Investigation of the One-Particle Approximation in the ANTARES simulation package KM3

Bachelorarbeit

vorgelegt von
Mona Dentler

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

1. Gutachter: Prof. Dr. Gisela Anton
2. Gutachter: Dr. Thomas Eberl

Tag der Abgabe: 12. 10. 2012

Intoduction

Since their first postulation by W. Pauli in 1930, neutrinos have developed into one of the most fascinating and challenging objects of research. The fascination arises from their pivotal role in the research of some of the most important unsolved questions of modern physics, reaching from the subatomic scale of particle physics to the intergalactic scale of astrophysics.
Among these numerous fields of study is the search for the sources of the cosmic radiation. The extremely small interaction cross section of neutrinos makes them ideal messenger particles traveling thousands of light years to our planet, possibly pointing back to powerful cosmic particle accelerators, like Supernova Remnants or Gamma Ray Bursts. The flip side of this property is what makes the work with neutrinos so challenging: the difficulty in measuring those particles on Earth requiring a new dimension of detectors and calling for the human inventiveness.
In the range of remarkable technical achievements facing this challenge stands the ANTARES neutrino telescope at the bottom of the Mediterranean sea, a surrounding hostile to any precise electronic instruments. ANTARES is constructed for detecting the reaction products of neutrinos which come from the other side of the Earth, using the whole mass of the Earth as a filter against other particles and the water layer as a shield against disruptive radiation.
This work deals with the computational simulation of the detector response. The simulation of the expected events has become a very important part of the research improving both the fine-tuning of the instruments determining the response of the detector and identifying possible systematic errors. With instruments becoming more and more complex and expensive, a reliable simulation is nowadays a crucial part of the experiment.

The aim of this work is to improve the simulation of non-muonic events in the detector by refining a method known as the "one-particle approximation". The detection principle of ANTARES is based on the measurement of photons produced by charged particles due to a process called Cherenkov radiation. While the simulation of the detector-response to muonic events, done by a program called *KM3*, gives satisfactory results, the default program for simulating non-muonic events, *geasim*, has the disadvantage of neglecting scattered photons and being computational complex. A workaround is the one-particle approximation, which makes *KM3* capable of simulating non-muonic events by replacing hadronic particles, which *KM3* cannot handle, by electrons. A preliminary version of the one-particle approximation is already implemented. In this work, a improved version is developed.
This is achieved in several steps: In the first chapter, the ANTARES detector is introduced, the physical background of the events which are searched for and their signatures are discussed, and the detection principle is explained. In the second chapter, *KM3*, the standard program used for simulating these events, is introduced and its current limitations are discussed. In chapter three the concept of the one-particle approximation is introduced, its current state is described and the range of applicability is discussed. Since a very important part of this work is to create data in a wide energy range using the simulation program *Gen*, in chapter four a description of the production of the data and their evaluation is given. Subsequently it is shown how, based on these data, the current version of the one-particle approximation is improved by constructing a better-adapted approximating function considering also the energy dependence. After having found such an improved function, chapter five evaluates the new one-particle approximation and compares it to the current one as well as to *geasim*, showing that the new one-particle approximation is an improvement over both in the distances relevant to ANTARES.

Contents

1 ANTARES **5**
- 1.1 Overview . . . 5
- 1.2 Detector Design . . . 5
- 1.3 Event Classes . . . 6
- 1.4 Detection Principle . . . 9

2 The KM3 simulation package **13**
- 2.1 Overview . . . 13
- 2.2 Program Layout . . . 13
  - 2.2.1 Gen . . . 13
  - 2.2.2 Hit . . . 14
  - 2.2.3 KM3mc . . . 14
- 2.3 Limitation . . . 14

3 The One-Particle Approximation **16**
- 3.1 General Idea . . . 16
- 3.2 Physical Background . . . 16
  - 3.2.1 Electromagnetic Cascades . . . 16
  - 3.2.2 Hadronic Cascades . . . 17
  - 3.2.3 Comparison . . . 18

4 Refining the One-Particle Approximation **20**
- 4.1 Generation and Evaluation of Data . . . 20
- 4.2 Assessing the Range of Validity . . . 21
  - 4.2.1 Electromagnetic Showers . . . 21
  - 4.2.2 Hadronic Showers . . . 24
- 4.3 Fit of the Showers and Construction of a new One-Particle Approximation . . . 26

5 Evaluation and Conclusion **30**
- 5.1 Comparing different Methods of simulating Shower Events . . . 30
- 5.2 Conclusion . . . 33

A *GEANT* particle ID **35**

B *Gen* Parameters **37**

Chapter 1

ANTARES

1.1 Overview

The acronym *ANTARES* stands for **A**stronomy with a **N**eutrino **T**elescope and **A**byss environmental **RES**earch. This is a good summary of the main features of the giant instrument at the bottom of the mediterean sea some kilometers offshore from Toulon, France: ANTARES is looking for *neutrinos*, exclusively weak-interacting particles belonging to one of three generations of *Leptons*, and being assumed to be one of the elemental components of matter. The neutrinos of interest are highly-energetic neutrinos of astrophysical origin, justifying the "astronomy" in ANTARES. [1] But not only information on astronomical processes like particle acceleration mechanisms in energetic astrophysical objects such as active galactic nuclei and gamma-ray bursts is hoped for, but also particle physic's topics like neutrino oscillation and even possible traces of dark matter in the form of weakly interacting massive particles (WIMPs) are explored in the ANTARES data.
The detector's surface area is $0.1\,\mathrm{km}^2$, but in some sense it is a *planetar-sized* experiment, since the the whole mass of the earth is used as a *shield* or filter, being only transparent to neutrinos below a certain energy. Figure 1.5 shows the increase of the cross-sections of selected neutrino types with energy. This is the reason for the earth becoming more and more opaque for increasing neutrino energy. Neutrinos below that energy transverse the earth, and a small fraction of them reacts in or near the instrumented area, producing secondary particles which can be detected with ANTARES, those are mainly muons. For suppressing background, which is mostly made up by *atmospheric muons*, the experiment is located at the bottom of the deep sea. Nevertheless it has proved that atmospheric muons coming from above and crossing the 2.5 km of sea water make up an important part of the experiment's background. [1]

1.2 Detector Design

Figure 1.1 shows the setup of the ANTARES detector. The basic unit of the detector is the *optical module* (OM), consisting of a photomultiplier tube, various sensors and the associated electronics, all housed in a pressure-resistant glass sphere. The electronics include a custom-built digital electronic circuit which captures and stores waveforms, pulse heights and timing information, as well as the high-voltage power supply for the photomultiplier tubes and the network nodes for data transmission and *slow control*, a automatic system to monitor all relevant parameters which change slowly compared to the measured values. [1] The optical modules are grouped together into *storeys* (shown in the circle in the middle of figure 1.1) of three modules and interconnected via an *electro-mechanical cable*. Twenty-five of those storeys are lined up on each of twelve *strings* having a total height of about 350 m each. The strings are distant by about

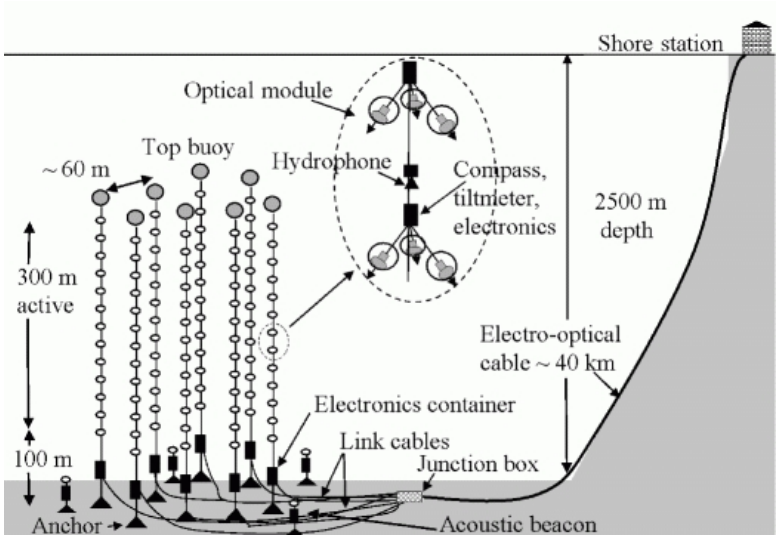


Figure 1.1: *Scheme of the setup of the ANTARES detector. [1]*

60 m, anchored on the sea floor and vertically supported by a buoy. This arrangement is not fixed but floats with a certain leeway with the sea current. Therefore a precise and continuous positioning system is required. This is archived by *acoustic triangulation*. In this system, every two minutes rangemeters placed on the string send an acoustic signal to a minimum of three transponders fixed to the sea bed. Each transponder replies with its characteristic frequency. A global fit of the measured acoustic paths gives the precise three-dimensional position of the rangemeters, provided that the positions of the transponders and the sound velocity in water are known.
The optical modules in a storey are arranged with the axis of the photomultiplier tubes 45° below the horizontal. This arrangement of the OMs detects light in the lower hemisphere with high efficiency, and has some acceptance for muon directions above the horizontal. In the lower hemisphere there is an overlap in angular acceptance between the OMs, permitting an event trigger based on coincidences from this overlap. ANTARES follows the concept "all data to shore", therefore the electronics of all the individual OMs are linked to a common junction box by electro-optical cables. A standard deep-sea telecommunication cable links the junction box with a shore station where the data are filtered and recorded. The trigger logic in the sea is designed to be as simple and flexible as possible. After three levels of triggering the readout rate is expected to be on the order of several kHz, and the corresponding data recording rate less than 100 events per second.[1]

1.3 Event Classes

The main interaction of neutrinos with matter is *inelastic scattering* on target nucleons, the cross section for which is generally several orders of magnitude larger than the cross section cross section for interaction with electrons. An exception to this is the *Glashow Resonance*, which will be discussed below.
A nucleus N and a neutrino ν can either interact by exchanging a charged $W^{\pm}$ boson, being called *Charged-Current Interaction* (CC), or by exchanging a neutral Z^0 boson, being called a

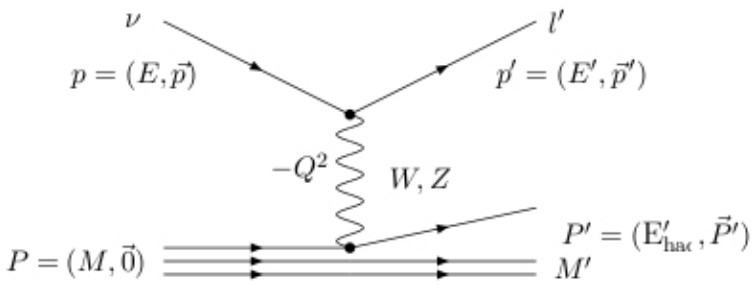


Figure 1.2: *inelastic scattering of a neutrino ν with momentum p on a nucleon N with momentum P. The interaction is either transmitted by exchanging a $W^{\pm}$ boson (charged current channel) producing a charged lepton ($e^{\pm}$, $\mu^{\pm}$, $\tau^{\pm}$) with momentum p', or by exchanging a neutral Z^0 boson (neutral current channel), producing a neutrino with momentum p'. In each case a hadronic shower is produced due to the large momentum of the recoiling quark. Following [2, p. 16].*

Neutral-Current Interaction (NC): [1] [2, p. 41-43]

$$\begin{aligned} &\text{CC:} && (\nu_l, \bar{\nu}_l) + N \rightarrow l^{\mp} + \text{hadronic shower} \\ &\text{or} \\ &\text{NC:} && (\nu_l, \bar{\nu}_l) + N \rightarrow (\nu_l, \bar{\nu}_l) + \text{hadronic shower} \end{aligned}$$

Because of the competing electro-weak coupling terms which must be taken into account for NC interactions, their cross section is only about one third of the CC cross section, which can be inspected in figure 1.5, showing the cross-section for both the NC and the CC channel.

The neutrino cannot be detected directly but only through the products of its interaction. So after having described qualitatively which interactions occur and which particles are produced, it is very important to look more quantitatively on the kinematics of the neutrino interaction to estimate the fraction of energy which gets transferred to the outputs of the interaction. This is done in the following part by introducing the kinematic variables Q, x and y which describe the momentum transferred, the inelasticity of the interaction, and the energy transferred between the neutrino and the nucleon.

