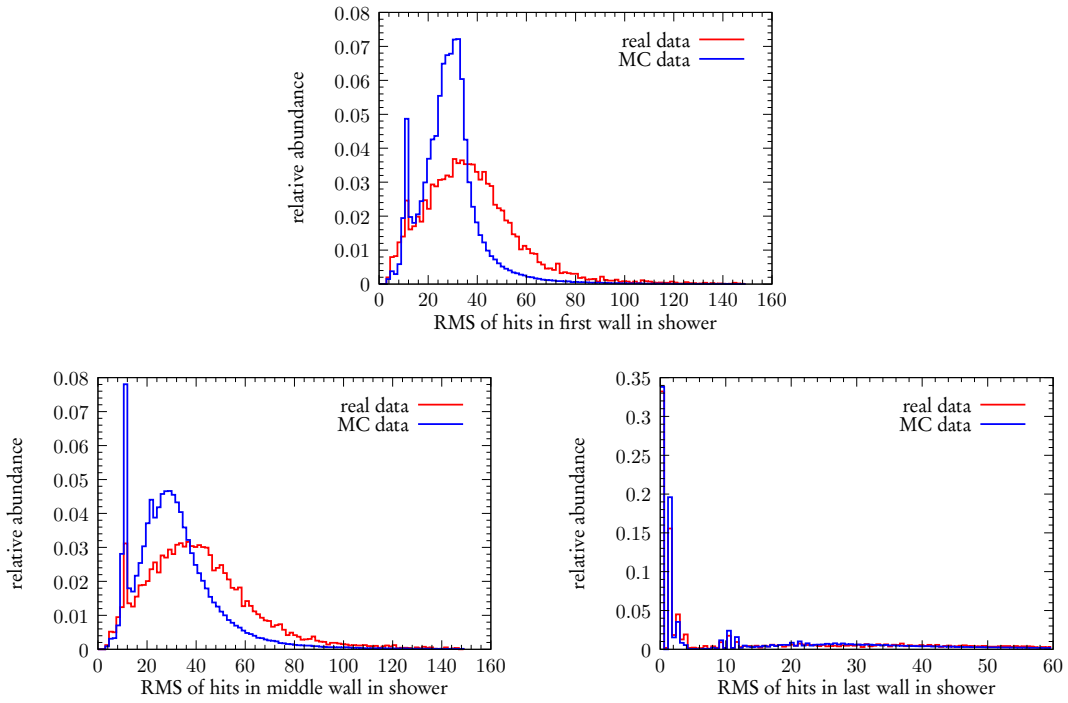

Figure B.6: RMS of the hit distribution in the first, middle and last target tracker wall hit by the shower.

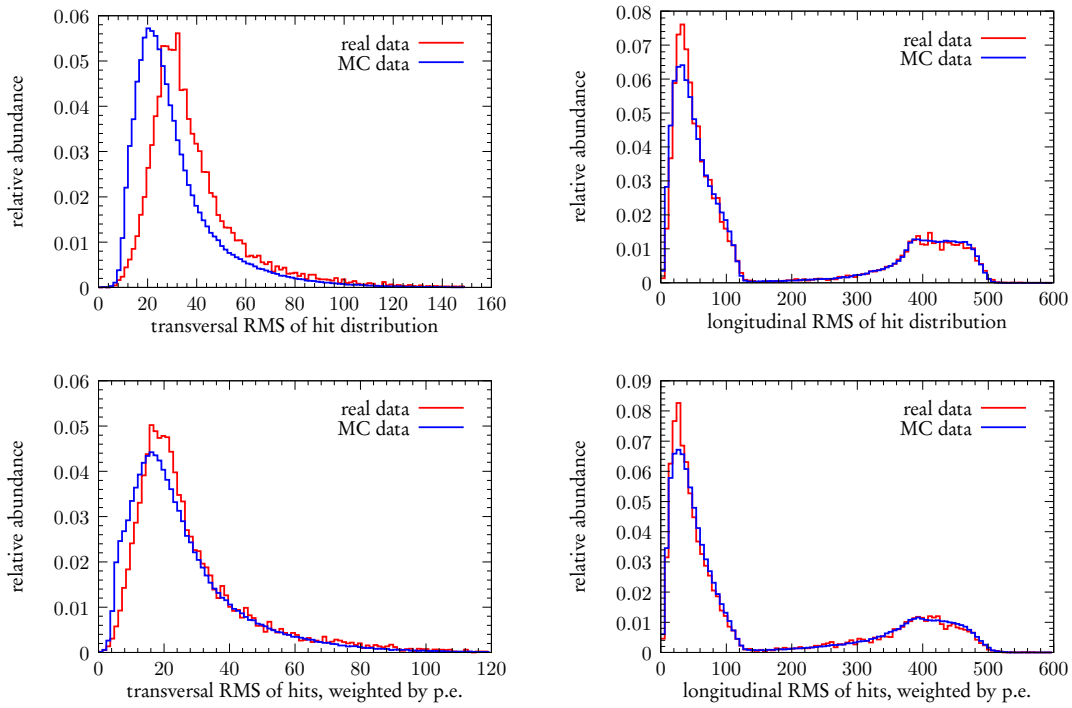

Figure B.7: Longitudinal and transversal RMS of the distributions of all target tracker hits, unweighted and weighted with the number of photoelectrons and.

