Search for cosmic neutrino emission from Milagro sources with ANTARES

Der Naturwissenschaftlichen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg
zur Erlangung des Doktorgrades Dr. rer. nat.

vorgelegt von
Stefanie Lucia Wagner
aus Forchheim

Als Dissertation genehmigt
von der Naturwissenschaftlichen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung:	06. Mai 2015
Vorsitzender des Promotionsorgans:	Prof. Dr. Jörn Wilms
Gutachter/in:	Prof. Dr. Ulrich Katz
	Prof. Dr. Gisela Anton

Zusammenfassung

Im Jahr 2013 wies das IceCube Experiment erstmalig kosmische Neutrinoemission nach, indem kaskadenartige hochenergetische Neutrinoereignisse gemessen wurden [1]. Kosmische Neutrinos werden unter anderem erzeugt, wenn Protonen, aus welchen die kosmische Strahlung bei Energien oberhalb von 100 GeV hauptsächlich besteht, mit Photonen, Protonen oder anderen Kernen wechselwirken und dabei Kaonen oder Pionen erzeugen. Neutrale Pionen zerfallen zu Gammaquanten, geladene hingegen zerfallen in Myonen und Myon-Neutrinos. Die Myonen können weiter zerfallen und so entstehen Neutrinos. Gammastrahlung kann auch in elektromagnetischen Prozessen, wie dem inversen Compton-Effekt oder Bremsstrahlung, erzeugt werden, wohingegen Neutrinos nur in hadronischen Wechselwirkungsprozessen erzeugt werden können. Daher könnte die parallele Messung von Gammastrahlung und Neutrinos von den selben Quellkandidaten die Fragen nach den unbekannten Quellen kosmischer Strahlung und nach der Zusammensetzung und der Natur leptonischer und hadronischer Prozesse innerhalb kosmischer Beschleuniger lösen. Mögliche Neutrino Quellkandidaten können in zwei Gruppen eingeteilt werden: galaktische Quellen, z.B. Supernovaüberreste (SNRs), und extra-galaktische Quellen, z.B. Gammastrahlungsblitze (GRBs). Milagro detektierte Gammastrahlungsemission von mehreren galaktischen Quellen, sogenannte PeVatrons, von denen man annimmt, sie würden kosmische Teilchen bis zu Energien von mehreren PeV beschleunigen. Einige von diesen Quellen liegen in der Cygnus Sternenbildungsregion, die ein Bereich mit hoher Sternbildungs- und Sternsterberate ist. Solche Regionen weisen eine hohe Dichte sowohl an molekularen Wolken als auch Supernovaüberresten auf. Wegen der hohen molekularen Dichte der Wolken steigt die Wahrscheinlichkeit einer Wechselwirkung von Protonen, die von SNRs erzeugt und beschleunigt wurden, mit den Atomen der Wolke und somit steigt die Wahrscheinlichkeit für Neutrinoemission im Vergleich zu weniger dichten Regionen. Eine Untermenge der acht mit Milagro gemessenen galaktischen TeV-Gammastrahlungsquellkandidaten ([2],[3]) wurde in [4] untersucht, mit der Vorhersage Neutrinos von diesen Quellen innerhalb von einigen Jahren Datennahme mit dem IceCube Detektor nachweisen zu können. Die Milagro Quellen MGRO J2019+37, MGRO J2031+41, MGRO J1908+06, C1 und C2 sind auch mit dem ANTARES Detektor sichtbar und wurden in dieser Arbeit für die Untersuchung von punktförmiger Neutrinoemission mit ANTARES ausgewählt.

Der ANTARES Detektor, der für den Nachweis von Neutrinos mit Energien im GeV bis in den PeV Bereich optimiert wurde, befindet sich im Mittelmeer und besteht aus 885 optischen Modulen (OMs), die an insgesamt 12 vertikalen Strukturen, sogenannten Lines, installiert sind. Die Lines sind 480 m hoch und sind in einem Abstand von 60 bis 70 m am Meeresboden in einer Tiefe von ca. 2.5 km verankert. Da Neutrinos neutral geladen sind und nur schwach mit Materie wechselwirken, benötigt man ein großes Detektionsvolumen, welches von der Energie der zu untersuchenden Neutrinos sowie vom Fluss derselben abhängt. Der ANTARES Detektor instrumentiert ein Volumen von 0,01 km^3 in der Tiefsee. Ein indirektes Nachweisprinzip wird genutzt, da Neutrinos nur über die schwache Wechselwirkung mit Atomen oder Molekülen inter-

agieren können und dabei hochenergetische Leptonen in geladenen Stromreaktionen erzeugen. Der Fokus liegt hier auf Myon-Neutrinos, die in der geladenen Stromreaktion wechselwirken, da für eine Punkquellenanalyse die gute Winkelauflösung der langen Spur im Detektor nötig ist, welche die erzeugten Myonen erfüllen. Diese relativistischen Teilchen erzeugen Licht über den Cherenkov-Effekt indem sie sich schneller als Licht in dem transparenten Medium Wasser bewegen. Das transmittierte Licht formt einen Lichtkegel unter einem typischen Winkel von $\Theta_c = 42^\circ$ für Wasser. Indem man die Position und die Ankunftszeit der Photonen, die mit den OMs detektiert wurden (=Hits), bestimmt, ist es möglich die Spur der erzeugten Myonen und damit der Neutrinos zu rekonstruieren, da der Zwischenwinkel zwischen Myonen und Neutrinos für Energien oberhalb von mehreren TeV vernachlässigbar ist.