The interaction between the neutrino and the nucleus is shown in figure 1.2: Q^2 is the negative squared four-momentum transfer $p - p'$ between the incoming and the outgoing lepton:

$$Q^2 := -(p - p')^2 = 2M(E - E') + M^2 - M'^2. \tag{1.1}$$

For highly-energetic neutrinos, the nucleus can be considered to be at rest and its rest frame can be identified with the laboratory system. A new variable, the *Bjorken variable* x is defined by the following equation:

$$x := \frac{Q^2}{2M(E - E')}. \tag{1.2}$$

In the case of a totally elastic interaction, $Q^2 = 2M(E - E')$, so that $x \equiv 1$. On the contrary, for inelastic interactions $0 < x < 1$ as can be seen from eq. 1.1. Therefore x can be considered as an indicator for the inelasticity of the interaction. As x approaches zero ($x \equiv 0$ would violate the momentum conservation and is therefore forbidden), all the energy transferred to the hadronic shower approaches the total energy of the neutrino, such that $E' = 0$. In this case M' would reach its maximum value $M' = \sqrt{M^2 + 2ME}$ and Q^2 would be zero, thus, the larger M', the smaller x. Summarizing the discussion above, x measures the elasticity of the

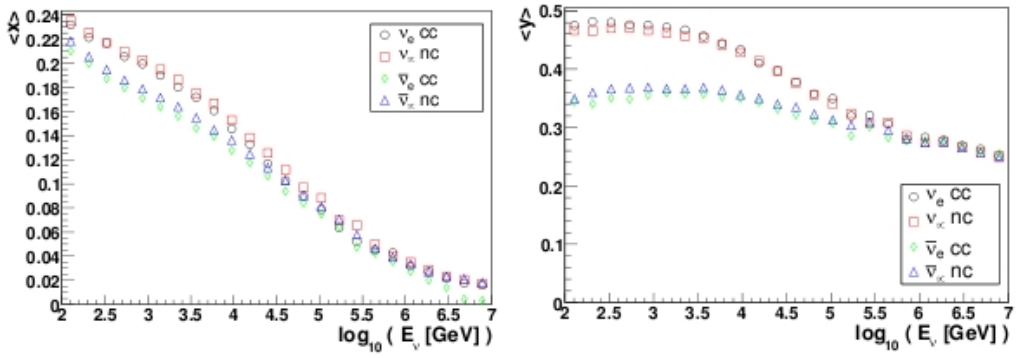


Figure 1.3: *Mean values of the Bjorken variables x and y as a function of the neutrino energy [2, p. 19].*

interaction, ranging from zero (totally elastic) to one (totally inelastic).
By contrast, the *Bjorken variable y* gives the relative energy transfer from the neutrino to the hadronic system. It is defined as

$$y := \frac{E - E'}{E}, \tag{1.3}$$

which implies that the larger y, the more energy is transferred on the nucleon. Applying this on the products of the CC- and NC-interaction of the neutrino it can be seen that small x-values mean that the reaction is inelastic, which is the condition for the creation of hadrons, while and large y- values mean that most of the energy of the neutrino is transferred on the hadronic shower and less on the outgoing lepton or neutrino and vice versa. [2, p. 16][4, p. 77-93]
Figure 1.3 shows the mean values of the Bjorken variables x and y as a function of the neutrino energy. The relation between the neutrino energy and the bjorken variables is not as obvious as one might suspect: it can be seen that $< x >$ decreases with energy, therefore the inelasticity of the interaction rises with energy. $< y >$ is depended from the neutrino type, but generally decreases with energy as well, and therefore the relative energy transfer on the nucleon decreases. It is important to keep this relation in mind during the following section describing the signatures of the produced particles.

If the energy transferred by the boson (measured by x and y) to the nucleon is sufficient, a quark gets knocked out of the nucleon, and for a brief moment the quarks fly apart as free particles, but when they reach a separation distance of around 10^{-15} m (the diameter of a hadron), their (strong) interaction is so great that new quark-antiquark pairs are produced mainly from gluons emitted by the quarks. Eventually those quarks recombine building a *jet* of mesons and baryons: Because of *confinement*, color-charged particles cannot be isolated singularly. The by-far biggest fraction of the jet is made up by pions. Those, together with the other *secondary particles* in the jet, can again interact with nuclei producing more hadrons; decay, if their lifetime is rather short; or, if they are charged, interact electromagnetically via bremsstrahlung or ionization. A cascade of particles, a *hadronic shower*, evolves just until all particles are stopped or absorbed by the surrounding medium. [3, p. 275-276]
For NC events, where the produced lepton is a neutrino, those hadronic showers are the only visible signature of a neutrino interaction. In case of CC interactions, the signature depends on the outgoing lepton, and thereby on the initial neutrino. In the case of a muon-neutrino, the outgoing lepton will be a muon. Those are the events detected most often, as the path-length of a muon in water exceeds the length of a shower by more than three orders of magnitude for energies above $\simeq 2\,\mathrm{TeV}$. This is the reason why those events, as discussed in sec. 1.4, can be

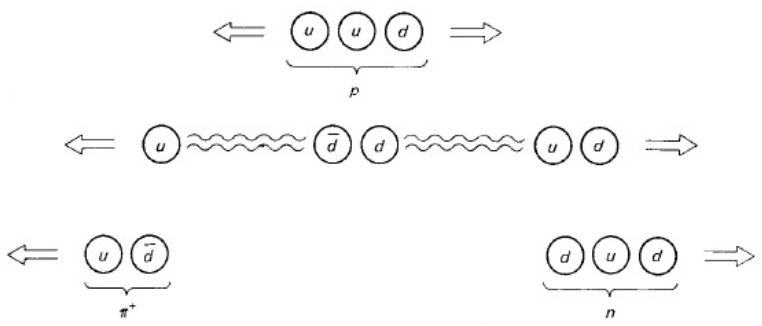


Figure 1.4: *Hadronization due to the confinement of quarks: an u quark gets knocked out of the proton, then a quark-antiquark pair is created, and finally, instead of an free quark, a pion and a neutron is created. [3, p. 72]*

registered, even if the interaction has taken place several kilometers outside the instrumented volume, because the muon travels a long distance before it decays due to its highly relativistic velocity.
In the case of the outgoing lepton being a tauon, a track will only be detected if the tauon has a total energy above $\simeq 1\,\mathrm{PeV}$. In that case, the expected signature will be a "double bang event", e.g. a hadronic shower, followed by the track of the tauon, and finally again a hadronic shower initialized by one of the decay modes of the tauon, or a "lollipop event", if the first interaction occurs outside the instrumented area.
In the case of the lepton being an electron, there will be no track, but an electromagnetic shower, because the electron will, due to its small mass, undergo bremsstrahlung and lose energy rapidly. [2, p. 41-51]
As mentioned above, the cross section for neutrino-nucleon interactions is much larger than the cross section for neutrino-electron interactions, except for $\bar{\nu}_e$ in the range of the *Glashow resonance* at $6.3\,\mathrm{PeV}$, clearly visible in figure 1.5. At this energy a W boson is produced resonantly by the CC interaction of the $\bar{\nu}_e$ with an electron of one of the target atoms. The following signatures are possible for that resonance:

$$\begin{aligned}
&\bar{\nu}_e + e^- \rightarrow \bar{\nu}_e + (e^- + \text{electromagnetic shower}) \\
&\bar{\nu}_e + e^- \rightarrow \bar{\nu_\mu} + \mu^- \\
&\bar{\nu}_e + e^- \rightarrow \bar{\nu_\tau} + (\tau^- \rightarrow \tau^- \text{-decay modes}) \\
&\bar{\nu}_e + e^- \rightarrow q + \bar{q}' \rightarrow \text{hadronic shower}
\end{aligned}$$

For the energy range of most of the events, the resonance channel constitutes only a small fraction of the total cross section. Furthermore, it will not be distinguishable from the neutrino-nucleon channels anyway. [2, p. 43]
So summing up this section, it can be said that in reality it is expected to observe hadronic cascades, electromagnetic cascades, muon tracks, and possibly some tauon tracks, which could very rarely be identified by their special shape; therefore they are treated as indistinguishable from muon tracks. The path-length of a muon track exceeds the length of the showers by far, though with increasing initial energy of the hadronic showers the shower length increases.

1.4 Detection Principle

As described in section 1.2, the ANTARES detector consists of arrays of photomultipliers, sensitive to photons in the visible frequency range. These are produced by charged energetic particles travelling through a medium, such as seawater, in a process called *Cherenkov radiation*,

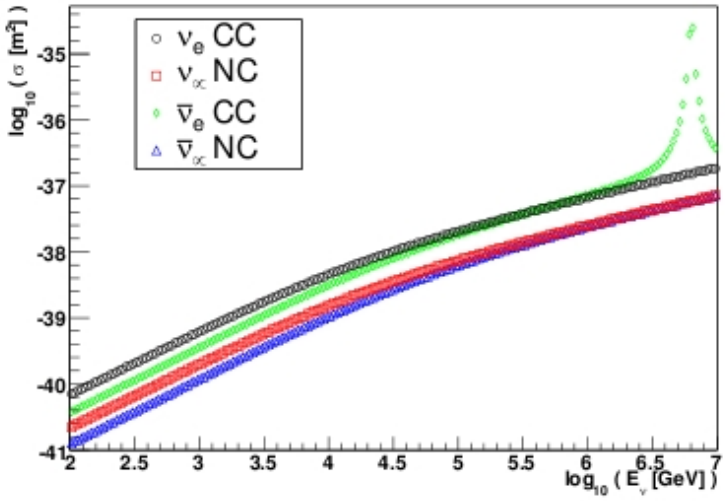


Figure 1.5: *Cross sections for selected interactions producing showers: Neutral current for the ν_μ and $\bar{\nu}_\mu$ and charged current for the ν_e and $\bar{\nu}_e$. For the latter, note the Glashow resonance at 6.3 PeV. [2, p. 18]*

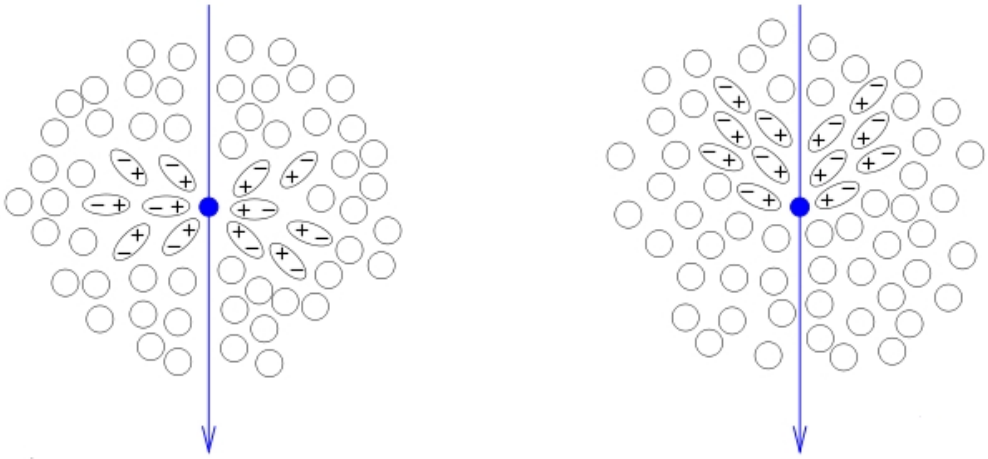


Figure 1.6: *Left: If the speed v of a charged particle travelling in a medium is less than the speed of light $\frac{c}{n}$ in this medium, the dipoles induced by polarization are distributed symmetrically around the particle. The overall dipol momentum is zero. Right: If the speed of the charged particle v is more than $\frac{c}{n}$, the dipole symmetrie is broken, and Cherenkov radiation is emitted. [2, p. 20]*

being based on the polarization of the surrounding medium along the track length of the energetic charged particle (see fig. 1.6 for illustration). As long as the particle's speed v is smaller than the speed of light in the medium, the polarization affects symmetrically all the atoms surrounding the particle, so that the overall dipole momentum is zero and no radiation will be emitted. But if the particle's speed exceeds the speed of light in the medium, it "overtakes" the effect of its force on the surrounding atoms, and the symmetry will be broken and dipole radiation will be emitted, forming a so-called *Cherenkov-cone*, evolving along the particle track.