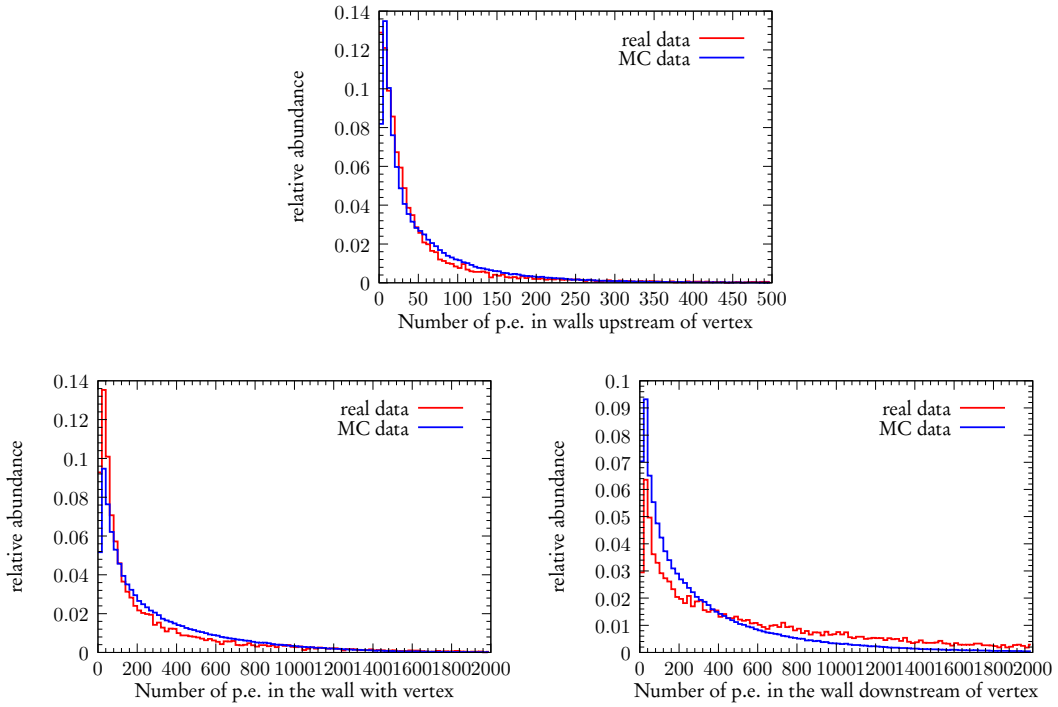

Figure B.8: Number of photoelectrons in the wall with the reconstructed interaction point, in all walls upstream and in the next wall downstream to the interaction point.

Appendix C

Quality control of the neural network training

As described in Section 5.1.3, the training process of the neural network was monitored for negative effects. The two most important cross checks are the observation of the cross-entropy error as defined in Section 4.4 in dependence of the training epoch, as well as the unity of the output y of the neural network and the a-posteriori probability $p(\mathcal{C}_1|y)$. Figures C.1 to C.4 show the cross-entropy error over the epoch, Figures C.5 to C.8 the true likelihood over the output of the neural network. The cross-entropy error for the test and for the training samples should match (see Section 5.1.3). Deviations from the unity $y = p(\mathcal{C}_1|y)$ indicate that the neural network is not large enough to achieve the desired precision of the approximation. As long as the mapping is monotonic ranking of events is not effected by possible deviations.

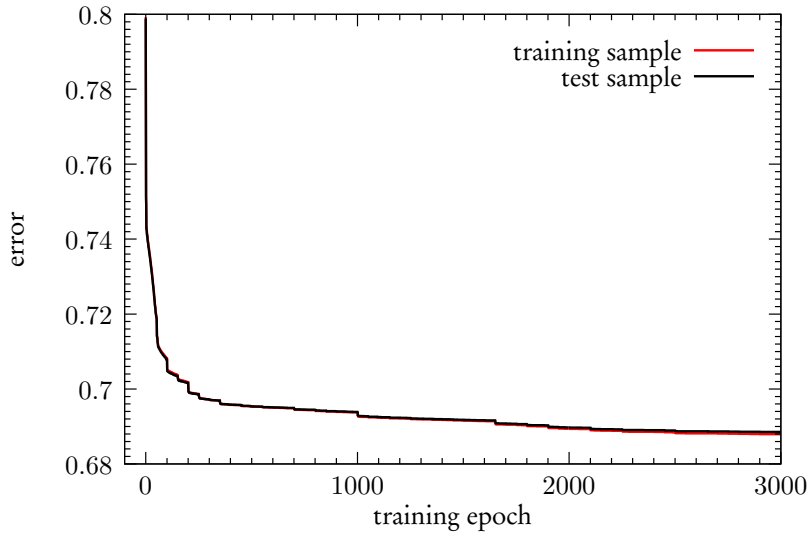

Figure C.1: Learning curve of the ν_τ network.