Eine neue Myonen-Spurrekonstruktionsmethode wurde in dieser Arbeit entwickelt um alternative bzw. zusätzliche Ereignisinformationen im Vergleich zu den verwendeten Standardalgorithmen zu bekommen. Die Strategie besteht aus einer Hit-Selektion, einem vorläufigen Fit (Pre-Fit), einen Haupt-Fit und Qualitätsschnitten. Der Pre-Fit basiert auf einer Abrasterung des kompletten Himmels in diskreten Winkelschritten. Indem jede untersuchte Richtung festgehalten wird, muss nur noch die Position der Spur mit einer linearen Anpassung gefunden werden. Für jede Richtung wird eine eigene Hit-Selektion durchgeführt, die eine Kausalitätsbeziehung zwischen Hit-Paaren benutzt. Direkte Signal-Hits müssen vorselektiert werden, da optischer Untergrund, der durch den Zerfall von im Meereswasser vorhandenem ^{40}K oder durch lichtemittierende Fauna (Biolumineszenz) erzeugt wird, im Detektor zufällig verteilt ist und gestreute Signal-Hits, die von der Kegelform abweichen, das Verhalten des Fits beeinflussen können. Jede gefittete Richtung am Himmel bekommt ihre eigenen Qualitätswert, der auf χ^2 und der Anzahl an Hits, die in der linearen Fit Prozedur verwendet wurden, basiert. Die zehn Spurhypothesen mit den höchsten Qualitätswerten werden als Basis für weitere Selektionskriterien verwendet. Zunächst werden Spuren die offensichtlich das Hit-Muster nicht gut beschreiben, ausgefiltert. Das dafür verwendete Kriterium ist das sogenannte Zylinderkriterium, wobei ein Zylinder um die Spur gelegt wird, der alle Hits mit geringen Zeitresiduen (um gestreute Hits zu unterdrücken) enthält. Falls das Verhältnis von Zylinderradius r zu Zylinderhöhe h größer als 1 ist, wird die Spur ausgefiltert. Diese Prozedur ist sehr effizient: ungefähr 80% der rekonstruierten Spuren mit Winkelfehlern größer als 60° werden ausgefiltert, wobei auf der anderen Seite 98,8% der rekonstruierten Spuren mit Winkelfehlern kleiner als 5° behalten werden. Aus der Untermenge an Spuren, die das Zylinderkriterium erfüllen, wird nun eine Spur ausgewählt, die ein Oberflächendichtekriterium der Hits um die Spur am besten erfüllt. Um eine bessere Auflösung zu erhalten wird zudem ein sogenannter MEstimator-Fit benutzt, welcher den Einfluss von Hits mit großen Zeitresiduen unterdrückt. Der Haupt-Fit basiert auf der Methode der maximalen Wahrscheinlichkeit (Maximum-Likelihood), welche die Wahrscheinlichkeitsdichtefunktion (PDF) der Zeitresiduen der Hits verwendet und maximiert. Der Qualitätsparameter, welchen diese Methode zu Verfügung stellt, ist der reduzierte Logarithmus der Wahrscheinlichkeit, kurz rlogL, der sich berechnet als der Logarithmus der Wahrscheinlichkeit geteilt durch die Anzahl an Freiheitsgraden. Hierauf sind Schnitte möglich, je kleiner der rlogL-Wert ist, desto besser ist das Fit-Ergebnis. Die komplette Fit-Prozedur wird KrakeFit genannt.

Mögliche Qualitätsschnittparameter wurden abgeleitet um nur gut rekonstruierte Ereignisse auszuwählen und zusätzlich den atmosphärischen Myonenuntergrund für fälschlicherweise als von unten kommend rekonstruierte atmosphärische Myonen zu unterdrücken. Für von oben kommende Ereignisse ist der amtosphärische Myonenuntergrund um einen Faktor von ca. 10^6 größer als für atmosphärische Neutrinos. Das Verhalten von KrakeFit wurde mit den standard ANTARES Spurrekonstruktionen BBFit und AAFit verglichen. Der Median des Winkelfehlers

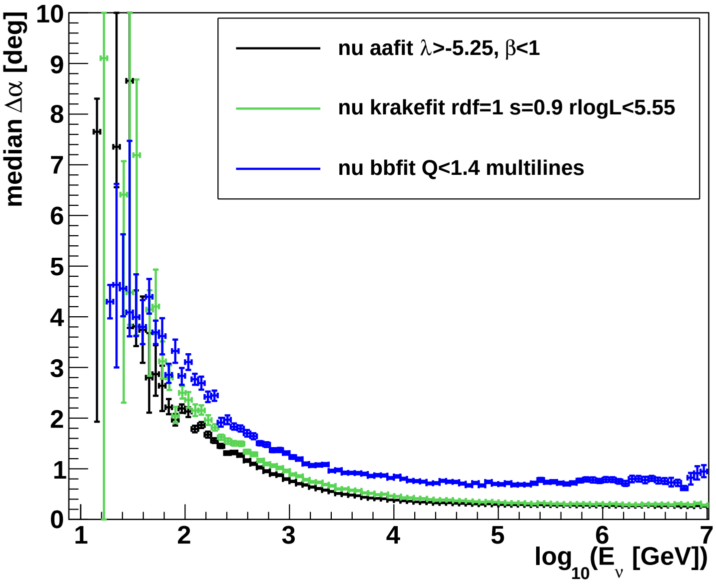


Abbildung 1: Median des Winkelfehlers $\Delta\alpha$ für von unten kommende Neutrinos als Funktion der Neutrinoenergie für AAFit ($\lambda > -5.25$, $\beta < 1°$) (black), BBFit ($Q < 1.4$, NLines > 1) (blue) und KrakeFit (rdf=1, $s > 0.9$, rlogL < 5.55) (green) für ein E^{-2} Spektrum.

in Abhängigkeit von der Energie ist in Abb. 1 gezeigt. BBFit, welches am effizientesten für niederenergetische Neutrinoereignisse ist, hat ein besseres Verhalten als KrakeFit für Ereignisse mit Energien unterhalb von 50 GeV. Für hochenergetische Neutrinoereignisse ist das Verhalten von KrakeFit und AAFit, welches die standard hochauflösende Spurrekonstruktionsstrategie von ANTARES ist, vergleichbar, nachdem Qualitätsschnitte angewendet wurden, die für beide eine vergleichbare Anzahl an falsch rekonstruierten atmosphärischen Myonen ergibt. Ein Vergleich der Ereignisse, die mit KrakeFit und AAFit nach Anwendung von Qualitätsschnitten rekonstruiert wurden, ergibt, dass 35,9% aller getriggerten Ereignisse mit beiden Strategien rekonstruiert werden, 20,8% werden nur mit AAFit rekonstruiert und 5,6% werden zusätzlich mit KrakeFit rekonstruiert. Daher bekommt man durch KrakeFit zusätzliche Neutrinoereignisinformation zu AAFit, was dazu genutzt werden kann um die Neutrinoausbeute zu erhöhen. Der mittlere Zwischenwinkel zwischen den Ergebnissen dieser beiden Spurrekonstruktionen für Ereignisse, die mit beiden als von unten kommend rekonstriert wurden beträgt 1,4°.