The opening angle θ of the cone is determined by the velocity of the particle on one hand, and the refraction index n on the other, by the following formula: [17, p. 241]

$$\cos\theta = \frac{1}{n\beta}, \tag{1.4}$$

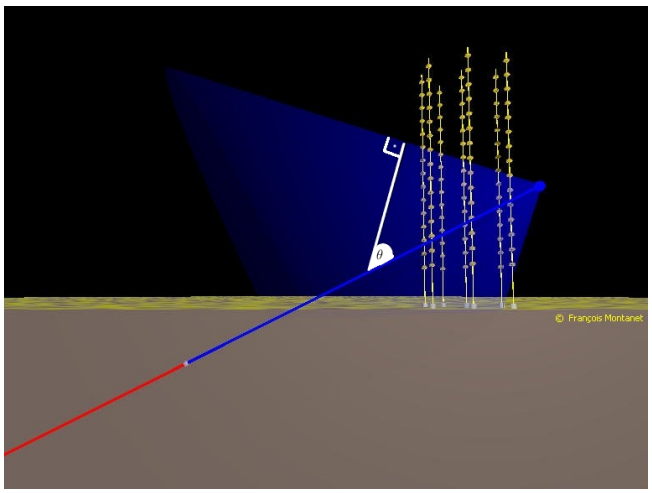


Figure 1.7: *Detection priciple: a neutrino (red) reacts near the detector, producing a muon (blue). The muon produces Cherenkov radiation, which is detected by the optical modules. [1]*

with $\beta = \frac{v}{c}$. In the energy range of interest for ANTARES (E > 10 GeV), particles will generally be ultra-relativistic with $\beta = 1$. The refractive index of sea water is $n = 1.35$ for a wavelength of 450 nm (400 - 500 nm is the wavelength at which the photomultipliers are most effective), therefore the Cherenkov light is emitted under 42° for this wavelength. This light cone will be registered by the OMs. The detector is optimized for registering *muon events*, e.g. for registering muons, which are produced by a muon- (anti)neutrino interacting with a nucleon by exchanging charged current. The number of photons N produced along a flight path dx in a wave length bin $d\lambda$ is: [17, p. 242]

$$\frac{d^2N}{d\lambda dx} = \frac{2\pi\alpha Q^2 \sin^2\theta}{\lambda^2}, \tag{1.5}$$

with $\alpha = \frac{2\pi e^2}{hc}$, e the elemental charge, Q the charge of the particle and θ the Cherenkov angle. For $\beta = 1$ the Cherenkov light yield is independent of the energy of the charged particle. Since at high energies the track length of a muon is normally much longer than the diameter of the detector, only a small part of the energy will be deposited inside the detector. This means the light output of a single particle does not allow its energy to be measured. On the other hand, due to the simple geometric pattern or the Cherenkov cone, it is possible to reconstruct precisely the direction of the muon, and therefrom the neutrino, from the measurement of only few hits at different spatial points [1]. Figure 1.7 illustrates the detection principle in this case: an upcoming neutrino reacts near the detector; the created muon produces Cherenkov radiation, which is detected by the optical modules. From the arrival time of the photons it is possible to reconstruct the direction of the muon. [1]

However, when hadronic showers are produced which actually occurs more often and are sometimes even the only signature for an neutrino event, the total light yield of the shower will be proportional to the total track length of charged particles in the shower and therefore to its initial energy. This allows some calorimetric measurements if the neutrino vertex is inside the active detector volume. By contrast, the shower direction is harder to reconstruct due to the many secondary particles which do not fly exactly in the shower direction, so that the Cherenkov cone is kind of "washed out", with the angular distribution of the photons emitted having a maximum, but no sharp peak at 42°, as can be seen in figure 1.8. Unfortunately, the shower length for a hadronic cascade is only about 10 m, which makes them appear as virtually point-like events in the detector. Obviously those events can only be detected if they occur inside or within $\simeq$ 100 m of the detector, while a muon can be registered even if it is created

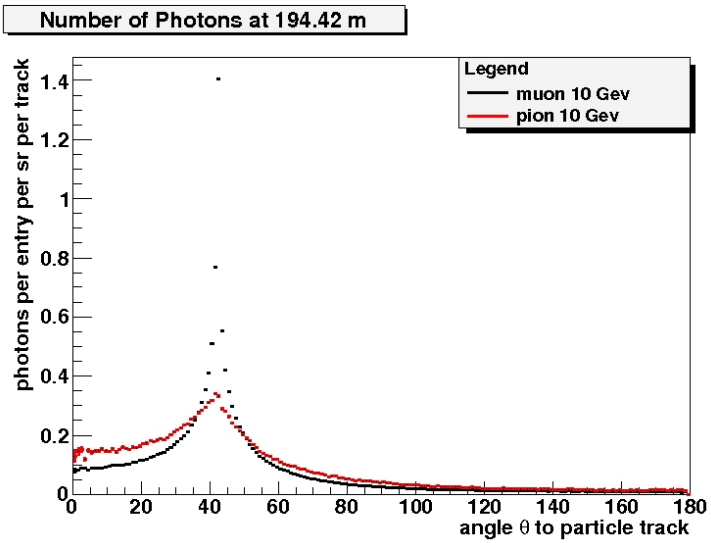


Figure 1.8: *Angular distribution of photons from a muon track (black) and a hadronic shower induced by a pion (red) of the same energy. Note the different shapes of the Cherenkov peak at 42°. (The data are produced with the simulation program gen, refer to section 4.1 for details)*

kilometers away from the detector. This makes the detector much more sensitive to muons and hence ν_μ events than shower events.
In this chapter the physical background of the detection of neutrino events within ANTARES has been discussed. For this work it is also important to understand how these events are modeled in the ANTARES simulation packages. Therefore an introduction of these simulation packages is given in the following chapter, building on the physical background explained in the sections above.

Chapter 2

The KM3 simulation package

2.1 Overview

In order to evaluate and fine-tune analysis techniques, synthetic data which is as realistic as possible is attempted to be reproduced by *Monte Carlo* computer simulations. An important part of simulating the detector response involves calculating the number of Cherenkov photons emitted from relativistic particles in seawater, their scattering and absorption in seawater, and how these photons are detected by the Antares optical modules. The standard program package used for those simulations is called *KM3*. It is designed to handle relativistic muons and their secondary particles. An important feature of *KM3* is that it also incorporates the scattering of light. *KM3* is actually a suite of three different `FORTRAN-90` programs designed to fulfill independent tasks [5]. A description of the individual programs is given in the following sections.

2.2 Program Layout

2.2.1 Gen

Gen [6, p. 5-9][5, p. 12-16] simulates the generation of Cherenkov photons by a particle in a given medium, including those produced by any secondary particles. Furthermore, it tracks the Cherenkov photons through space with wavelength-dependent absorption and scattering taken into account, recording the position, direction and arrival time of photons at spherical *shells*, concentric sphere surfaces of given radii around the origin.
Gen requires, amongst other things, the following input data: particle type to be simulated, number of events to be processed, the particle momentum, the medium type (e.g. seawater in the case of ANTARES) and the physics to be regarded in the simulation (pair production, δ-ray, bremsstrahlung etc). Another input has to be passed by the user: a file which specifies the shell radii and the *scale factors* applied on each shell. In order to keep the size of the output file acceptable, the photons on each shell are only stored with a certain probability specified by the user. Thus, those factors should be smaller the nearer the shell is located to the origin and closer to one the further away the shell is located, in order to keep the number photons recorded on each shell approximately the same.
The initial particle can be any particle known by *GEANT* with any energy. *GEANT* [7] ("GEometry ANd Tracking") has been developed at CERN and originally designed for simulating high-energy physics experiments, but today it has found applications in many other fields. It is a program package simulating the passage of particles through matter by using Monte-Carlo methods.
A complete *GEANT* simulation [5, p. 13] is used at this step. The *GEANT* code gets initialized

and the particle is propagated along the positive z-axis through the medium. In the case of muons, *Gen* will stop tracking the muon after 1 m. All secondary particles are also processed. The number of Cherenkov photons is computed according to a *Poisson distribution* with mean given by equation 1.4. For each photon, the position, the direction and the energy are given. Each photon is propagated through the volume, suffering scattering and absorption, and finally, if the photon crosses the boundaries of a shell, the mentioned parameters are recorded with a certain probability, as explained above.
The output of the program is a file containing specific parameters of the produced photons on the eighteen shells. These data generated by *Gen* are subsequently used by the program *Hit* (see following section 2.2.2)

2.2.2 Hit

Instead of the actual output of *Gen*, which contains a complete record of the photon properties on each shell, the information of interest includes the probability and time distribution of direct and scattered photons arriving at the optical modules (OM), and their response. The final photon tables are computed by dividing each shell into twenty *bands* defined by intervalls of $\cos\theta$, θ measuring the angle to the particle track. In each band a generic OM is placed, and by looping over different OM orientations given by θ_{OM}, Φ_{OM} , weighting each hit for each OM direction by the fraction of the *effective area* of the PMT and the *geometrical area* of the band, an inverse cumulative probability distribution $N_{det} = N_{det}(shell, band, \theta_{OM}, \Phi_{OM})$ is calculated. Thereafter, the time distribution t_{dist} for each θ_{OM}, Φ_{OM} bin is computed, providing the time distribution $t_{dist} = t_{dist}(shell, band, \theta_{OM}, \Phi_{OM})$. The output is therefore a two four-dimensional histogram of the probability respectively a five-dimensional histogram of the arrival time distribution, folding the shell bands and the different OM orientations with the OM efficiency as a function of wavelength.[5, p. 16-19][6, p. 10-12]

2.2.3 KM3mc

The *KM3mc* program takes the photon tables generated by *Gen* and *Hit* and uses them to simulate the detector response to the passage of high-energy muons. Both the *Gen* and *Hit* programs are usually run only once to generate the relevant tables of OM hit probabilities. These tables are then stored to disc for subsequent use by *KM3*, resulting in a program which includes the effects of scattering and yet is fast. [6, p. 12]
After reading the user's inputs and the photon tables created by *Gen* and *Hit* for muons and electrons, the *MUSIC* package is initialized. *MUSIC*, MUon SIulation Code, is a `FORTRAN-90` program which propagates muons to some distance in rock or water and returns the final muon parameters [5, p. 10]. The processes taken into account include bremsstrahlung, pair production, inelastic scattering, ionization and multiple Coulomb scattering. A muon is processed if its distance to the detector is smaller than $\simeq 150\,\mathrm{m}$. If the energy lost by a muon in a *segment* (e.g. the length of the current tracking step) is well above the average energy loss by ionization (typically the threshold is set $\simeq 0.3\,\mathrm{GeV/m}$), an *independent* electromagnetic shower is assumed to occur at a random location along the length of the segment. This case is treated separately as an *electron*. Finally the programs computes the direct and scattered hits produced by the muon and electrons using the photon tables. [6, p. 12-14][5, p. 19-23]

2.3 Limitation

KM3mc is designed to only calculate the Cherenkov photons emitted by an energetic muon and the secondary electromagnetic cascades it generates. But as discussed in section 1.3,

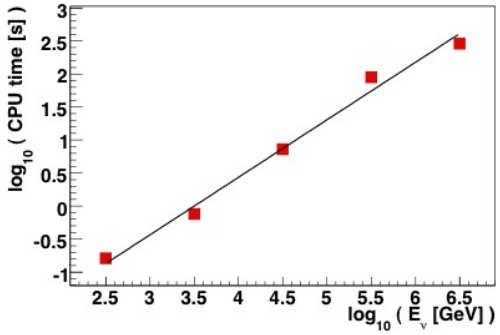


Figure 2.1: *Average CPU time for processing one event as a function of the neutrino energy. [2, p. 139]*

muon tracks are the signature of only one possible neutrino interaction, the CC-channel of the muon-neutrino interacting with a nucleus. Since the cross-section for the weakly-interacting neutrinos is notoriously small, it is important to take also rarer, but still important interactions into account, such as a neutrino interacting directly inside ANTARES, to increase the event rate of the detector. Furthermore, cascade events have the advantage that they allow much better calorimetric measurements, since all the energy transfered to the nucleon will be deposited by the shower. However, *KM3* is not able to simulate those events by default. Those events are handled by a program called *geasim*. This program takes into account all relevant physics processes that can occur during the passage of a particle through the medium [8]. The angular distribution of electromagnetic showers is parametrized, since Monte Carlo studies have shown that the distribution is independent from the energy of the shower. However, for hadronic showers, the distribution is energy-dependent and subject to large fluctuations (compare also section 3.2.2), therefore every single particle in the shower is tracked down to Cherenkov level. This is very time-consuming, and the CPU time rises with the shower energy, as can be seen in figure 2.1. It can be seen from there that it takes $\simeq 0.2\,\mathrm{s}$ to calculate one event for a neutrino of 315 GeV but over two and half minutes for one single event for a neutrino of $\simeq 315\,\mathrm{PeV}$. Another disadvantage of *geasim* is that it does not, unlike *KM3mc*, account for scattered photons. Therefore one possible alternative would be to get *KM3mc* to be able to handle these events. This idea will be pursued in the following chapters.

Chapter 3

The One-Particle Approximation

3.1 General Idea

As explained in section 2.3, the simulation package *KM3* is designed to only calculate the Cherenkov photons emitted from an energetic muon and the secondary electromagnetic cascades it generates; all the other possible events being potentially seen in the real detector are not simulated in the *KM3mc* simulation package. In order to make *KM3mc* capable of handling events other than muon events, the wide range of secondary particles which get produced in neutrino interactions must be made to "look like" the electrons and muons which *KM3mc* can handle. This means approximating all the hadronic and electromagnetic cascades produced by the secondary particles with a single electron cascade at some energy E_e, such that the total number of Cherenkov photons produced remains the same. In other words, any secondary particle p_i of an initial energy E_i produced in a neutrino interaction is replaced by an electron of the initial energy $E_e = w_i \cdot E_i$, where w_i is a weight function which is specific to the produced particle p_i. This is called the "one particle approximation". [9][10]
A preliminary version of the one-particle approximation has already been implemented in *KM3*. The current state is that if the secondary particle is an electron, positron, gamma, kaon or neutral pion, it gets replaced by an electron of the same energy; if the secondary particle is a charged pion $\pi^{\pm}$, it gets replaced by an electron of only 20% of the its energy; all other particles are not regarded [10]. The aim of this work will be to improve that approximation by modeling the equivalent electromagnetic energy more accurately. In order to do that, the photon yield of both electromagnetic and hadronic showers will be studied and a parametrization of each will be given in chapter 4, but first the physical background underlying the one-particle approximation will be discussed. This is a very important step, because the quality of the approximation depends significantly on the shapes of the shower types: because the one-particle approximation considers only the difference in energy between the showers, the photon yield at any distance can only be calculated exactly if the shower shapes were the same. This assumption does not correspond to the physical facts, therefore the physical background regarding the two different shower types will be described followed by an discussion of the applicability of the approximation.