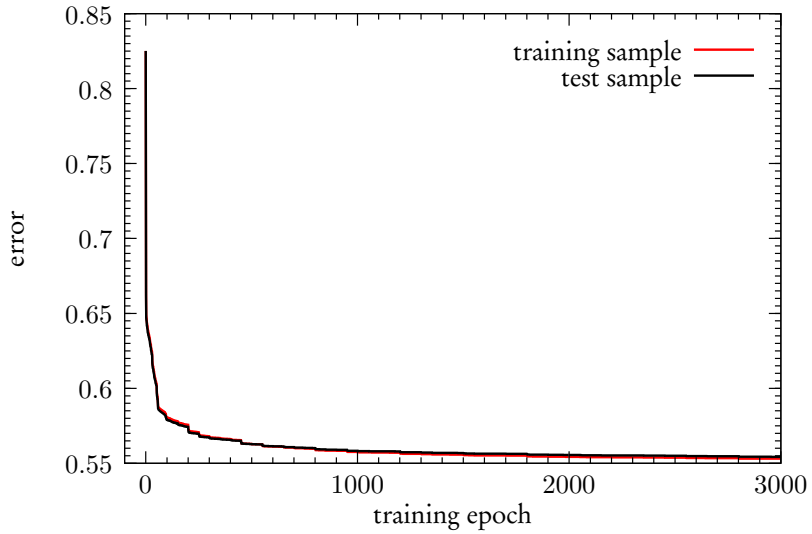

Figure C.2: Learning curve of the ν_e network.

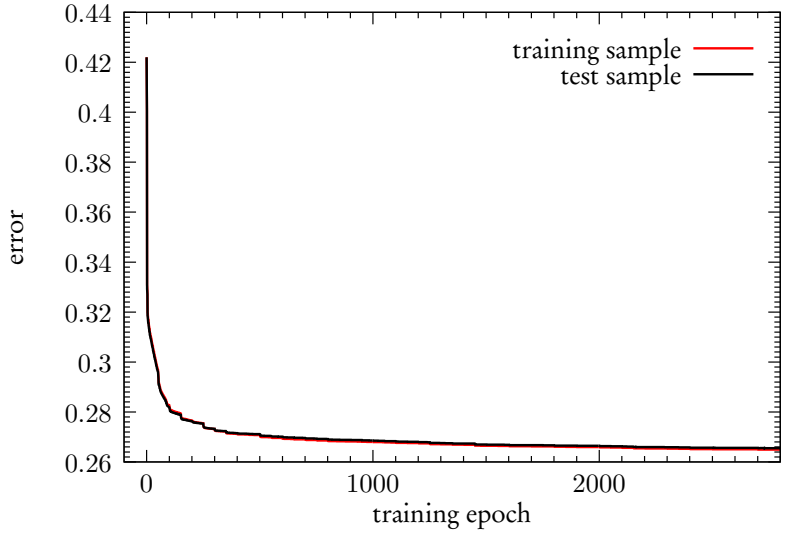

Figure C.3: Learning curve of the combined ν_τ and ν_e network.

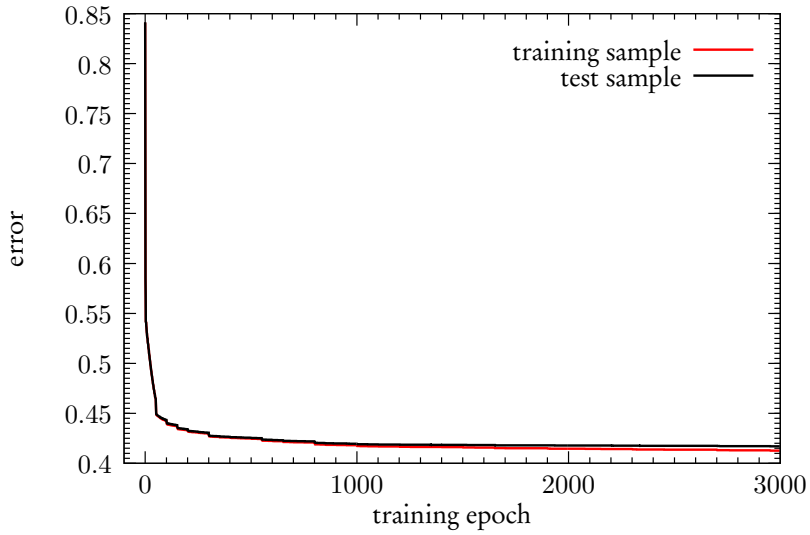

Figure C.4: Learning curve of the NC network.