KrakeFit ist eine Rekonstruktionsstrategie mit hoher Winkelauflösung - für Neutrinos mit Energien über 1 TeV (10 TeV) ist die Auflösung besser als 0,88° (0,45°) - und einer guten Effizienz für die Rekonstruktion von (hochenergetischen) Neutrinos. Um zu überprüfen ob die Ereignisrekonstruktion arbeitet wie erwartet und ob die Run-für-Run (rbr) basierte Monte Carlo-Simulation (MC) die Daten gut beschreibt wurde ein Daten-MC-Vergleich durchgeführt. Dabei wurden generelle Observablen von KrakeFit wie die Verteilung des Zenit- und Azimut-Winkels, der Anzahl der Hits und Lines, die vom Fit verwendet wurden, für eine Run-Auswahl untersucht. Diese enthält ANTARES Runs, die von 2008 bis 2012 mit der kompletten Detektorkonfiguration aufgezeichnet wurden und einer Lebenszeit von 744,68 Tagen entspricht. Nachdem Qualitätsschnitte angewendet wurden, mit Hauptqualitätsschnitt auf den rlogL-Wert, wurde eine gute Übereinstimmung zwischen Daten und MC beobachtet. Zusätzlich zu den Standard-Observablen wurden weitere Parameter zur Unterscheidung zwischen atmosphärischen Myonen und Neutrinos untersucht. Einige davon sind gute Kandidaten als Schnittparameter. Das sind die Länge der direkten Hits (ausgewählt als Hits mit geringen Zeiresiduen zur rekonstruierten Spur) projeziert auf die Spur, der Sicherheitsparameter des rdf (random decision forest zur Klassifizierung von von unten kommenden und von oben kommenden Myonenspuren) und der Zwischenwinkel zwischen Ereignissen die mit KrakeFit und AAFit rekonstruiert wurden. Da jede Analyse ihre eigenen Anforderungen hat, müssen diese Parameter für jede Analyse separat analysiert und optimiert werden, eventuell auch in einer Kombination von mehreren Parametern.

In dieser Arbeit wird für die Untersuchung der Milagro Quellen die KrakeFit Spurrekonstruktion verwendet. Um den Parameterraum auf eine untersuchbare Anzahl an Parametern zu beschränken werden für die Sensitivitätsstudie einer punktförmigen Neutrinoemission von den Milagro Quellkandidaten nur der rlogL-Wert und ein zusätzlicher Schnitt auf rdf=1 (rdf=1 für von unten kommenden Spuren) untersucht. Die optimalen Schnittwerte wurden für jeden Milagro Quellkandidaten und seiner Neutrinoflussvorhersage im Hinblick eines größtmöglichen Modell-Entdeckungspotentials (MDP) für eine 3σ Entdeckung separat bestimmt. In [4] und [5] wurde die Neutrinoflussvorhersage für MGRO J2019+37, MGRO J2031+41, MGRO J1908+06, C1 und C2 von Messungen des Gammastrahlungsspektrums abgeleitet. Man erwartet, dass dieser einem Potenzgesetz mit exponentiellem Abfall für hohe Energien folgt:

$$\frac{dN_\nu(E_\nu)}{dE_\nu} = k_\nu \left(\frac{E_\nu}{\mathrm{TeV}}\right)^{-\alpha_\nu} \exp\left(-\sqrt{\frac{E_\nu}{E_{cut,\nu}}}\right). \tag{1}$$

Für jede Quelle und Neutrinoflussvorhersage wird eine punktförmige Neutrinosimulation mithilfe der Run-für-Run Simulationsmethode durchgeführt. Da es mehrere CPU-Jahre dauern würde um eine vollständige rbr-Simulation für die Run-Auswahl für eine Quelle durchzuführen wird stattdessen eine kleine Untermenge an Runs verwendet, die definierte Standard-Detektorbedingungen erfüllt und somit den Hauptanteil an Runs repräsentiert. Mit diesen Simulationen wird die Signal Wahrscheinlichkeitsdichtefunktion (PDF) erzeugt, die ein Bestandteil der Maximum-Likelihood-basierten Punktquellenanalysemethode ist. Als Signal PDF S_i wird die Winkelverteilung zur Quellenposition der als von unten kommend rekonstruierten Signal-Ereignisse nach Qualitätsschnitten verwendet. Das ist die sogenannte PSF (point spread function). Als Untergrund PDF B_i wird die Ereignisrate in Abhängigkeit von $\sin(\delta)$ der Deklination δ nach Qualitätsschnitten verwendet. Beide PDFs sind normiert, sodass jeweils ihr Integral über den Raumwinkel 1 ergibt. Eine Maximum-Likelihood-basierte Methode wird gebraucht, da nach Messung einer Spur, die als von unten kommend rekonstruiert wurde, nicht zwischen kosmischen Neutrinos, atmosphärischen Neutrinos und falsch rekonstruierten atmosphärischen Myonen unterschieden werden kann. Deshalb ist ein statistisches Verfahren nötig um zu interpretieren wie signal- oder untergrundartig das Ergebnis ist. Dies wird hier durch einen Hypothesentest gemacht, wobei die Nullhypothese lautet, dass die gemessenen Daten nur aus Untergrundereignissen bestehen. Anstatt die Wahrscheinlichkeit L zu maximieren wird die Teststatistik Q maximiert, welche als

$$\log L(n_s) = \sum_{i=0}^{N} \log\left(\frac{n_s}{N} S_i + \left(1 - \frac{n_s}{N}\right) B_i\right) \tag{2}$$

bzw.

$$Q = \log\left(\frac{\max(L(n_s))}{L(n_s = 0)}\right) = \max(\log L(n_s) - \log L(n_s = 0)) \tag{3}$$

definiert sind, wobei N die Gesamtanzahl an Ereignissen in Daten nach Qualitätsschnitten und n_s die Anzahl an vermuteten Signalereignissen ist. Der Wert μ_s, der der Wert von n_s ist, welcher die Wahrscheinlichkeit bzw. die Teststatistik maximiert, ist die beste Abschätzung für die wahre Anzahl an Signal-Ereignissen. Pseudo-Experimente wurden generiert indem nur Untergrund- bzw. Untergrundereignisse mit vorhandenen Signalereignissen generiert wurden. Das Ergebnis wurde statistisch interpretiert. Anhand der Verteilung von Q für reinen Untergrund wurden die Grenzen von Q für eine 3σ und eine 5σ Entdeckung bestimmt. Es wurde gezeigt, dass der mittlere gefittete Wert von μ_s der Anzahl der injizierten Signal-Ereignisse entspricht und daher der Hypothesentest durchgeführt werden kann. Die besten Schnittwerte (rlogL-Wert mit und ohne rdf=1) wurden optimiert in Hinblick auf das Modell-Entdeckungspotentials (MDP) um die beste Wahrscheinlichkeit einer Entdeckung für die Anzahl der zu erwartenden Ereignisse (die aus der punktförmigen Quellensimulation für jede Quelle und Neutrinoflussvorhersage bestimmt wurden) zu erhalten. Die entsprechende Sensitivität wurde durch Multiplikation des erwartenden Neutrinoflusses mit dem Verhältnis aus der Anzahl der nötigen Ereignisse für eine 50% Chance für eine 3σ Entdeckung $n_{3\sigma}$ und der erwartenden Anzahl an Ereignissen bestimmt. Eine mittlere obere Grenze für 90% C.L. (confidence level) $\bar{\Phi}_{90}$ wurde auf den Neutrinofluss gesetzt. Der Fokus liegt hier bei den Werten, die für die 3σ Optimierung bestimmt wurden, da die Anzahl an erwarteten Neutrinoereignissen in der Größenordnung von 10^{-3} liegt.