3.2 Physical Background

3.2.1 Electromagnetic Cascades

The main interactions of an energetic electron are ionization and bremsstrahlung [2, p. 44-46]. In this context, the *critical energy* is defined as the energy of an electron, at which the energy loss due to ionization on one hand and bremsstrahlung on the other are equal. Above that

critical energy, an electron will induce an electromagnetic shower, because a photon produces an electron and a positron, a process called "pair-production", if its energy is higher than the rest masses of the electron and the positron. The resulting positron and electron will again emit photons due to bremsstrahlung, and so a shower of particles will evolve until the energy of the electrons and positrons falls below the critical energy and the photon's energy below the energy needed for pair-production. In a simplified model, the relevant interactions happen after each radiation length given as [11, p. 63]

$$\chi_0 = \frac{716.4 \cdot A}{Z(Z+1)\ln\frac{287}{\sqrt{Z}}}\ \mathrm{g \cdot cm^{-2}}, \tag{3.1}$$

with A the atomic mass and Z the atomic charge of the atoms in the medium. It is 35 cm in salt water. In this simplified model, the energy is distributed equally between the generated particles. Therefore after a distance $s = t \cdot x_0$ the number of particles N is $N = 2^t$, the energy per particle is $E = \frac{E_0}{N} = \frac{E_0}{2^t}$. The maximum number N_{max} of particles being produced in the cascade is $N_{max} = \frac{E_0}{E_{crit.}}$, the shower length is approximately $x_0 \cdot t_{max} = x_0 \cdot \frac{ln(E_0/E_{crit.})}{ln(2)}$.
In reality, neither the energy ratio given to each particle is always the same nor is there a fixed distance after which the interactions occur, but on an average over the many particles in the shower induced by an electron of relatively high energy, the model gives an useful approximation. The *longitudinal profile* of the shower, that is the energy loss per path length, is given by a parametrization called the *Longo-formula* [11, p. 71] [17, p. 239-240]:

$$\frac{dE}{dt} = E_0 b \frac{(bt)^{a-1}e^{-bt}}{\Gamma(a)}, \tag{3.2}$$

where t is the penetration depth in units of radiation lengths, $t = z/x_0$. The shower maximum according to this formula is at

$$t_{max} = \frac{(a-1)}{b}$$

with a, b depending on E_0, Z. Knowing the parameters a, b, the equation can be integrated numerically, giving the distance at which a certain amount of energy has been deposited. A numerical computation of the longitudinal shower profile of both electromagnetic and hadronic showers of varying energies is given in figure 3.2 in section 3.2.3.

3.2.2 Hadronic Cascades

The description of a hadronic shower is much more complex than that of an electromagnetic shower since it does not only consist of two sorts of particle, like the electromagnetic shower, but of many types of different fractions, depending on the initial energy. A hadronic shower is produced by a high-energy hadron such as a nucleon, pion, or atomic nucleus. As hadrons are color-charged, they mainly interact via the strong force, thus, the exchange of *gluons*. Due to confinement, the inelastic scattering of hadrons results in a jet of new hadrons, an effect called hadronization [3, p. 275-276]. This effect is also discussed in section 1.3. The produced hadrons can then again interact with each other and the nucleons of the surrounding medium until they decay or are stopped or absorbed, resulting in a cascade of particles. If the lifetime of the particles is rather short compared to the radiation length, they decay before they interact. The most prominent example is the π^0 which decays into two γ-photons: $\pi^0 \to \gamma\gamma$; decays are the cause of most of the electromagnetic subshowers. Some hadrons are also electrically charged – these particles will also interact electromagnetically via bremsstrahlung or ionization. This gives rise to electromagnetic subshowers.

Most of the particles produced in the shower will be pions, but other hadrons like kaons,

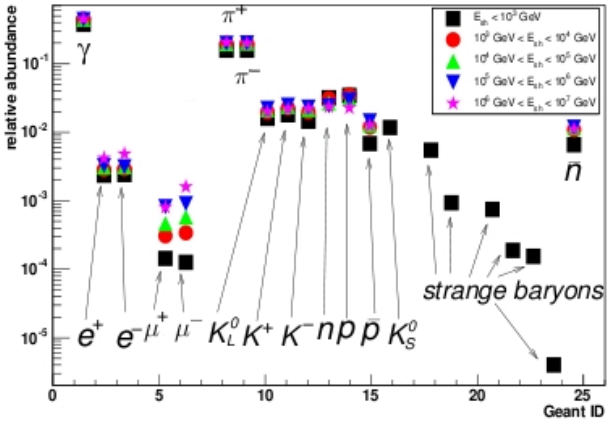


Figure 3.1: *Composition of hadronic shower: relative abundance of secondary particles in hadronic showers with different primary energies. Note that only relatively long-lived particles appear. [2, p. 48]*

protons or neutrons may also occur. Figure 3.1 shows the relative abundance of secondary particles in hadronic showers of different energies. The by-far biggest fraction is made up by γ, π^+ and π. For the primary energy being between 10 PeV and 100 PeV, the fraction of $\pi^\pm$ of the solid particles, e.g. all particles except the γ, is $\simeq 60\%$. The next biggest fraction is made up by neutrons and protons, lying between 2-3%. The production of muons contributes significantly to the fluctuations of the hadronic showers, since the muons usually leave the shower depositing their energy along long tracks.
For a shower initialized by a pion, it can be stated that if the initial energy does not exceed 139,57 MeV (the rest mass of the pion) by much, the pion has so little kinetic energy, that it almost instantly decays, without a significant cascade. On the other side of the energy range, the shower approximates more and more an electromagnetic shower, because the number of π^0 increases significantly due to the interaction of the charged pions. Since the π^0 have a very short lifetime, they decay almost instantly into two photons, which will start an electromagnetic cascade. The percentage of electrons and positrons of the total track length of a hadronic shower exceeds 90% for a hadronic shower of 1 TeV, which means that above that energy, the largest part of the photon yield will be generated by electromagnetic sub-showers. [12]

3.2.3 Comparison

In the two previous subsections, electromagnetic and hadronic showers have been characterized. Although electromagnetic showers can be regarded as a subunit of hadronic showers and appear from a certain energy onwards always as a component of the hadronic showers, there are significant differences between the two types of showers. These comprise different interaction types - electromagnetic and strong - as well as the number of occurring particle types. Consequently, they differ both in the angular and the longitudinal shower profile. Furthermore, the hadronic showers show strong fluctuations, which exceed 30% with respect to the longitudinal profile in the region of some hundred GeV, as can be seen in figure 3.1. Figure 3.2, showing the parametrization of the longitudinal shower profile of electromagnetic and hadronic showers in units of the electromagnetic radiation length in saltwater, $x_0 = 35\,\text{cm}$, illustrates these differences. Below 100 TeV, the electromagnetic shower has a shorter longitudinal extension than the hadronic shower. Above 100 TeV, the shower maximum of the electromagnetic shower still is further from the origin, but the overall length becomes approximately equivalent for both shower types.

Due to the increasing electromagnetic component, as discussed in section 3.2.2, the fluctuations in the hadronic shower shape are less significant at higher energies. They add up to only

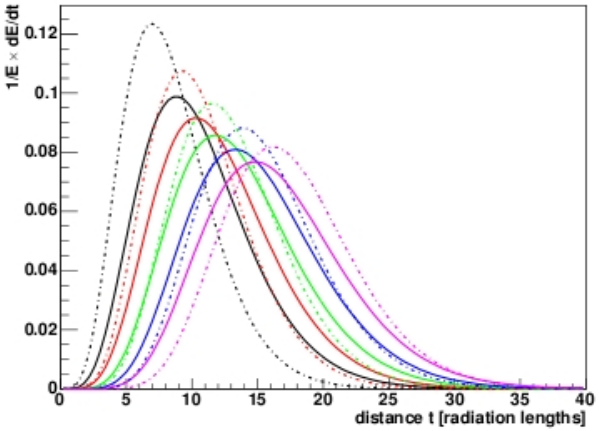


Figure 3.2: *Parametrisation of the longitudinal shower profiles for pion-induced hadronic showers (solid lines) and electromagnetic showers (dashed lines) at energies of 100 GeV (black), 1 TeV (red), 10 TeV (green), 100 TeV (blue) and 1 PeV (magenta). [2, p. 49]*

around 10% with respect to the longitudinal profile at some 10 TeV [2, p. 48]. This shows that the one-particle approximation is justifiable at least for high initial energies. Another aspect is the angular distribution of the photons emitted by electromagnetic and hadronic showers of different energies. Due to the larger number of particles and the higher fluctuations, it is expected that the angular distribution of hadronic showers is broader then that of electromagnetic showers. Again this effect is expected to decrease as the fraction of the electromagnetic subshowers increases.
In summary it can be stated that, physically, the one-particle approximation seems absolutely applicable for photons produced by showers of higher initial energies and in some distance from the origin. It remains to be determined however at which distances and energies the approximation can be reasonably applied. This assessing the range of validity is therefore the first goal of the following chapter.

Chapter 4

Refining the One-Particle Approximation

4.1 Generation and Evaluation of Data

In the previous section the need for the one-particle approximation has been shown and the general compatibility of the approximation with the physical conditions has been discussed, though no explicit range of validity has been specified. By contrast the aim of the following sections is to treat the approximation quantitatively. This will be carried out in two steps: the first step is the quantitative assessment of the range of validity of the one-particle approximation; the second step then is the parametrization of the photon yield of both the electromagnetic and the hadronic showers as a function of the initial energy in this established range.
Finally, a new one-particle approximation taking the energy dependence of the weight-function w_i (refer to the introduction of the one-particle approximation in 3.1 for details) into account will be constructed based on these parameterizations.

Undeniably the basis of these quantitative investigations are data of electromagnetic and hadronic showers in a sufficient large energy range. Because these data are underlying all the further sections, a short description of how the data are produced and evaluated is given at this point.
As described in chapter 2.2, the simulation program *Gen* is used to simulate particle showers in ANTARES and to calculate the resulting yield of photons. An overview of the basic structure of *Gen* is given in section 2.2.1. Using *Gen*, the parameters initial particle type, total energy of the particle, the number of events, e.g. the number of times the event is to be simulated, as well as the physical model describing the interactions of the particles are passed to the program. This is done by a *tcsh-scipt*, which calls *Gen* and passes the parameters. The particles specified were of course electrons and $\pi^{\pm}$, though in the highest energy regime ($> 1\,\mathrm{TeV}$) only π^{+} were used, because at high energies no significant difference between the photon yield induced by π^{+} and π^{-} is expected (as indicated by preliminary investigations). In total, the number of events simulated for this work was nearly 20000.
The energy range was chosen from $10\,\mathrm{MeV}$ up to $10\,\mathrm{TeV}$. The lower end of the energy range was chosen because below this energy, the kinetic energy of the pions is too small to allow them to travel even a single interaction length. Therefore they decay into a muon before they can interact with matter, and no shower evolves. The higher end of the energy range was chosen because above that energy, no change in the trend of the photon yield was expected considering a reasonable extrapolation of the data based on the discussion given in section 3.2.3.

The number of photons calculated rises at a rough estimate linearly with the initial energy of the particle (a detailed discussion of that relation is given in the following sections) and hence increases the size of the output. Therefore the number of events was chosen depending

on the energy: on one hand this was due to the fact that, because of the format, the size of the output file is restricted; on the other hand are the statistics for high-energy showers much better. This is due to the larger amount of particles produced: random fluctuations are less significant and average out on the total shower length, which means that a smaller number of events is needed. In the low energy regime there was chosen a higher number of events (100), which was successively decreased to three events for the highest energies (above 1 TeV).
The physical model used to calculate scattering of photons is the *partic-model*, which takes into account the effect of scattering on water molecules (*Rayleigh component*) as well as on particulate matter (*Mie-component*). *Gen* then propagates the initial particle along the z-axis, calculates the secondary particles and the created Cherenkov-photons on eighteen shells. For the values of the other parameters refer to appendix B.
The output of *Gen* is, as described in section 2.2.1, a file containing specific parameters of the produced photons on the eighteen shells. Additionally, the source code of *Gen* was changed, such that the particle code and the position of each particle produced in the simulated showers were written to file. 2.2.2) [6, p. 5-9][5, p. 12-16]

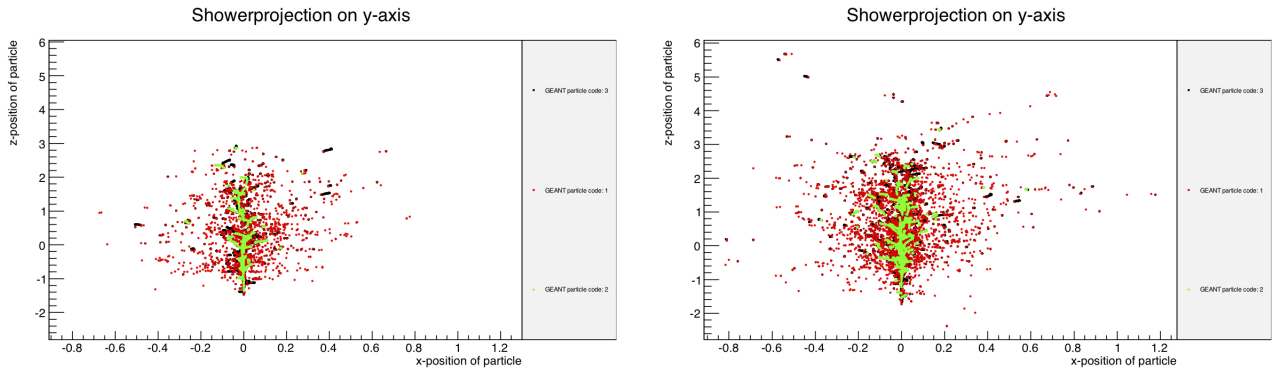
For the evaluation of the data, the output files were converted to *.root*-files, since in this work, *ROOT* was used to evaluate the data simulated by *Gen*. ROOT [13] is a framework for data processing, developed by CERN, which is designed for particle physics data analysis.
The first step when evaluating the data was to calculate the error on the number of photons counted. This was done by calculating the mean by averaging over the the number of photons created per event. The error was then estimated as the root-mean-square deviation from this mean.
The most important application of ROOT for this work were the fitting algorithms. The ROOT fitting algorithms are included in the *MINUIT* class [14]. By default a chi-square function is used for fitting. It calculates the chi-square between the user fitting function and the data for given values of the parameters. It is the task of *MINUIT* to find those values of the parameters which give the lowest value of chi-square. However, for functions which are linear in parameters, like polynomials, a linear fitting algorithm is applied by default.