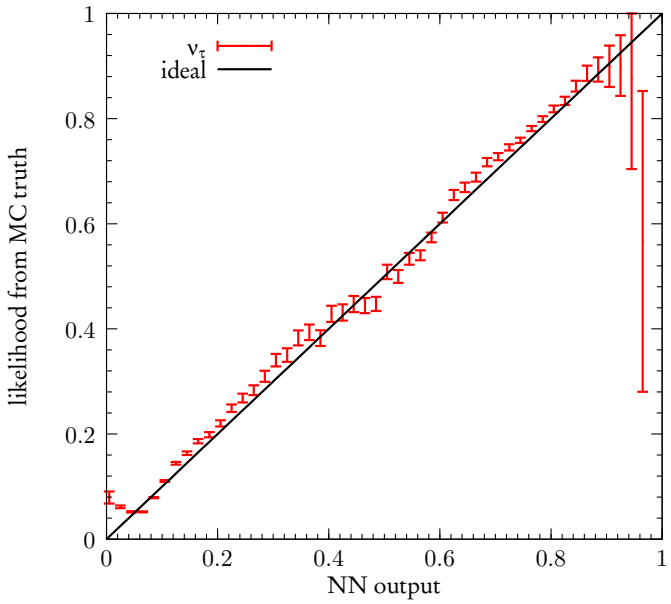

Figure C.5: True likelihood over output of the ν_τ neural network. No significant deviation from the ideal line is visible.

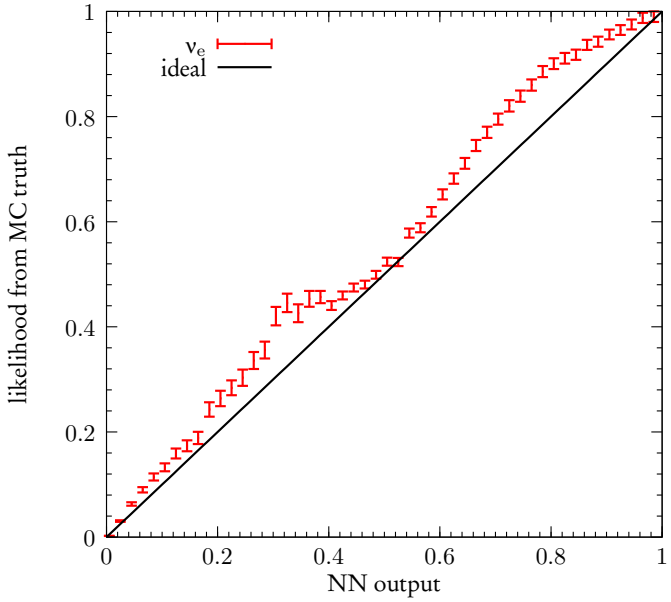

Figure C.6: True likelihood over output of the ν_e neural network. No problematic deviation from the ideal line is visible, as the mapping is monotonic.

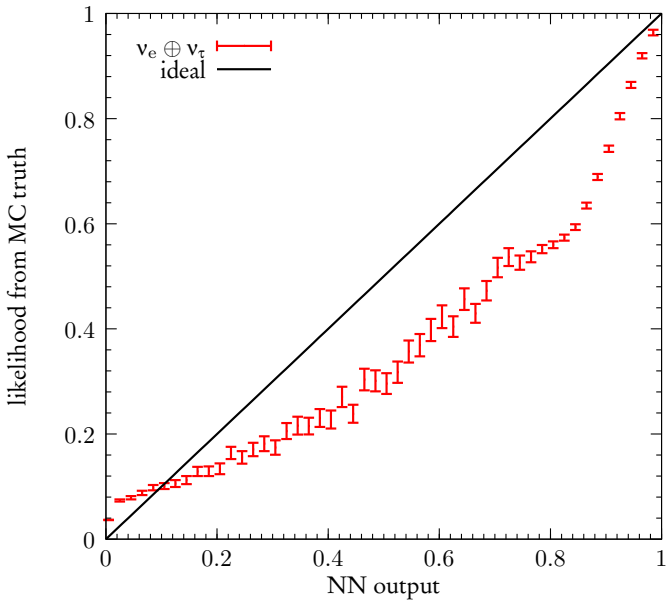

Figure C.7: True likelihood over output of the combined ν_τ and ν_e neural network. No problematic deviation from the ideal line is visible, as the mapping is monotonic.