Da für die Quellen MGRO J1908+06, MGRO J2019+37 und MGRO J2031+41 mehr als eine Neutrinoflussvorhersage vorhanden ist, werden die finalen Ergebnisse in Tab. 1 für die Neutrinoflüsse und Schnitte mit dem höchsten MDP angegeben. In Abb. 2 werden die Sensitivitäten der fünf Milagro Quellen für eine 3σ Entdeckung gezeigt. Die Sensitivität ist für alle Quellen ungefähr einen Faktor 10^2 höher als der erwartete Neutrinofluss. Es ist daher unwahrscheinlich eine signifikante Anzahl an Neutrinos für eine 3σ Entdeckung zu messen. Das kommt daher, dass die Sichtbarkeit der Quellen im ANTARES Detektor mit einer Sichtbarkeit von 23% bis 30%, außer MGRO J1908+06 mit einer Sichtbarkeit von 48%, klein ist und durch das erwartete Energiespektrum mit exponentiellem Abfall, die meisten Neutrinoereignisse eher niederenergetisch sein werden und damit im atmosphärischen Untergrund verschwinden.

Quelle	Fluss	rlogL	rdf	$n_{3\sigma}$	$\Phi_{3\sigma}$ ($\mathrm{TeV^{-1}cm^{-2}s^{-1}}$)	$\bar{\Phi}_{90}$ ($\mathrm{TeV^{-1}cm^{-2}s^{-1}}$)
MGRO J1908+06	halzen3	5.1	-	1.81	$6.3 \cdot 10^{-10}$	$1.35 \cdot 10^{-9}$
MGRO J2019+37	halzen3	5.2	1	0.93	$3.5 \cdot 10^{-10}$	$9.70 \cdot 10^{-10}$
MGRO J2031+41	kappes	5.3	-	1.07	$8.0 \cdot 10^{-10}$	$2.20 \cdot 10^{-9}$
C1	kappes	5.2	1	1.01	$8.7 \cdot 10^{-10}$	$2.28 \cdot 10^{-9}$
C2	kappes	5.2	1	0.94	$6.3 \cdot 10^{-10}$	$1.76 \cdot 10^{-9}$

Tabelle 1: Zusammenfassung der Neutrinoflussvorhersage und optimierte Schnitte für das beste 3σ Entdeckungspotential pro Quelle. Die Anzahl an nötigen Ereignissen für eine 3σ Entdeckung $n_{3\sigma}$, die Normierung der Sensitiviät $\Phi_{3\sigma}$ und die mittlere obere Grenze für 90% C.L. $\bar{\Phi}_{90}$ sind angegeben.

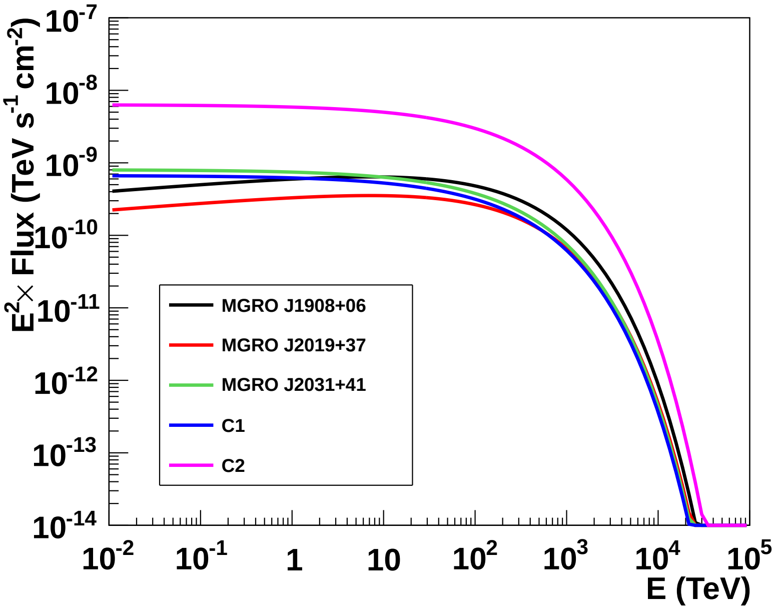


Abbildung 2: Sensitivitäten der fünf Milagro Quellen für eine 3σ Entdeckung für die Neutrinoflussvorhersagen und optimierten Schnitte von Tab. 1.

Für MGRO J1908+06 ist die Neutrinoflussvorhersage, die Sensitivität und die 90% C.L. mittlere obere Grenze, die in dieser Arbeit bestimmt wurden, sowie ein Vergleich zur 90% C.L. oberen Grenze, die mit ANTARES [6] bzw. IceCube [7] für diese Quelle bestimmt wurden in Abb. 3 gezeigt. Die letzteren beiden verwendeten ein E^{-2} Energiespektrum ohne exponentiellem Abfall. Die offizielle ANTARES Analyse berücksichtigt eine größere Run-Auswahl, die einer um das 1,8-fache größeren Lebenszeit als der in dieser Arbeit berücksichtigten entspricht, und das vermutete Neutrinospektrum ohne exponentiellem Abfall ergibt eine höhere Anzahl an erwarteten Neutrinos und daher eine bessere obere Grenze. Es muss erwähnt werden, dass der vermutete E^{-2} Neutrinofluss für diese Quellen nicht physikalisch motiviert ist. Des Weiteren ist der Energiebereich mit den meisten erwarteten Ereignissen mit diesem Fluss nicht angegeben. Der Energiebereich von IceCube für Quellen der nördlichen Hemisphäre wird mit 1 TeV und 1 PeV angegeben. Weder IceCube noch ANTARES konnte mit ihrer neuesten Punktquellenanalyse eine signifikante Anzahl an Neutrinos von MGRO J1908+06 messen.