After having explained the generation of data and having introduced the evaluation methods used, the next sections deal with the analysis of the so-produced data.

4.2 Assessing the Range of Validity

4.2.1 Electromagnetic Showers

Following the model described in section 3.2.1, an electron of half of the initial energy should lead to a cascade of approximately half of the (charged) particles, and therefore to half of the (Cherenkov-) photons. In figure 4.1 there is shown a plot of two electromagnetic showers. The initial energy of the electron in (a) is half the size of the initial energy in (b). The number of electrons and positrons and can be estimated by eye to be about twice as big in (b) then in (a), illustrating the argument above.



(a) Initial energy: 5 GeV (b) Initial energy: 10 GeV

Figure 4.1: *Plot of electromagnetic showers induced by electrons of different energies. Red: γ, black: electrons, green: positrons*

In figure 4.2 the photon yield in dependence from the initial electron's energy for shell eighteen ($\hat{=}$ 194.4 m) is shown. Obviously the data follow a linear relationship, consistent with the argument above.

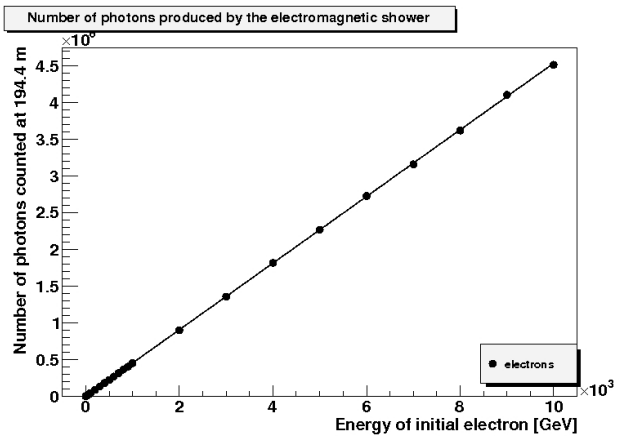
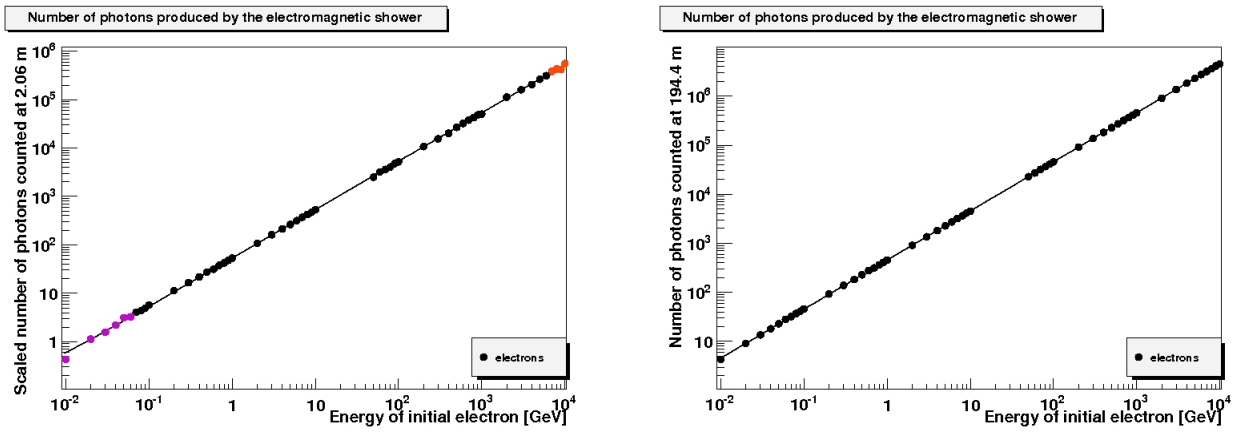


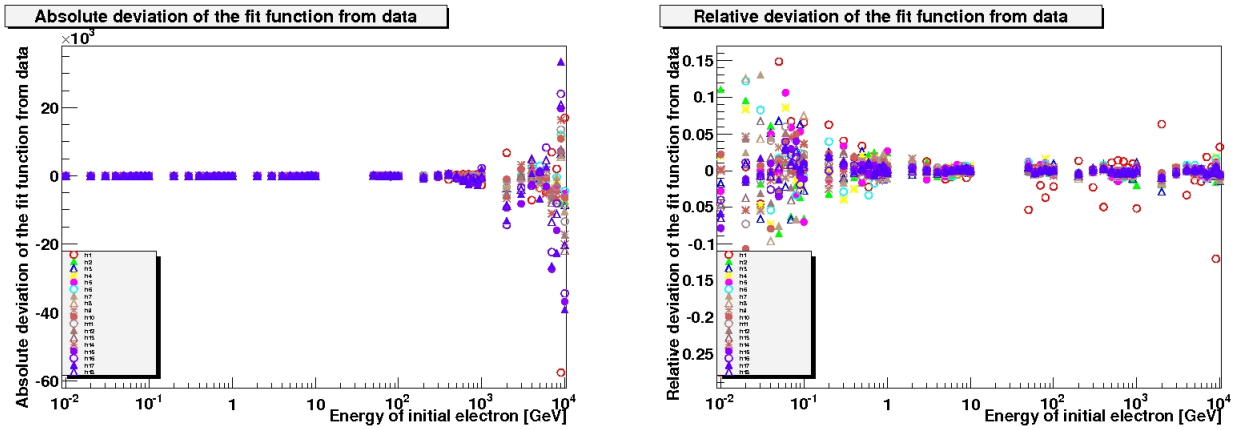
Figure 4.2: *Photon yield of an electromagnetic cascade induced by electrons of different energies at shell eighteen.*

In figure 4.3 (b) the same data are presented, but this time in log-scale, and for comparison the number of photons at shell one ($\hat{=}$ 2.06 m) is shown in 4.3 (a). Note that for shell one, the number of photons is scaled by the factor $3.69 \cdot 10^{-4}$ (refer section 2.2.1 for details). The black line in each plot shows a linear fit of the data. The deviation of the data points in the low energy regime (colored purple in figure 4.3 (a)) at shell one (refer figure 4.3 (a)) is probably due to the poor statistics for low energy showers; for low initial energies, only a few photons are produced. Due to the small scaling factor, only a very small fraction of these photons gets recorded, for an initial energy of 10 MeV for example, on average 0.43 photons on shell one are recorded. The deviations in the high-energy regime (colored orange in figure 4.3 (a)) are due to the shower length, which is longer than the inner shells' radii, and which therefore sees only a fraction of the produced photons. Figure 3.2, which has been discussed in section 3.2.1, shows the shift of the shower maximum and the extension of the shower length. It is obvious that if the shower length exceeds the shell radius significantly, the number of photons will be distorted. Table 4.1 shows the shower length, given as the distance in which 95% of the initial energy has been deposited and the radius of the three innermost shells.



(a) innermost shell (b) outermost shell

Figure 4.3: *Photon yield of an electromagnetic cascade induced by electrons of different energies at different shells: (a) innermost shell, colored: deviations in the low- and high-energy regime. (b) outermost shell; note that the number of photons is scaled as explained in section 2.2.1.*



(a) Absolute deviation of the number of photons from the fit-function (b) Relative deviation of the number of photons from the fit-function

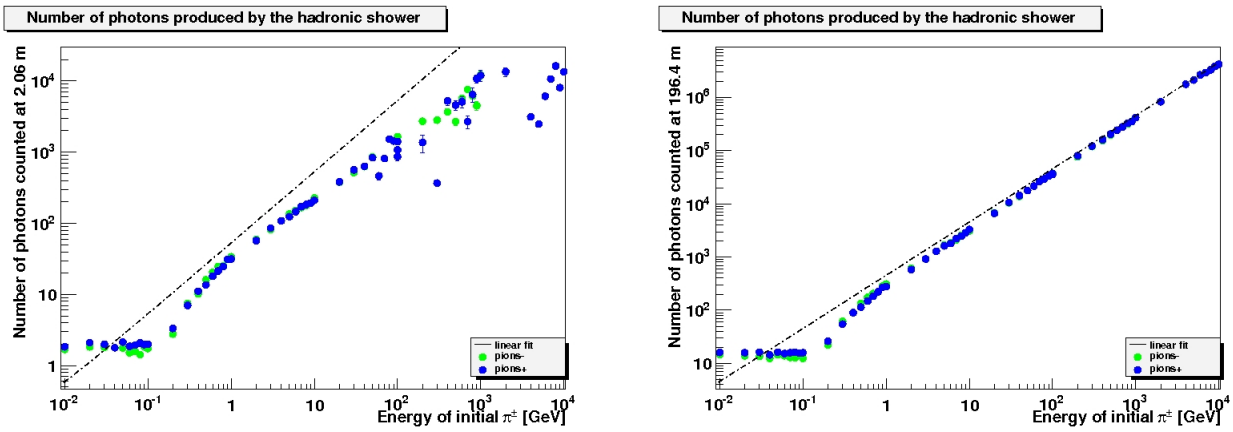
Figure 4.4: *Absolute (a) and relative (b) deviation of the number of photons from the linear fit on the eighteen shells*

initial energy E_0	shower length
1 GeV	1.8 m
10 GeV	3.6 m
100 GeV	5.4 m

shell number	radius
1	0.8 m
2	2 m
3	4 m

Table 4.1: *shower length dependent on the initial energy and for comparison the radii of the inner shells [5, p. 13].*

By looking at figure 4.3 it seems that the discussed factors do not bias the recorded number of photons considerably. Therefore the absolute deviation $(N_{\text{scaled}}(E_e) - \text{Fit}(N_{\text{scaled}}(E_e)))$ and relative deviation $([N_{\text{scaled}}(E_e) - \text{Fit}(N_{\text{scaled}}(E_e))]\,/\text{Fit}(N_{\text{scaled}}(E_e)))$ of the logarithmized data from the fit-function is shown in figure 4.4. This figure confirms the points discussed above: Shell one (red circles) has the highest relative as well as absolute deviation from the fit. The most significant relative deviations occur in the low energy regime below $\simeq 1\,\text{GeV}$ (due to the small data set) and the high-energy regime above $\simeq 1000\,\text{GeV}$ (shower size exceeding the shells' diameter). It is the aim to find a fit-function, which represents most of the shells as accurate



(a) innermost shell (b) outermost shell

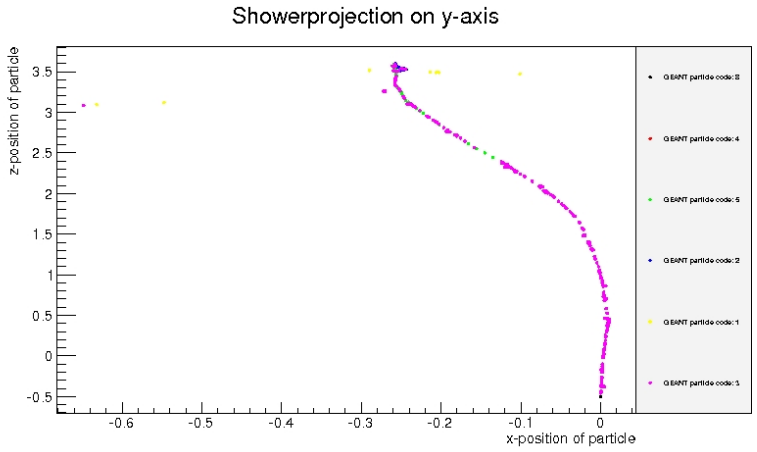
Figure 4.5: *Photon yield of a hadronic cascade induced by π^+ and π^- of different energies at different shells: (a) innermost shell, note that the number of photons is scaled as explained in section 2.2.1. (b) outermost shell*

as possible. Therefore it was decided based on these observations to neglect the two innermost shells for finding the fit-function.
In this case (with the innermost shells left out) the electromagnetic shower can be considered a point-like event, which produces the more photons the higher the initial energy is. Since the absorption of the photons is not depended from the number of photons, the function of the photon yield on each shell should be linear with energy, only differing in the slope due to the different number of photons reaching at the different distances and being counted there according to the scale factors.