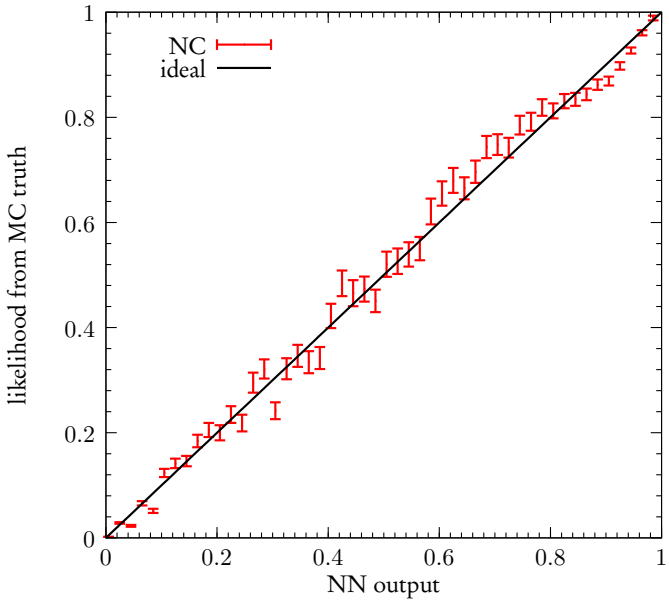

Figure C.8: True likelihood over output of the NC neural network. No significant deviation from the ideal line is visible.

Bibliography

[1] Final report of the EURISOL design study (2005-2009), a Design study for a European Isotope-Separation-On-Line radioactive ion beam facility. 2009. Full report on accompanying CD in cover pocket.

[2] Momentum measurement by the Multiple Coulomb Scattering method in the OPERA lead emulsion target. 2011, physics.ins-det/1106.6211.

[3] K. Abe, T. Abe, H. Aihara, Y. Fukuda, Y. Hayato, et al. Letter of Intent: The Hyper-Kamiokande Experiment — Detector Design and Physics Potential —. 2011, hep-ex/1109.3262.

[4] K. Abe et al. Solar neutrino results in Super-Kamiokande-III. *Phys.Rev.*, D83:052010, 2011, hep-ex/1010.0118.

[5] K. Abe and J.B. Spitz. The T2K Experiment. *Nucl.Instrum.Meth.*, A659:106–135, 2011, physics.ins-det/1106.1238.

[6] R. Acquafredda, T. Adam, N. Agafonova, P. Alvarez Sanchez, M. Ambrosio, et al. The OPERA experiment in the CERN to Gran Sasso neutrino beam. *JINST*, 4:P04018, 2009.

[7] T. Adam et al. Measurement of the neutrino velocity with the OPERA detector in the CNGS beam. 2011, hep-ex/1109.4897.

[8] P. Adamson et al. Measurement of neutrino velocity with the MINOS detectors and NuMI neutrino beam. *Phys.Rev.*, D76:072005, 2007, hep-ex/0706.0437.

[9] P. Adamson et al. Improved search for muon-neutrino to electron-neutrino oscillations in MINOS. *Phys.Rev.Lett.*, 107:181802, 2011, 1108.0015.

[10] P. Adamson et al. Measurement of the neutrino mass splitting and flavor mixing by MINOS. *Phys.Rev.Lett.*, 106:181801, 2011, hep-ex/1103.0340.

[11] N. Agafonova et al. Observation of a first ν_τ candidate in the OPERA experiment in the CNGS beam. *Phys.Lett.*, B691:138–145, 2010, hep-ex/1006.1623.

[12] A. Aguilar et al. Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam. *Phys.Rev.*, D64:112007, 2001, hep-ex/0104049.

[13] A.A. Aguilar-Arevalo et al. A Search for electron neutrino appearance at the $\Delta m^2 \sim 1\text{eV}^2$ scale. *Phys.Rev.Lett.*, 98:231801, 2007, hep-ex/0704.1500.

[14] B. Aharmim et al. Electron energy spectra, fluxes, and day-night asymmetries of B-8 solar neutrinos from measurements with NaCl dissolved in the heavy-water detector at the Sudbury Neutrino Observatory. *Phys.Rev.*, C72:055502, 2005, nucl-ex/0502021.

[15] B. Aharmim et al. An Independent Measurement of the Total Active B-8 Solar Neutrino Flux Using an Array of He-3 Proportional Counters at the Sudbury Neutrino Observatory. *Phys.Rev.Lett.*, 101:111301, 2008, nucl-ex/0806.0989.

[16] Carl H. Albright and C. Jarlskog. Neutrino Production of m+ and e+ Heavy Leptons. 1. *Nucl. Phys.*, B84:467, 1975.

[17] G. Alimonti et al. The Borexino detector at the Laboratori Nazionali del Gran Sasso. *Nucl.Instrum.Meth.*, A600:568–593, 2009, 0806.2400.

[18] C. Andreopoulos et al. The GENIE Neutrino Monte Carlo Generator. *Nucl. Instrum. Meth.*, A614:87–104, 2010, hep-ph/0905.2517.

[19] M. Apollonio et al. Limits on neutrino oscillations from the CHOOZ experiment. *Phys.Lett.*, B466:415–430, 1999, hep-ex/9907037.

[20] B. Armbruster et al. Upper limits for neutrino oscillations muon-anti-neutrino $\rightarrow$ electron-anti-neutrino from muon decay at rest. *Phys.Rev.*, D65:112001, 2002, hep-ex/0203021.

[21] Arrabito, L. and others. Track reconstruction in the emulsion-lead target of the OPERA experiment using the ESS microscope. *Journal of Instrumentation*, 2(05):P05004, 2007.