Um die Auswirkung unterschiedlicher Detektorbedingungen auf die Sensitivität der Quellen systematisch zu untersuchen wurden zusätzliche fünf Detektorbedingungen definiert, wobei in drei von ihnen die Grundlinie der Rate des optischen Untergrunds und in zwei von ihnen die Anzahl an aktiven OMs verändert wurden. Diese Parameter repräsentieren die Arbeits- und Umweltbedingungen des ANTARES Detektors. Mit den simulierten Neutrinoereignissen punktförmiger Neutrinoemission von den fünf Milagro Quellen mit Berücksichtigung dieser Detektorbedingungen wurde die Analyse wiederholt und die Sensitivität für die unterschiedlichen Detektorbedingungen bestimmt. Dies wurde für die Neutrinoflussparametrisierung, die in [4] bestimmt wurde und hier mit kappes bezeichnet wird, da diese für alle fünf Quellen zu Verfügung steht, druchgeführt. Um eine Abschätzung der Unsicherheit auf die Sensitivität für die

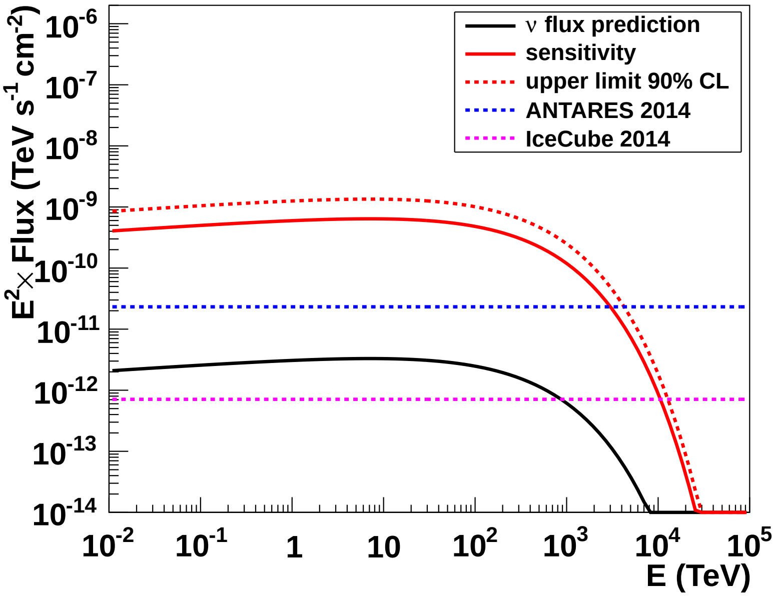


Abbildung 3: Neutrinoflussvorhersage (black), Sensitivität (rote durchgezogene Linie) für eine 3σ Entdeckung und mittlere obere Grenze 90% C.L. (rote gestrichelte Linie) für MGRO J1908+06 für die halzen3 Neutrinoflussvorhersage. Zwei 90% C.L. obere Grenzen für MGRO J1908+06 werden gezeigt, die ein E^{-2} Spektrum ohne exponentiellem Abfall annehmen für ANTARES in blau (bestimmt in [6]) und für IceCube in pink (bestimmt in [7]).

verschiedenen Detektorbedingungen abzuleiten, wurde die Tendenz der Sensitivität jeweils für die Grundlinie der Rate und für die Anzahl an aktiven OMs unabhängig approximiert und ein Intervall um den Median der Verteilung, das 68% der Runs der Run-Auswahl enthält, wurde als Basis für die Fehlerberechnung definiert. Der Fehler auf die Sensitivität wird durch die Anzahl an aktiven OMs mehr beeinflusst als durch die Grundlinie der Rate, da ihre gewählten Werte für die Standard-Detektorbedingungen den Median der Verteilung gut repräsentieren, wohingegen die Verteilung der Anzahl der aktiven OMs breiter ist. Die Sensitivität der Quellen für die kappes Neutrinoflusserwartung konnte für jede Quelle mit Unsicherheit zu

- MGRO J1908+06: $\Phi_{3\sigma} = 9.72^{+10.75}_{-2.41} \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
- MGRO J2019+37: $\Phi_{3\sigma} = 9.60^{+13.06}_{-3.79} \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
- MGRO J2031+41: $\Phi_{3\sigma} = 8.03^{+12.01}_{-4.85} \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
- C1: $\Phi_{3\sigma} = 9.15^{+12.22}_{-4.38} \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
- C2: $\Phi_{3\sigma} = 6.68^{+8.68}_{-2.39} \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$

bestimmt werden. Die relativen Fehler sind mit einer mittleren relativen oberen Unsicherheit von 131% und einer unteren Unsicherheit von 41% ziemlich groß. Deshalb ist es unabdingbar eine bessere Run-Auswahl für die rbr-Simulation punktförmiger Neutrinoquellen zu wählen oder die volle rbr-Simulation, welche mehrere CPU-Jahre pro Qulle dauert und damit sehr zeitaufwändig ist, durchzuführen bevor die Daten angesehen werden. Da die Anzahl an erwartenden Ereignissen in der Größenordnung 10^{-3} liegt und die Sensitivität für diese Quellen einen Faktor 10^2 größer als der erwartete Neutrinofluss ist, ist es unwahrscheinlich eine signifikante Anzahl an Neutrinos von diesen Quellen für eine 3σ Entdeckung zu messen. Um zukünftige Analysen nicht mit einem Trialfactor zu belegen werden die Daten nicht angesehen. In der Zukunft könnte man untersuchen, ob eine andere Kombination zusätzlicher oder anderer Schnittparameter, die bereits oben erwähnt wurden, bessere Ergebnisse liefern können. Es wäre sinnvoll eine energieabhängige Komponente in der Signal PDF aufzunehmen, welches die Anzahl an nötigen Ereignissen reduziert [8] und dieses vielleicht mit einer Stacking-Analyse-Methode kombinieren. Zusätzlich kann die Information von KrakeFit durch hinzugefügte Ereignisse, die mit AAFit rekonstruiert werden, erweitert werden, um die Anzahl an Neutrinoereignissen zu erhöhen. Diese Bemühungen könnten zu einer besseren oberen Grenze im Bereich der Neutrinoflussvorhersage führen, aber es wird mit einer hohen Wahrscheinlichkeit nicht zu einer Entdeckung von Neutrinoemission von diesen Quellen mit ANTARES führen.

In dieser Arbeit wurde gezeigt, dass ANTARES nicht in der Lage ist einen physikalisch motivierten Neutrinofluss, der aus Gammastrahlungsmessungen dieser Quellen abgeleitet wurde, und besser motiviert ist als ein generisches E^{-2} Energiespektrum nachzuweisen oder einzugrenzen. Durch Anpassung der Methode, die in dieser Arbeit benutzt wurde, ist es eventuell möglich die obere Grenze auf den Neutrinofluss zu verbessern, jedoch ist ein Nachweis der Quellen mit dem gewählten Modell mit ANTARES unwahrscheinlich.