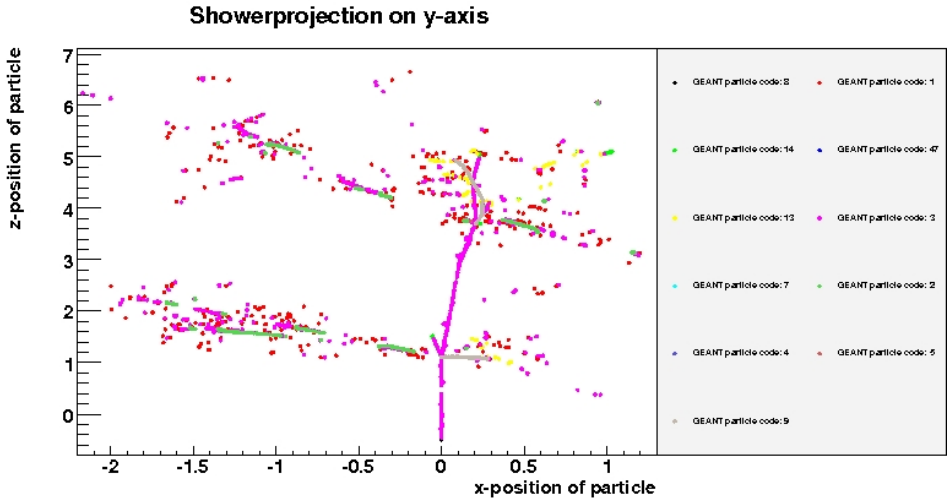
4.2.2 Hadronic Showers

As stated in section 3.2.1, the dependence of hadronic showers on the initial particle's energy is not as simple as for the electromagnetic case. Running *Gen* as described in section 4.1 produces the photon distributions shown in figure 4.5. The behavior in the low-energy regime has been discussed in section 3.2.2: if the initial energy of the pion is small, it decays almost instantly into a muon. For very small energies below $\simeq 100\,\mathrm{MeV}$ the energy of the created muon is also very small and it decays also virtually at the point of its origin. The created particles, mostly electrons in case of μ^- (if the initial particle was a π^-) or positrons in case of μ^+ (if the initial particle was a π^+), do not have enough energy to induce subshowers, so that in this case no energy-dependence of the photon yield can be observed in this region. It is noticeable that in this region the number of photons produced by π^+ is greater than the number of photons produced by π^-. This is possibly due to the fact that in the case of the initial particle being a π^+, positrons are produced in the decay series $\pi^+ \to \mu^+\nu_\mu$, $\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu$. These positrons finally annihilate producing γ. This would also explain why this effect is more important near the shower origin (figure 4.5 (a)) than further away (figure 4.5 (b)).

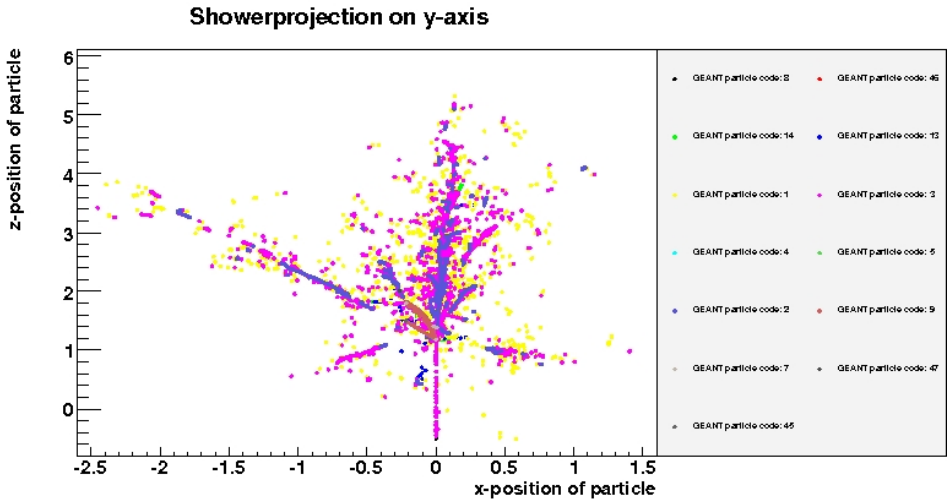
For the higher-energy part of the curve, nearly no difference between the two types of pions can be observed. This is because at higher energies there are produced so many particles that some slight differences at the beginning of the cascade are negligible.
For energies above 200 MeV the photon yield rises with the initial energy. For high initial energies, an asymptotic approximation to the electromagnetic linear curve is expected, because of the significant increase of the fraction made up by electromagnetic subshowers. This can be observed in figure 4.5 (b): for high energies the relation between the initial energy and the photon yield approximates a linear relation.The deviation from the linear curve in the lower



(a) Initial energy: 1.5 GeV



(b) Initial energy: 5 GeV



(c) Initial energy: 10 GeV

Figure 4.6: *Hadronic showers: initial energy 1.5 GeV (a): the initial pion (black) decays nearly instantly into a muon (green). The electrons (pink) along the muon track are probably delta electrons; initial energy 5 GeV (b): a wider range of particles is produced, but no compact shower shape is visible due to significant fluctuations; initial energy 10 GeV (c): a shower with hadronic and electromagnetic components evolves. The shape is compact and comparable to the shape of the electromagnetic showers as shown in 4.1. (For decoding the GEANT particle ID, refer to A.*

part (below $\simeq 10\,\mathrm{GeV}$) is due to the more significant fluctuations. The influence of these fluctuations is illustrated by figure 4.6. In (a) no shower evolves, the total energy is drawn away by a muon. In (b) more particles are produced, but the shape of the shower is not as compact as in the electromagnetic case (fig. 4.1 (a)). In (c) the shape is comparable to shape of an electromagnetic shower, fluctuations are less significant.
As in the case of electromagnetic showers, there are great differences between the inner and outer shells: In figure 4.5 quite strong deviations can be seen for the inner shells for higher energies, which is again due to the shower length exceeding the shell's radius. Because of that, the three most inner shells were neglected in finding the fit-function.

4.3 Fit of the Showers and Construction of a new One-Particle Approximation

The outcome of the last two sections has been that for the electromagnetic showers a linear model is appropriate. By contrast the model for the hadronic showers is more complex, since the curve approximates only for high energies a linear curve. Therefore the first step is to find the parameters for the linear fit of the electromagnetic showers. The fit-function for the hadronic showers is then modeled with the side condition that for high energies the functions approximates the linear curve of the electromagnetic case.
As a result of the assessment of validity done above, for both fits the three innermost shells are neglected.

Since the data are quite widely spread over several magnitudes of energy, the data will now be logarithmized before fitting to improve the results of the fit. For computational reasons and for modeling the asymptotic behavior of the hadronic showers versus the electromagnetic showers, it would be favorable to also fit the logarithmized data with a linear function. This would be possible if the photon yield would follow the relation $N_\gamma(E_e) = a_n \cdot E_e$, with $N_\gamma(E_e)$ the number of photons counted, E_e the energy of the initial electron, and a_n the characteristic slope for each shell, because then the transformed equation would be linear, too: $N'_\gamma(E_e) = \log N_\gamma(E_e) = \log a_n \cdot E_e = \log E_e + \log a_n = E'_e + k_n$. Though the the rest mass of an electron is not zero and it therefore needs to have at least an energy well above 511 keV for being able to travel relativistically in water and make Cherenkov radiation, the rest mass of an electron is so small compared to the energy range regarded here that it will be neglected, and the linear relation $N'_\gamma(E_e) = E'_e + k_n$ will be used, with $N'_\gamma(E_e) = \log N_\gamma(E_e)$ and $E'_e = \log E_e$ and $k_n = \log a_n$.
Figure 4.7 shows this fit applied to the all the shells larger than shell two, the data have been logarthmized before the fit. The fit-function used is the linear function

$$e(E'_e) = E'_e + k_n, \tag{4.1}$$

with $E'_e = \log E_e$.
For modeling the fit-function of the pion-induced hadronic shower, there was used a rational-function model. There are several reasons for that: mathematically, rational functions have the advantage of being able to take on an extremely wide range of shapes, accommodating a much wider range of shapes than does the polynomial family. Also, rational-function models have better interpolatory properties than polynomial models. Rational functions are typically smoother and less oscillatory than polynomial models. Rational function models can often be used to model complicated structure with a fairly low degree in both the numerator and denominator. This in turn means that fewer coefficients will be required compared to the polynomial model. They are also moderately easy to handle computationally. Although they are nonlinear models, rational function models are particularly easy nonlinear models to fit.

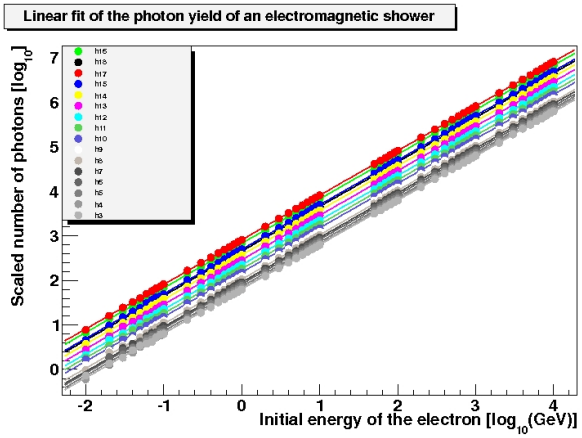


Figure 4.7: *Linear fit of logarithmized data for each shell.*

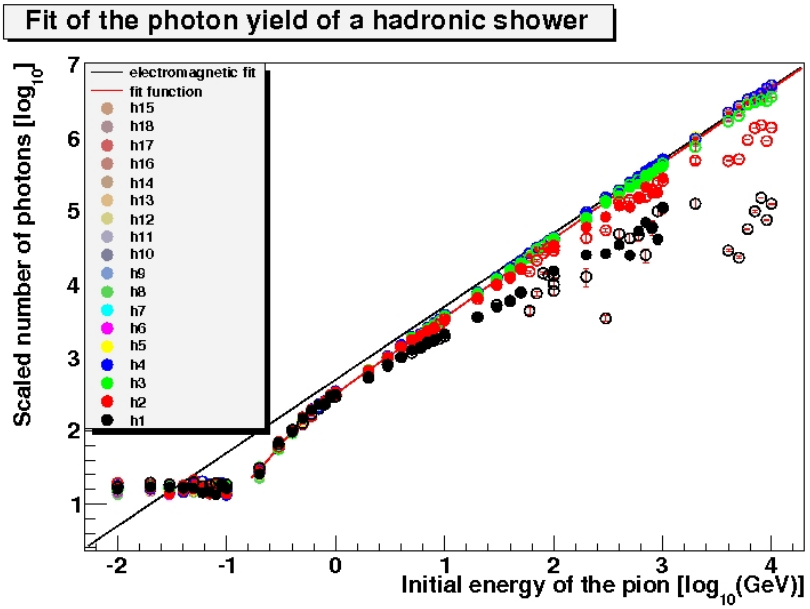
[15, p. 197]
But most important, rational functions turned out to be very useful for modeling the asymptotic character of the data at the high-energy range, since linear asymptotes can be modeled easily by choosing the degree of the numerator larger then the degree of the denominator by one.
The following fit-function was chosen:

$$p(E'_\pi) = \frac{\frac{a-f}{k_{ref}} \cdot (E'_\pi)^5 + a \cdot (E'_\pi)^4 + b \cdot (E'_\pi)^3 + c \cdot (E'_\pi)^2 + d \cdot (E'_\pi) + e}{\frac{a-f}{k_{ref}} \cdot (E'_\pi)^4 + f \cdot (E'_\pi)^3 + g \cdot (E'_\pi)^2 + h \cdot (E'_\pi) + i}, \tag{4.2}$$