[22] S. Assylbekov, B.E. Berger, H. Berns, D. Beznosko, A. Bodek, et al. The T2K ND280 Off-Axis Pi-Zero Detector. 2011, physics.ins-det/1111.5030.

[23] D. Autiero, I. Efthymiopoulos, A. Ferrari, E. Gschwendtner, A. Guglielmi, et al. The CNGS neutrino beam: Status. *Nucl.Phys.Proc.Suppl.*, 188:188–190, 2009.

[24] D. Ayres et al. Letter of intent to build an off-axis detector to study $\nu_\mu \to \nu_e$ oscillations with the NuMI neutrino beam. 2002, hep-ex/0210005.

[25] D.S. Ayres et al. NOvA: Proposal to build a 30 kiloton off-axis detector to study nu(mu) —> nu(e) oscillations in the NuMI beamline. 2004, hep-ex/0503053. Updated version of 2004 proposal. Higher resolution version available at Fermilab Library Server.

[26] Ch. J. Backhouse. Measuring neutrino oscillation parameters using ν_μ disappearance in minos, 2011.

[27] A. Badertscher, A. Curioni, U. Degunda, L. Epprecht, S. Horikawa, et al. Giant Liquid Argon Observatory for Proton Decay, Neutrino Astrophysics and CP-violation in the Lepton Sector (GLACIER). 2010, 1001.0076.

[28] John N. Bahcall, Aldo M. Serenelli, and Sarbani Basu. New solar opacities, abundances, helioseismology, and neutrino fluxes. *Astrophys.J.*, 621:L85–L88, 2005, astro-ph/0412440.

[29] R. Barate et al. Measurement of the Z resonance parameters at LEP. *Eur.Phys.J.*, C14:1–50, 2000.

[30] R. Michael Barnett. Evidence for new quarks and new currents. *Phys. Rev. Lett.*, 36(20):1163–1166, May 1976.

[31] A. Bertolin and N.T. Tran. OpCarac: an algorithm for the classification of the neutrino interactions recorded by the OPERA experiment. Technical report, OPERA note 100, 25/05/2009.

[32] Ch. M. Bishop. Neural Networks: A Pattern Recognition Perspective. Technical report, Neural Network Research Group, Aston University, Birmingham, UK, January 1996.

[33] R. Brun and F. Rademakers. ROOT - An Object Oriented Data Analysis Framework. In *proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997)*, pages 81–86. See also http://root.cern.ch/.

[34] The Borexino Collaboration. First evidence of pep solar neutrinos by direct detection in Borexino. 2011, hep-ex/1110.3230.

[35] The OPERA Collaboration. Search for $\nu_\mu \rightarrow \nu_\tau$ oscillation with the OPERA experiment in the CNGS beam. 2011, hep-ex/1107.2594.

[36] C.L. Cowan, F. Reines, F.B. Harrison, H.W. Kruse, and A.D. McGuire. Detection of the free neutrino: A Confirmation. *Science*, 124:103–104, 1956.

[37] Alessandro Curioni. GLACIER and related R&D. *J.Phys.Conf.Ser.*, 309:012012, 2011, 1107.0825.

[38] R. Dharmapalan et al. A Request for Further MiniBooNE Antineutrino Running, 2010.

[39] Marcos Dracos. Neutrinos from CERN to Gran Sasso: The CNGS project. *J.Phys.Conf.Ser.*, 203:012013, 2010.

[40] Marcos Dracos. The Double Chooz experiment. *AIP Conf.Proc.*, 1382:135–137, 2011.

[41] J. Altegoer et al. The nomad experiment at the cern sps. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 404(1):96 – 128, 1998.

[42] K. Nakamura et al. (Particle Data Group). *J. Phys.*, G 37:075021, 2010.

[43] K. Nakamura et al. (Particle Data Group). *J. Phys.*, G 37:075021, 2010. and 2011 partial update for the 2012 edition.

[44] A Ferrari, A M Guglielmi, M Lorenzo-Sentis, S Roesler, P R Sala, and L Sarchiapone. An updated Monte Carlo calculation of the CNGS neutrino beam. Technical Report AB-Note-2006-038. CERN-AB-Note-2006-038, CERN, Geneva, Aug 2007.

[45] G.L. Fogli, E. Lisi, A. Marrone, A. Palazzo, and A.M. Rotunno. Hints of $\Theta_{13} > 0$ from global neutrino data analysis. *Phys.Rev.Lett.*, 101:141801, 2008, hep-ph/0806.2649.

[46] Tsutomu Fukuda. The Nuclear Emulsion Technology and the Analysis of the OPERA Experiment Data. 2009, hep-ex/0910.3274.

[47] Y. Fukuda et al. Measurements of the solar neutrino flux from Super-Kamiokande's first 300 days. *Phys.Rev.Lett.*, 81:1158–1162, 1998, hep-ex/9805021.