Contents

Chapter 1

Summary

In 2013 the IceCube experiment found the first evidence for cosmic neutrino emission by detecting cascade-like high-energy neutrino events [1]. Cosmic neutrinos are produced - amongst other mechanisms - when protons, of which cosmic rays with energies above 100 GeV mainly consists, interact with photons, protons or other nuclei and produce kaons or pions. Neutral pions decay to gamma rays, charged pions decay to muons and muon-neutrinos. The muons can decay further so that again neutrinos are produced. Gamma rays can be produced also in electromagnetic processes, such as inverse Compton scattering or Bremsstrahlung, neutrinos can only be produced in the hadronic interaction process. Therefore, measuring gamma rays and neutrinos from the same source candidates could solve the questions of the unknown sources of cosmic rays and of the composition and nature of leptonic and hadronic processes within cosmic accelerators. Possible neutrino source candidates can be divided into two groups: galactic sources, for example supernova remnants (SNRs), and extragalactic sources, for example gamma ray bursts (GRBs). Milagro detected gamma ray emission from several galactic sources, so-called PeVatrons, as they are expected to accelerate cosmic particles up to several PeV. Some of these sources are located in the Cygnus star-forming region, which is a region where the star formation rate and the star death rate is high. Such regions exhibit a high density both of molecular clouds and of SNRs. Because of the higher molecular density of the molecular cloud, the probability for the interaction of protons, accelerated and ejected by the SNRs, with the atoms of the cloud is enlarged and therefore the probability of neutrino emission is enlarged compared to less dense media. A subsample of the eight galactic TeV gamma ray candidate sources measured with Milagro ([2],[3]) was studied in [4], where it is stated that neutrinos from these sources might be detected within several years of data taking with the IceCube neutrino telescope. The Milagro sources MGRO J2019+37, MGRO J2031+41, MGRO J1908+06, C1 and C2 are also observable with the ANTARES detector and are selected for the study of point-like neutrino emission with ANTARES in this thesis.

The ANTARES detector, which is optimized to detect neutrinos from the GeV up to the PeV range is located in the Mediterranean Sea and consists of 885 optical modules (OMs) placed at 12 vertical structures, so-called lines. The lines are 480 m high and are anchored with a distance of 60 to 70 m at the seabed in a depth of about 2.5 km. As neutrinos are neutrally charged and weakly interacting with matter, a large detection volume is necessary depending on the energy of the studied neutrinos and the neutrino flux. The ANTARES detector instruments a volume of 0.01 km^3 in the deep sea. An indirect detection principle is used as neutrinos can interact only in the weak-interaction processes with atoms or molecules producing highly energetic leptons in the charged-current interaction process. Here we focus on the muon-neutrino charged-current channel as for a point source analysis the good angular resolution of long trajectories within the detector is necessary. The produced muons have this property. These relativistic particles generate light via the Cherenkov-effect by moving faster than light in the transparent medium water. The transmitted light forms a light-cone under a typical angle $\Theta_c = 42^\circ$. Knowing the positions and the arrival time of photons detected by the OMs, it is possible to reconstruct the trajectory, or track, of the produced muon and therefore the produced neutrino as the space angle between them is negligible for neutrino energies above several TeV.

A new muon track reconstruction strategy has been developed in this thesis to obtain alternative or additional event information with respect to the standard algorithms. The strategy consists in a hit selection, prefit, main fit and quality cuts. The prefit is based on a scan of the whole sky in discrete angular steps. Taking each probed direction fixed, only the position of the track has to be fitted with a linear fit. For each direction an hit selection is performed using a causality criterion between hit pairs. Direct signal hits have to be preselected as optical background hits, which are caused by the decay of ^{40}K present in seawater or light emitting fauna (bioluminescence), are randomly distributed within the detector, and scattered signal hits, that deviate from the cone form, can influence the performance of the fit. Each fitted direction at the sky gets its quality value, based on the χ^2 and the number of hits used in the linear fit procedure. The ten track assumptions with the highest quality values are taken as basis for further selection criteria. Firstly, tracks that obviously do not fit the hit sample well are filtered out. The criterion used for this decision is the so-called cylinder criteria, where a cylinder around the track is calculated including all hits with low time residuals to suppress scattered hits. If the fraction of the radius r to the height h of the cylinder has a relation greater than 1, the track is filtered out. This procedure is very efficient: about 80% of tracks reconstructed with angular errors greater then 60° are filtered out, where on the other hand 98.8% of tracks reconstructed with angular errors lower then 5° are kept. With the subsample of fits fulfilling the cylinder criteria only one track is selected based on the surface density of hits around the track. To get better precision an MEstimator fit suppressing hits with large time residuals is used. The final fit is based on a maximum likelihood method using the probability density function of the time residuals of the hits. The quality parameter given by that procedure is the reduced log likelihood value, short rlogL, which is the logarithm of the likelihood divided by the number degrees of freedom. This can be used as cut parameter, the lower the rlogL, the better is the fit result. The whole fitting procedure is called KrakeFit.

In order to select only well reconstructed events and additionally to suppress the background of falsely as upgoing reconstructed atmospheric muons, possible quality cut parameters are derived for KrakeFit. The background of atmospheric muons is about a factor 10^6 higher than the number of atmospheric neutrinos for downgoing events. The performance of KrakeFit is compared to the standard ANTARES track reconstruction strategies BBFit and AAFit. The median angular error depending on the energy is shown in Fig. 1.1. BBFit, which is most efficient for low-energetic neutrino events, has a better performance than KrakeFit for energies below 50 GeV. For high-energy neutrino events, the performance of KrakeFit and AAFit, which is the standard high-resolution track reconstruction strategy of ANTARES, are comparable after applying cuts for which both give comparable numbers of misreconstructed atmospheric muons. A comparison of the events reconstructed with KrakeFit and AAFit, applying quality cuts, gives that 35.9% of all triggered events were reconstructed with both strategies, 20.8% were only reconstructed with AAFit and 5.6% were additionally reconstructed with KrakeFit. So KrakeFit gives additional neutrino event information to AAFit and can be used to enhance the amount of neutrinos. The mean of the space angle between the results of the two reconstruction strategies