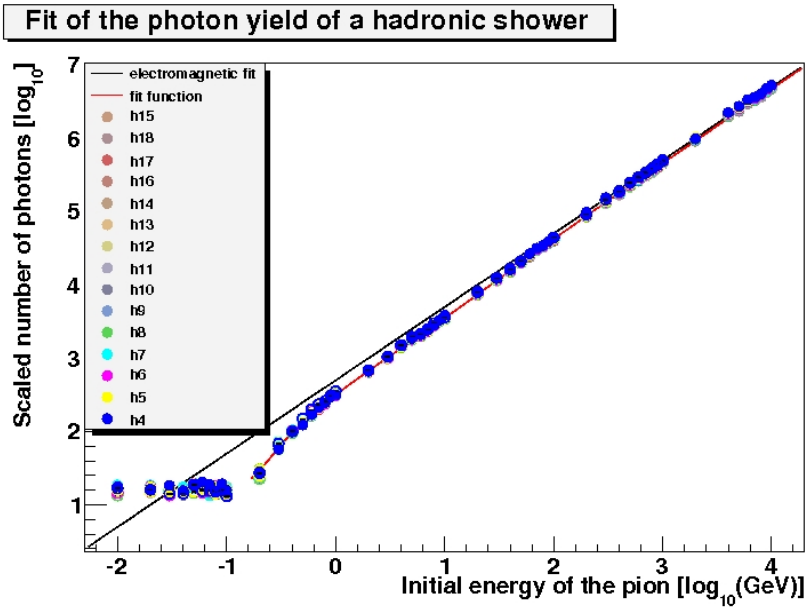
with $E'_\pi = \log(E_\pi)$. The first coefficient, $\frac{a-f}{k_{ref}}$, was formed that way to guarantee the asymptotic behavior versus the electrons' curve, e.g.

$$\lim_{E'_\pi \to \infty} p(E'_\pi) = e(E'_e) = E'_e + k_{ref}. \tag{4.3}$$

The degree of five for the numerator was chosen because the degree of four had not shown the desired fit, moreover, an even degree in the denominator seems more favorable, since singularities caused by roots of the denominator can be avoided. The data were again lorithmized to improve the fit's result, as in the case of the electromagnetic cascade, explaining the argument $E'_\pi = \log(E_\pi)$ of the function.
Since the difference between the yield of photons produced by π^+ and π^- in the range in which the fit-function shall be applied (>100 MeV) seems not significant (as discussed in section 4.1) only one fit for both types of pions will be made.
Having chosen the fit-function, the next step was to determine the function's coefficients by fitting the function to the data calculated by *Gen*. In contrast to the case of the fit of electromagnetic showers, where only one parameter had been varied, the function now has nine parameters. Nevertheless the aim is to find one fit-function for all shells, which leaves only one degree of freedom being the fraction of photons reaching at the different distances and being counted there according to the scale factors. This parameter, however, has already been determined with the fit of electromagnetic shower, corresponding the offset k_n between the individual shells. Therefore, once the nine parameters of the fit-function $p_{ref}(E'_\pi)$ have been found for one shell, the function is fixed for all shells.
But now the question arises which shell is to be used as "*reference shell*", e.g. which shell should be used for determining the nine fit parameters.
The following method was applied: to find the reference shell, all shells were for a test used as reference shell, then the RMS between the fit-function and the data was calculated for each shell, and the reference shell with the smallest overall RMS was used. Figure 4.8 shows the



(a) All shells including shell one to three



(b) All shells but shell one to three

Figure 4.8: *The fit-function with the parameters determined for shell fifteen (red line) and the data for each shell (solid: π^+, hollow π^-). In (a) all shells are plotted, including the three innermost shells which are not suited to being fit at the highest energies. The significant deviations of these shells from the fit-function can be seen clearly. (b) shows the good accordance of the outer shells with the fit function.*

result: shell fifteen was used as reference shell, e.g. it was used to calculate the parameters of the fit-function, which is applied to all shells by shifting it with an offset $c_n = k_n - k_{ref}$ which corresponds to the scaled fraction of photons counted on each shell:

$$N_\gamma(\text{scalefactor}, \text{shell}, E_\pi) = c_n \cdot N_\gamma(\text{scalefactor}, \text{referenceshell}, E_\pi). \tag{4.4}$$

In figure 4.8 the inverse is drawn: the data points of each shell are shifted with this offset c_n, so that they lie on the fit-function of the reference shell.

parameter	value for $\pi^{\pm}$	error
a	72.425	2.365
b	-49.417	9.376
c	5.858	6.872
d	207.252	10.143
e	132.784	0.312
f	-10.277	2.784
g	-19.441	2.374
h	58.598	4.109
i	53.161	0.118
k_{ref}	2.698	0.164

Table 4.2: *Parameters chosen for the fit-function 4.2. Shell fifteen was used as reference shell, e.g. the shell for which the overall RMS value is the smallest. The error is calculated by root. The ROOT class TMINUIT was used for fitting.*

Table 4.2 shows the parameters chosen for the fit-function. In this section the parameters of the linear fit-function $e(E'_e)$ of the electromagnetic showers and the rational fit-function $p(E'_\pi)$ of the hadronic showers have been determined. The last step is to join them together to gain the new one-particle approximation. It is divided into two parts:
The first part is valid for initial energies $> 100\,\mathrm{MeV}$. Setting $e(E'_e) = p(E'_\pi)$ gives for any arbitrary shell:

$$\begin{aligned} e(E'_e) &= p(E'_\pi) \Rightarrow \\ E'_e + k_n &= p(E'_\pi) + c_n \\ E'_e + k_{ref} + c_n &= p(E'_\pi) + c_n \\ E'_e &= p(E'_\pi) - k_{ref}. \end{aligned} \tag{4.5}$$

With $E'_e = \log E_e$ the one-particle approximation can be expressed by

$$E_e = 10^{(p(E'_\pi) - k_{ref})}. \tag{4.6}$$

For energies $< 100\,\mathrm{MeV}$, e.g. the flat part of the curve, no energy dependence of the photon yield was observed. Therefore in this range, the one-particle approximation is simply set as:

$$E_e = 10^{(p(E'_\pi = 100\,\mathrm{MeV}) - k_{ref})}. \tag{4.7}$$

Chapter 5

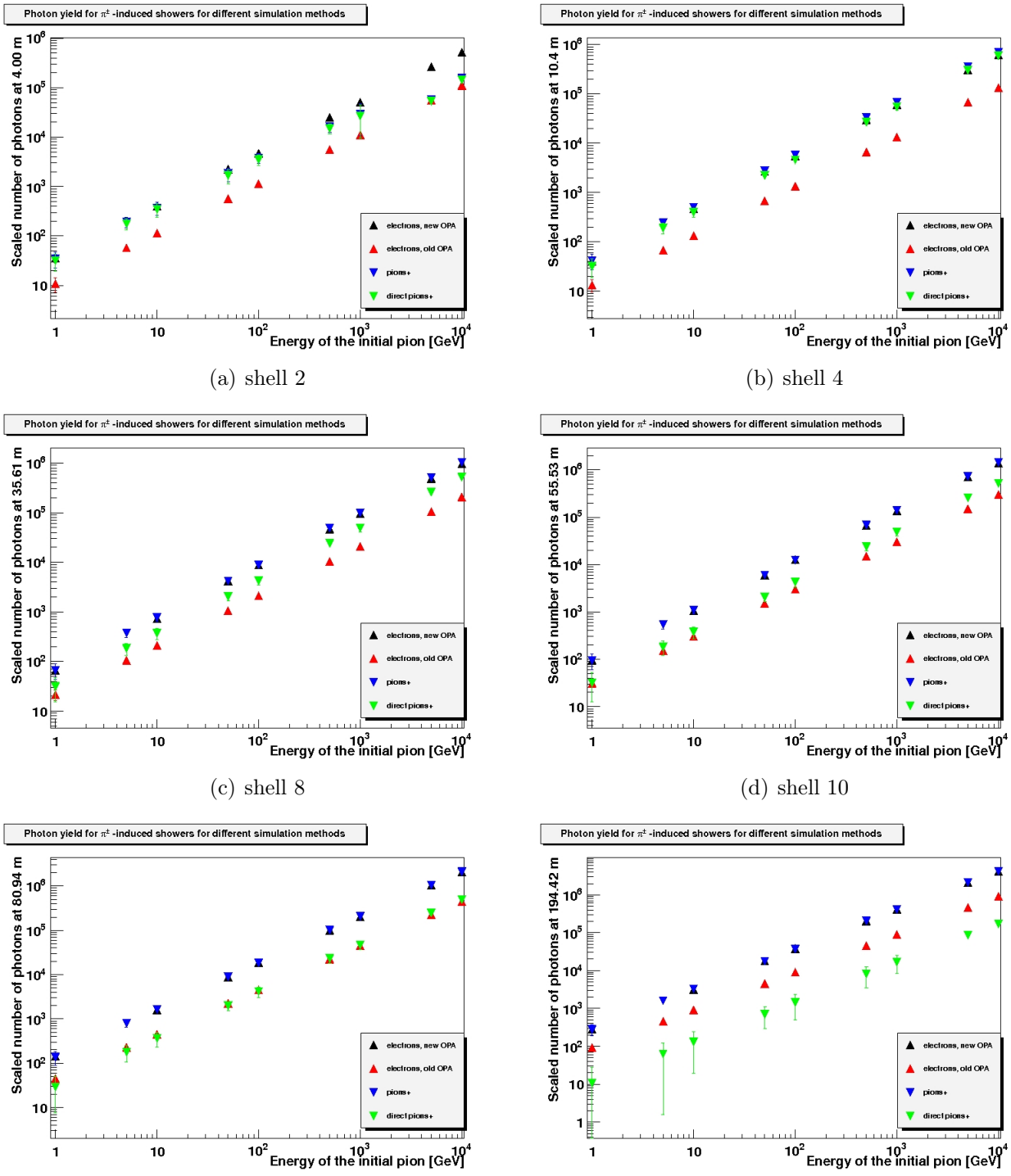
Evaluation and Conclusion

5.1 Comparing different Methods of simulating Shower Events

In this chapter a conclusive comparison between the methods used to simulate shower events[1] within ANTARES and the full simulation of a hadronic shower computed by *Gen* is made.
The first method is using *geasim*, a program which tracks all secondary particles of a hadronic shower down to Cherenkov level. The disadvantage of *geasim*, besides that it is very computationally complex and therefore time consuming, is that *geasim* does not account for scattered photons, which biases the photon distribution at high distances (refer section 2.3 for details).
The second method is the one-particle approximation within *KM3*, as discussed above. Here, two versions exist: the already-implemented one-particle approximation (*old OPA*), which replaces pions by electrons of only 20% the initial energy, and the new one-particle approximation (*new OPA*), which replaces pions by electrons of an energy computed according to the function modelled with respect to the parametrization done in chapter 4.
It is crucial to note though that the result of this comparison depends significantly of the distance at which the methods in question are evaluated. At low distances the number of photons will be made up nearly totally by *direct* photons, since the amount of scattering is correlated to the amount of target material. For low distances there is obviously less target material and therefore less scattering than for big distances. On the other hand, the actual shape of the shower, e.g. the longitudinal and angular profile as well as the fluctuations and the randomly deviations between showers have more influence on the number of photons counted at small distances, than they have for larger distances. This is why at small distances the fact that *geasim* does not account for scattered photons has little effect. By contrast the region closer to the origin, especially at high energies, is a critical range of the one-particle approximation, because here the intrinsic differences between electromagnetic and the hadronic shower become more evident than for larger distances at which both types appear virtually point-like. Therefore at small distances it is expected that *geasim* will give better results than the one-particle approximation.
Obviously at larger distances the opposite argument is true: here, scattering is more important, which means that the photon distribution computed by *geasim* becomes more biased. By contrast, the one-particle approximation works more accurately at larger distances, where the false assumption regarding the actual shower shape due to replacing a hadronic shower by an electromagnetic one has less impact.