[48] A. Gando et al. Constraints on θ_{13} from A Three-Flavor Oscillation Analysis of Reactor Antineutrinos at KamLAND. *Phys.Rev.*, D83:052002, 2011, hep-ex/1009.4771.

[49] Carlo Giunti and Marco Laveder. First Double-Chooz Results and the Reactor Antineutrino Anomaly. 2011. See also references cited within this publication., hep-ph/1111.5211.

[50] M.C. Gonzalez-Garcia and Yosef Nir. Neutrino masses and mixing: Evidence and implications. *Rev.Mod.Phys.*, 75:345–402, 2003, hep-ph/0202058.

[51] E Gschwendtner, K Cornelis, I Efthymiopoulos, A Ferrari, A Pardons, W Treberspurg, H Vincke, J Wenninger, D Autiero, A Guglielmi, and P Sala. Performance and operational experience of the cngs facility. oai:cds.cern.ch:1277917. (CERN-ATS-2010-153):3 p, Apr 2010.

[52] Xinheng Guo et al. A Precision measurement of the neutrino mixing angle theta(13) using reactor antineutrinos at Daya-Bay. 2007, hep-ex/0701029.

[53] A. Habig. MINOS neutrino oscillation results. *Nucl.Phys.Proc.Suppl.*, 218:320–325, 2011.

[54] Yoshifusa Ito. Approximation of continuous functions on rd by linear combinations of shifted rotations of a sigmoid function with and without scaling. *Neural Networks*, 5(1):105 – 115, 1992.

[55] Alexander Izmaylov. New Oscillation Results from the T2K experiment. 2011, hep-ex/1112.0273.

[56] Nick Jelley, Arthur B. McDonald, and R.G.Hamish Robertson. The Sudbury Neutrino Observatory. *Ann.Rev.Nucl.Part.Sci.*, 59:431–465, 2009.

[57] Benjamin J P Jones. The Status of the MicroBooNE Experiment. 2011, physics.ins-det/1110.1678.

[58] H.V. Klapdor-Kleingrothaus, A. Dietz, L. Baudis, G. Heusser, I.V. Krivosheina, et al. Latest results from the Heidelberg-Moscow double beta decay experiment. *Eur.Phys.J.*, A12:147–154, 2001, hep-ph/0103062.

[59] H.V. Klapdor-Kleingrothaus, A. Dietz, H.L. Harney, and I.V. Krivosheina. Evidence for neutrinoless double beta decay. *Mod.Phys.Lett.*, A16:2409–2420, 2001, hep-ph/0201231.

[60] H.V. Klapdor-Kleingrothaus and I.V. Krivosheina. The evidence for the observation of onu beta beta decay: The identification of onu beta beta events from the full spectra. *Mod.Phys.Lett.*, A21:1547–1566, 2006.

[61] K. Kodama et al. Final tau-neutrino results from the DONuT experiment. *Phys.Rev.*, D78:052002, 2008, hep-ex/0711.0728.

[62] Ch. Kraus, B. Bornschein, L. Bornschein, J. Bonn, B. Flatt, et al. Final results from phase II of the Mainz neutrino mass search in tritium beta decay. *Eur.Phys.J.*, C40:447–468, 2005, hep-ex/0412056.

[63] S. Kretzer and M. H. Reno. Tau neutrino deep inelastic charged current interactions. *Phys. Rev. D*, 66(11):113007, Dec 2002.

[64] S. Kretzer and M. H. Reno. Target mass corrections to electro-weak structure functions and perturbative neutrino cross sections. *Phys. Rev.*, D69:034002, 2004, hep-ph/0307023.

[65] Belina von Krosigk. Energierekonstruktion hadronischer Schauer in den Sampling-Kalorimetern des Neutrino-Detektors OPERA. *Diploma Thesis*, 2010.

[66] Hung-Liang Lai et al. New parton distributions for collider physics. *Phys. Rev.*, D82:074024, 2010, hep-ph/1007.2241.

[67] V.M. Lobashev. The search for the neutrino mass by direct method in the tritium beta-decay and perspectives of study it in the project KATRIN. *Nucl.Phys.*, A719:153–160, 2003.

[68] Andrea Longhin and for the OPERA Collaboration. OPERA: Waiting for the tau. 2008, hep-ex/0810.2664.

[69] L. Lyons. A method of reducing systematic errors in classification problems. *Nucl. Instrum. Meth.*, A324:565–568, 1993.

[70] J. Maneira. Status and prospects of SNO+. *Nucl.Phys.Proc.Suppl.*, 217:50–52, 2011.

[71] G. Meierhofer. GERDA: A new neutrinoless double beta experiment using Ge-76. *J.Phys.Conf.Ser.*, 312:072011, 2011.

[72] D.G. Michael et al. The Magnetized steel and scintillator calorimeters of the MINOS experiment. *Nucl.Instrum.Meth.*, A596:190–228, 2008, physics.ins-det/0805.3170.

[73] Lino Miramonti. Borexino: A Real time liquid scintillator detector for low-energy solar neutrino study. pages 570–577, 2002, hep-ex/0206063.