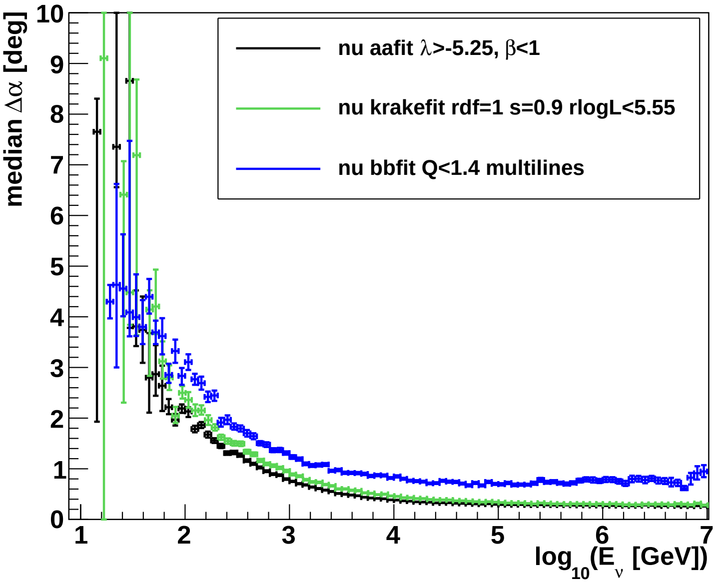


Figure 1.1: Median angular error $\Delta\alpha$ for upgoing neutrino events as a function of the neutrino energy for AAFit ($\lambda > -5.25$, $\beta < 1°$) (black), BBFit ($Q < 1.4$, NLines > 1) (blue) and KrakeFit (rdf=1, $s > 0.9$, rlogL < 5.55) (green) for an E^{-2} energy spectrum.

for events reconstructed with both as upgoing is 1.4°.

KrakeFit is a reconstruction strategy which has a high angular resolution - for neutrinos with energies above 1 TeV (10 TeV) the resolution is better then 0.88° (0.45°) - and a good efficiency for the reconstruction of (high-energy) neutrinos. To inspect if the event reconstruction works as expected and if the run-wise-based (rbr) Monte Carlo simulation (MC) describes the data well, a data-MC-comparison was performed. Here general observables of KrakeFit as the distribution of for example the zenith, azimuth, number of hits and number of lines used in the fit were studied for a run-selection. This includes ANTARES runs that were recorded from 2008 to 2012 with the full-detector configuration and correspond to a live-time of 744.68 days. After applying quality cuts, the main one on the rlogL value, a good agreement between data and MC is observed. In addition to the standard observables further parameters to distinguish between atmospheric muons and neutrinos are studied. Some of them are good candidates for cut parameters. These are the length of direct hits (determined as hits with low time residuals to the fitted track) projected on the track, the safety parameter of the rdf (random decision forest up-down classifier) and the space angle between KrakeFit- and AAFit- reconstructed events. These have to be studied and optimized, maybe also in combination, for each analysis separately as each analysis has different requirements.

The KrakeFit track reconstruction strategy is now applied to a study of the Milagro sources. To limit the parameter space to a feasible number of parameters for the study of the sensitivity of point-like high-energy neutrino emission from the Milagro source candidates, only the rlogL value and an additional cut on the rdf=1 (rdf=1 for upgoing tracks) was studied. The optimal cut values are determined for each Milagro source candidate and their neutrino flux predictions separately with respect to the best MDP (model discovery potential) for a 3σ discovery. In [4] and [5] the neutrino flux predictions of MGRO J2019+37, MGRO J2031+41, MGRO J1908+06, C1 and C2 are derived from measurements of their gamma ray spectrum and is predicted following a power law with exponential cut-off:

$$\frac{dN_\nu(E_\nu)}{dE_\nu} = k_\nu \left(\frac{E_\nu}{\mathrm{TeV}}\right)^{-\alpha_\nu} \exp\left(-\sqrt{\frac{E_\nu}{E_{cut,\nu}}}\right). \tag{1.1}$$

For each source and neutrino flux assumption a point-like neutrino simulation using the rbr simulation scheme was performed. As it takes several CPU-years to produce a rbr simulation for each of the selected runs for one source, a subsample of runs is used instead which fulfil a standard detector condition defined to represent the majority of runs. With these simulations the signal PDF (probability density function) was created, which is one input for the maximum-likelihood-based point source analysis method. As signal PDF S_i the spacing distribution to the source of upgoing reconstructed signal events after applying quality cuts is taken. This is the so-called PSF (point spread function). As background PDF B_i the event rate per $\sin(\delta)$ of declination δ is taken after applying quality cuts. Both are normalized so that the integral over the solid angle for each is equal to 1. A maximum-likelihood-based method is needed as after measuring a trajectory reconstructed as upgoing one cannot distinguish between cosmic neutrinos, atmospheric neutrinos and misreconstructed atmospheric muons. Therefore a statistical approach has to be used to interpret how signal- or respectively background-like the result is. This is done by hypotheses testing where the null hypothesis is that the measured data only consists of background events. The method of maximizing the test statistic Q is used instead of maximizing the likelihood L, which are defined as

$$\log L(n_s) = \sum_{i=0}^{N} \log\left(\frac{n_s}{N} S_i + \left(1 - \frac{n_s}{N}\right) B_i\right) \tag{1.2}$$

and

$$Q = \log\left(\frac{\max(L(n_s))}{L(n_s = 0)}\right) = \max(\log L(n_s) - \log L(n_s = 0)) \tag{1.3}$$

respectively, where N is the total number of events in data after applying quality cuts and n_s is the number of assumed signal events. The value μ_s, which is the value of n_s that maximizes the likelihood or respectively the test statistic, is the best estimate for the true number of signal events. Pseudo-experiments are generated by generating only background and background with present signal event samples. The outcome is interpreted statistically. From the distribution of Q for the background-only distribution the threshold of Q for a 3σ and a 5σ discovery can be determined. It is shown that the mean fitted value of μ_s corresponds to the number of injected events and thus the hypotheses test can be performed. The best cut values (rlogL value with and without rdf=1) are optimized due to the model discovery potential (MDP) to get the best probability for discovery for the expected number of events (determined from the point source simulation per source and neutrino flux assumption). The corresponding sensitivity is determined by multiplying the expected neutrino flux by the factor of the number of events necessary for a 50% chance of a 3σ discovery $n_{3\sigma}$ divided by the number of expected events. An average upper limit for 90% confidence level (C.L.) $\bar{\Phi}_{90}$ on the flux is set. The focus lies on the values determined for 3σ optimization as the number of expected events is in the order of 10^{-3}.