Figure 5.1 shows the results of a simulation done with *Gen*: over the range of four orders of magnitude four sets of data were evaluated. The blue triangles show the total photon yield of hadronic showers induced by π^+. For simulating the output of *geasim*, which makes a full

[1] The term *shower event* refers to a hadronic shower which evolves as result of the energy transfered to the hadronic system in a neutrino event, e.g. a neutrino-nucleon interaction (refer to section 1.3 for details)



(a) shell 2

(b) shell 4

(c) shell 8

(d) shell 10

(e) shell 12

(f) shell 18

Figure 5.1: *Photon yield of hadronic showers induced by π^- (blue) as simulated with Gen. Green shows the yield if only direct photons are regarded, as it would be the case in geasim. Red shows the photon yield for the old one-particle approximation, in which the initial pion gets replaced by an electron of 20% of the energy. Black shows the photon yield for the new one-particle approximation, in which the initial pion gets replaced by an electron of the energy calculated according to the method explained in section 4.3. For clarity purposes π^+ are not plotted, since they do not show a significantly different curve.*

simulation of the shower, but does not take scattered photons into account, the same simulation is used, but this time only the direct photons are plotted (green triangles). Finally the two one-particle approximations are shown: the old one-particle approximation replacing the initial pion by an electron of 20% of the initial energy (red), and the new one-particle approximation as it was modeled in section 4.3 (black).
As expected according to the discussion in section 3.2.3, the result of the simulation depends significantly from the distance at which the data are evaluated. Figure 5.1 (a)-(f) shows the result for six selected shells, ranging in distance from 4.00 m to 194.42 m .
At low distances the first thing to notice is that there is no significant difference between the total number of photons and the number of scattered photons, which is the expected behavior according to the discussion above. This means that *geasim* would reflect the data almost exactly at this range. For assessing the quality of the one-particle approximation, the low energy regime must be distinguished from the high-energy regime. In the low energy regime (up to $\simeq 100\,\mathrm{GeV}$) the accordance between the one-particle approximation and the actual data is good, whereas at energies higher than that the accordance decreases. This is due to the fact that, at smaller distances, the shower shape matters more than at larger distances, where the events appear virtually point-like. With rising energy, the shower length increases, for example at an energy of 10 GeV the shower length of both electromagnetic and hadronic showers is below 4.00 m, the radius of the second shell, shown in figure 5.1 (a), which "closes in" the shower, which has the effect that the actual shape of the shower does not affect the number of photons counted at this shell. For higher energies the shower length exceeds the radius of the shell, therefore the shape of the shower, especially the longitudinal profile, e.g. the energy deposition per path length becomes important. In section 3.2.3 there is shown a graph (figure 3.2) of the parametrization of electromagnetic and hadronic showers. It can be seen that the energy deposition of the electromagnetic showers at energies of 1 TeV (red) and 10 TeV (green) is higher than the energy deposition of the hadronic showers at $4.00\,\mathrm{m} \simeq 11.4$ units of radiation length, which explains the fact that the new one-particle approximation overestimates the data in this energy range.
However at larger distances (Figure 5.1 (c)-(f)), the actual shower shape gets less and less important, since the shower is enclosed by the shell, and the shower size becomes less important in comparison with the total track length of the photons. Therefore the quality of the new one-particle approximation becomes better and better. By contrast scattering effects are more important at greater distances because the amount of scattering depends on the amount of target material as discussed above. This means that the output of *geasim* becomes less and less accurate at greater distances.
The old one-particle approximation generally underestimates the photon yield by a factor of $\simeq 3-4$ in the low-energy regime, which can be seen in all figures, and a factor of $\simeq 2$ in the high-energy regime, for the shells greater than shell four. The problem is that a linear approximation is only valid for high energies from $\simeq 3\,\mathrm{GeV}$ onwards, which can be estimated by eye looking at figure 4.5. This explains why the accordance between the old one-particle approximation and the actual data becomes better at high energies, though the factor 0.2 is estimated too small, given the fact that the electromagnetic component in hadronic showers makes up 95% at energies of 10 TeV. Nevertheless the bias of the data from *geasim* due to the neglect of scattering becomes so significant at high distances, that the old one-particle approximation becomes more exact. In the low energy regime the relation between the initial energy and the photon yield is not linear as can be seen clearly in figure 4.5, which explains the particularly bad performance of the old one-particle approximation in the low energy regime.
Summarizing the discussion above, two facts can be stated: The first statement is that in general, the new one-particle approximation performs better than the old one-particle approximation at distances larger than 4.00 m. The second statement is that the decision as to which

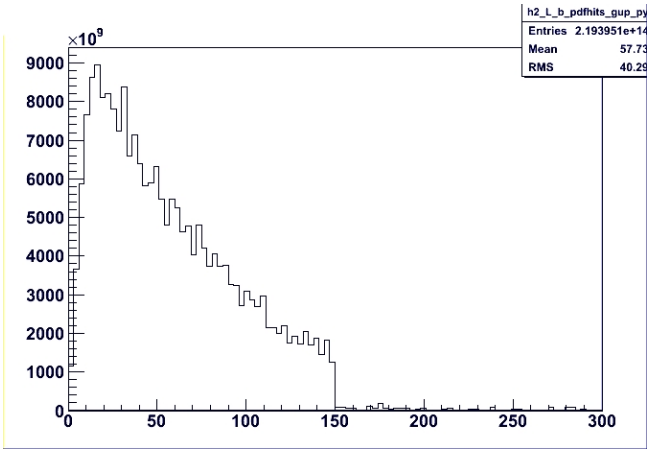


Figure 5.2: *Distribution of the photon track length from the origin to the OM in ANTARES. [16]*

method simulates the photon yield of a hadronic cascade better, the new one-particle approximation or *geasim*, depends on the distance at which the output is evaluated.

As a criterion to rate the two methods in question, figure 5.2 can be consulted. This figure shows the distribution of the photon track length from origin to the optical modules in ANTARES, reconstructed from Monte Carlo simulations assuming only direct photons. It can be seen in the figure that the mean distance between the event and the optical module in ANTARES is 57.73 m, while the maximum lies at about 35 m. Therefore the most relevant distances are shown in figure 5.1 (c) and (d). At these distances it can be seen clearly that the new one-particle approximation gives the best results, *geasim* is better than the old one particle approximation in the energy regime below $\simeq 1\,\mathrm{TeV}$, above that energy, no significant difference can be seen between the two simulation methods.

5.2 Conclusion

In this work a new version of the one-particle approximation in the ANTARES simulation software package *KM3* has been constructed. Before assessing this outcome, in the following the steps which have been taken to archive it are recapitulated.

The software for simulating neutrino events in the ANTARES detector as well as its limitations in regard to the simulation of hadronic events have been presented (see chapter 2). As a possibility to overcome these limitations, the one-particle approximation has been introduced and an overview of the current state of an already implemented one-particle approximation is given. The basic idea of the one-particle approximation is to replace all the hadronic and electromagnetic cascades produced by the secondary particles in a shower event with a single electron cascade at some energy, such that the total number of Cherenkov photons produced remains the same.

The feasibility of this idea, especially regarding the physical background of the approximation, has been discussed in chapter 3. The outcome of this discussion was that the one-particle approximation can give quantitatively a good approximation of the photon yield produced by a hadronic shower. Qualitatively however, with respect to the different shower shapes and the fluctuations of the hadronic showers in the low-energy regime (below $\simeq 10\,\mathrm{GeV}$), and the difficulties to apply a consistent parametrization at short distances (below $\simeq 10\,\mathrm{m}$) for high energies (above $\simeq 1\,\mathrm{TeV}$) for both shower types, the one-particle approximation will always be limited. These limitations are especially significant for observers near to the shower and at lower energies. By contrast, at further distances the shower event appears virtually point-like, thus, the differences in the shower shape are less significant. At higher energies, the electromag-

netic component of the hadronic showers increases, which obviously improves the results of the approximation. Therefore the one-particle approximation was judged a reasonable extension of the simulation software.
It was therefore the aim of this work to improve the already existing, preliminary one-particle approximation in the *KM3* software package. In order to do this, the photon yield of both electromagnetic showers and hadronic showers have been computed using the particle simulation and propagating program *Gen* as described in section 4.1. Subsequently, in section 4.2 a detailed assessment of the range of validity was given. Based on this evaluation a parametrization of both shower types with respect to the photon yield was computed. The resulting function was then used to implement a new version of the one-particle approximation.
In section 5.1 a comparison between this new one-particle approximation and the current default program *geasim* on one hand and the old one-particle approximation on the other is undertaken. Two facts are shown in that evaluation: firstly, the new one-particle approximation performs better than the old one-particle approximation at distances larger than 4.00 m; secondly, the decision which method simulates the photon yield of a hadronic cascade better, the new one-particle approximation or *geasim*, depends on the distance at which the output is evaluated. At the distances which have been judged to be most relevant to ANTARES (refer to section 5.1 for details), it can be seen clearly that the new one-particle approximation gives the best results, while *geasim* is better than the old one particle approximation in the energy regime below $\simeq 1\,\mathrm{TeV}$; above that energy, no significant difference can be seen between these two simulation methods.
Therefore it can be stated conclusively that the new particle approximation is definitely an improvement of the old one-particle approximation, especially at lower energies. Furthermore, the results of the one particle approximation are more exact than the output of *geasim*, which does not take into account scattered photons, from distances of $\simeq 4.00\,\mathrm{m}$ onwards, including the distances most relavant to ANTARES. Besides being less accurate, *geasim* has also the disadvantage of being computational complex and therefor time-consuming at high energies, as discussed in section 2.3. Therefore the implementation of the new one-particle approximation can be advised.
Perspectively though, there can be imagined further improvements of the new one-particle approximation. Although pions are by-far the biggest hadronic component of a hadronic shower induced by an neutrino, it could be worthwhile to find parametrizations of the other hadronic components of the shower, which are currently set to zero. It might be also possible to combine the one particle approximation with *geasim*. This could be done by a partly-one-particle approximation, which would not replace all secondary pions in the shower by electrons, but simulate a fraction of them using *geasim*. With all these considerations, the desired accuracy has to be balanced with the computational complexity and time-consumption of the simulation.
Having stated this, the new one-particle approximation is a good compromise between improving accuracy of the results of the simulation and keeping the computational complexity low, being more accurate than the old one-particle approximation on one hand and both more accurate in the relevant distances and less computional complex than *geasim*. Therefore the new one particle approximation is step forward although there remain always possibilities for improvement.

Appendix A

GEANT particle ID

GEANT part. ID	Particle
1	Gamma
2	Positron
3	Electron
4	Neutrino
5	Muon +
6	Muon -
7	Pion 0
8	Pion +
9	Pion -
10	Kaon 0 long
11	Kaon +
12	Kaon -
13	Neutron
14	Proton
15	Antiproton
16	Kaon 0 short
17	Eta
18	Lambda
19	Sigma +
20	Sigma 0
21	Sigma -
22	Xi 0
23	Xi -
24	Omega
25	Antineutron
26	Antilambda
27	Antisigma -
28	Antisigma 0
29	Antisigma +
30	Antixi 0
31	Antixi +
32	Antiomega +
45	Deuteron
46	Tritium
47	Alpha
48	Geantino
49	He3

Table A.1: *GEANT particle ID and the corresponding particles. [18]*

Appendix B

Gen Parameters

Card	Parameter	Description
SEED	3937	first random number seed
WMOD	partic	water model id.
ETA	0.17	fraction of rayleigh scatt. (partic model) (ignored if wlsc != 0)
WLSC	1	method for calculating eta (1=use kopelevich)
IABS	1	lambda abs. from input (if 0 from Geasim)
RNDM	8432 7493	second rndm. number group
MUNU	0	muon-nucleus interactions
HADR	4	hadr. interactions
PFIS	1	photon induced nuclear fission
DRAY	1	delta ray production
LOSS	1	continuous energy loss
ERAN	1.e-5, 1.e7, 120	emin emax nbins for cross-section tables
CUTS	0.0002	tracking cut for gammas, electrons,
		neutral and charg. hadrons, muons, bremmss., ...

Table B.1: *Variables (Cards) used for running Gen. For details refer to [6, p. 7]*

Bibliography

[1] http://antares.in2p3.fr/Overview/index.html

[2] B. Hartmann, *Reconstruction of Neutrino-Induced Hadronic and Electromagnetic Showers with the ANTARES Experiment*, Dissertation, 2006

[3] D. Griffiths, *Intoduction to Elementary Particles*, second edition, Wiley-VCH, 2008

[4] Povh · Rith · Scholz · Zetsche, *Teilchen und Kerne*, 8. Auflage, Springer, 2009

[5] S. Navas and L.Thompson, *KM3 User Guide and Reference Manual*, ANTARES-Soft, 1999

[6] D. Bailey, *KM3 v2r1 : User Guide*, ANTARES-Software, 2002

[7] CERN Application Software Group, *GEANT â Detector Description and Simulation Tool*, CERN Program Library Long Writeup W5013, 1993

[8] J. Brunner,*Geasim: User manual*, ANTARES Internal documentation, 2000

[9] G. De Bonis and F. Folger, *GEASIM OPA*, talk at the ANTARES Collaboration Meeting, Strasbourg, 2011

[10] S. Biagi, *Simulation of MC events: Electron Neutrinos and NC*, talk at the ANTARES Collaboration Meeting, Bologna, 2008

[11] H. Kolanoski, *Einfuhrung in die Astroteilchenphysik*, Vorlesung an der Humboldt-Universität zu Berlin, 2010

[12] J. Brunner, *Cherenkov light from HE electromagnetic and hadronic showers*, ANTARES-Soft, 2002

[13] http://root.cern.ch/drupal/content/documentation

[14] http://root.cern.ch/root/html/TMinuit.html

[15] V. Melas, *Functional Approach to Optimal Experimental Design*, Springer, 2006

[16] *Internal communication*, ANTARES cooperation, 08/2012

[17] K. Nakamura (Particle Data Group), *Particle Physics Booklet*, extracted from the Review of Particle Physics, Journal of Physics G 37 075021 (2010)

[18] http://wwwasdoc.web.cern.ch/wwwasdoc/geant/node72.html

Danksagung

An dieser Stelle möchte ich mich zu aller erst bei meinen Eltern bedanken, die mir dieses Studium überhaupt erst ermöglichen und mich stets auf jede erdenkliche Weise unterstützen und auf die ich mich immer verlassen kann.
Außerdem möchte ich dem ECAP für die freundliche Aufnahme und ganz besonders natürlich Clancy James für die super Betreuung danken. Thank you Clancy for your support and your patience, *sensitivly* helping me to make *sensible* English sentences without mixing up "derivation" with "deviation"...:-) Weiterhin möchte ich meinen netten Zimmergenossen für die gute und manchmal recht lustige Atmosphäre danken und schließlich meiner Schwester und allen Freunden die stets ein offenes Ohr für mich hatten und meine mitunter schlechte Laune ertragen haben!

Erklärung

Hiermit bestätige ich, dass ich diese Arbeit selbstständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe.

Erlangen,

Mona Dentler