[74] Otto Nachtmann. Positivity constraints for anomalous dimensions. *Nucl. Phys.*, B63:237–247, 1973.

[75] K. Nishikawa et al. Tokai-to-kamioka (t2k) long baseline neutrino oscillation experiment proposal, 2006.

[76] P. Novella and for the Double Chooz collaboration. Double Chooz: Searching for theta13 with reactor neutrinos. 2011, hep-ex/1105.6079.

[77] Marco Pallavicini. Borexino: Recent results, detector calibration and future perspectives. *Nucl.Phys.Proc.Suppl.*, 217:101–106, 2011.

[78] A. Pardons. Horn operational experience in K2K, MiniBooNE, NuMI and CNGS. *PoS*, NUFACT08:096, 2008.

[79] E. A. Paschos and J. Y. Yu. Neutrino interactions in oscillation experiments. *Phys. Rev.*, D65:033002, 2002, hep-ph/0107261.

[80] G. Rosa, A. Di Bartolomeo, G. Grella, and G. Romano. Automatic analysis of digitized TV images by a computer driven optical microscope. *Nucl.Instrum.Meth.*, A394:357–367, 1997.

[81] C. Rubbia, M. Antonello, P. Aprili, B. Baibussinov, M.Baldo Ceolin, et al. Underground operation of the ICARUS T600 LAr-TPC: first results. *JINST*, 6:P07011, 2011, hep-ex/1106.0975.

[82] C. Vigorito. SNEWS: The supernova early warning system. *J.Phys.Conf.Ser.*, 309:012026, 2011.

[83] F. Vissani, M.L. Costantini, W. Fulgione, A. Ianni, and G. Pagliaroli. What is the issue with SN1987A neutrinos? 2010, hep-ph/1008.4726.

[84] Christopher W. Walter. The Super-Kamiokande Experiment. 2008, hep-ex/0802.1041.

[85] Elena Wildner. Beta Beams for Neutrino Production. *AIP Conf.Proc.*, 1222:31–35, 2010.

[86] (ed.) Winter, K. Neutrino physics. *Camb.Monogr.Part.Phys.Nucl.Phys.Cosmol.*, 1991, 2nd edition 2000.

[87] J. Wolf. The KATRIN Neutrino Mass Experiment. *Nucl.Instrum.Meth.*, A623:442–444, 2010, physics.ins-det/0810.3281.

[88] L. Wolfenstein. Neutrino Oscillations in Matter. *Phys.Rev.*, D17:2369–2374, 1978.

[89] Michael Wurm et al. The next-generation liquid-scintillator neutrino observatory LENA. 2011, astro-ph.IM/1104.5620.

[90] E.D. Zimmerman. Updated Search for Electron Antineutrino Appearance at MiniBooNE. 2011, hep-ex/1111.1375.

Acknowledgement

I want to thank all people who helped me doing this analysis and writing this thesis. Special thanks deserves Caren Hagner for making this work possible in the first place and for her support, encouragement and the freedom to develop the subject of this work independently beyond the original topic.

For countless, fruitful discussions, many quick answers and sometimes just the right questions I have to thank Walter Schmidt-Parzefall. He was always willing to listen to my problems and had often just the right idea to look into a new direction.

In addition, I have to thank Dario Autiero, Giovanni De Lellis and Osam Sato for their input and support regarding the event ranking, and Thomas Strauss and Umut Köse for their quick help with the installation and use of OpEmuRec. For their help in many small (but not unimportant) aspects I also thank Björn Wonsak and Torben Ferber.

This thesis would not have been possible without my collegues from OPERA and from the neutrino physics working group in Hamburg. I thank Daniel Bick for being a nice and quiet (but not too quiet) colleague and friend sharing the office with me, Christoph "Klaus" Göllnitz for the countless hours of drinking coffee[1] with me, Björn Wonsak for the always delightful discussions during lunch, Jan Lenkeit for being *the senior* PhD student who always had a helpful opinion, Annika Hollnagel for being the blue lady and for the many cakes, Jan "Hortie" Timm for all the *Mett*, Dr. h.c. Michael Wurm for replacing me as worm, Joachim Ebert for solving all financial problems (at least of our group) and the foosball table which made many breaks during the last months joyful. I want to thank also the large number of former members of the group for their help and the good atmosphere they contributed to, especially Benjamin Janutta, Jörn Boysen, Belina von Krosigk and Stefanie Kietzmann, who still stay in touch with the group.

[1] I have to make a confession: sometimes it wasn't coffee, but tea in my cup. I hope you can forgive me...

I also have to thank all the lectors reading my thesis, namely Daniel Bick, Ingo Burmeister, Christoph Göllnitz, Annika Hollnagel, Jan Lenkeit and Michael Wurm.

Last but not least, I thank my family, especially my mother, who always supported me. Without her this thesis would not have been possible. I also have to thank my father who inspired me already in my early childhood for physics and who died much too early, so he cannot see the fruits of his encouragement (or can he?).