As there are more then one flux assumption per source available for the sources MGRO J1908+06, MGRO J2019+37 and MGRO J2031+41 the final results given in Tab. 1.1 are taken

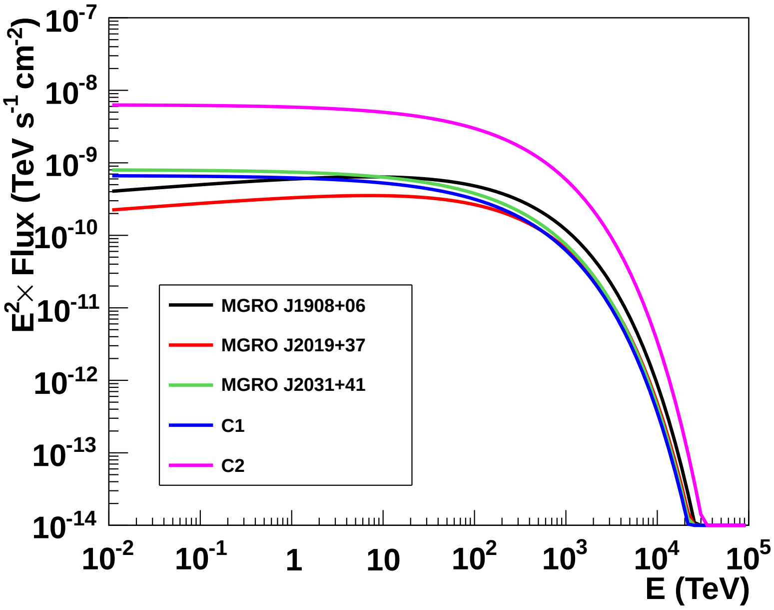


Figure 1.2: Sensitivity of the five Milagro sources for a 3σ discovery for the flux assumptions and optimized cuts given in Tab. 1.1.

from the cuts and neutrino flux assumptions with the highest MDP per source. In Fig. 1.2 the sensitivities for a 3σ discovery for the five Milagro sources is shown. The sensitivity is for all sources about a factor 10^2 higher then their expected neutrino flux. It is unlikely to measure a significant amount of neutrinos for a 3σ discovery. This is caused by the fact that the visibility of these sources within the ANTARES detector is low, it is about 23% to 30%, except for MGRO J1908+06 with a visibility of 48%, and by the fact that the expected energy spectrum has an exponential cut-off, so the neutrino events will be mostly low-energetic and therefore disappear in the atmospheric background.

For MGRO J1908+06 the neutrino flux assumption, the sensitivity and the 90% C.L. average upper limit determined in this thesis is shown in Fig. 1.3 in comparison to the 90% C.L. upper

source	flux	rlogL	rdf	$n_{3\sigma}$	$\Phi_{3\sigma}$ (TeV^{-1}cm^{-2}s^{-1})	$\bar{\Phi}_{90}$ (TeV^{-1}cm^{-2}s^{-1})
MGRO J1908+06	halzen3	5.1	-	1.81	$6.3 \cdot 10^{-10}$	$1.35 \cdot 10^{-9}$
MGRO J2019+37	halzen3	5.2	1	0.93	$3.5 \cdot 10^{-10}$	$9.70 \cdot 10^{-10}$
MGRO J2031+41	kappes	5.3	-	1.07	$8.0 \cdot 10^{-10}$	$2.20 \cdot 10^{-9}$
C1	kappes	5.2	1	1.01	$8.7 \cdot 10^{-10}$	$2.28 \cdot 10^{-9}$
C2	kappes	5.2	1	0.94	$6.3 \cdot 10^{-10}$	$1.76 \cdot 10^{-9}$

Table 1.1: Summary of the flux assumptions and optimized cuts for the best 3σ discovery potential per source. The number of $n_{3\sigma}$, the normalization of the sensitivity $\Phi_{3\sigma}$ and the average upper limit for 90% C.L. $\bar{\Phi}_{90}$ are also shown.

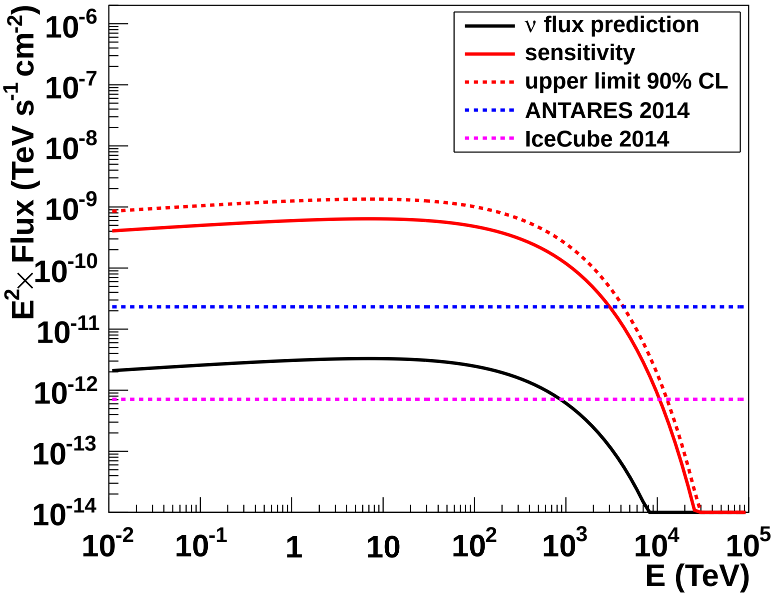


Figure 1.3: Neutrino flux assumption (black), sensitivity (red continuous line) for a 3σ discovery and average upper limit 90% C.L. (red dashed line) for MGRO J1908+06 for the halzen3 neutrino flux assumption. Two 90% C.L. upper limits are shown assuming an E^{-2} energy spectrum without cut-off energy determined for MGRO J1908+06 for ANTARES in blue (derived in [6]) and for IceCube in pink (derived in [7]).

limits determined by ANTARES [6] and IceCube [7] for this source, the latter two assume an E^{-2} energy spectrum without cut-off energy. The official ANTARES analysis considers a larger run selection with a live-time 1.8 times larger than considered in this thesis, and the assumed neutrino flux spectrum without cut-off energy gives an higher amount of expected neutrinos and therefore a better upper limit. It has to be mentioned that the assumed E^{-2} neutrino flux without cut-off is not physically motivated for theses sources. Further the energy range of the majority of events expected from that flux is not stated. The energy range of IceCube for northern hemisphere sources is between 1 TeV and 1 PeV. Neither IceCube nor ANTARES have seen any significant amount of neutrinos from MGRO J1908+06 in their latest point source analysis of this source.

To systematically study the effect of different detector conditions on the sensitivity, additionally five detector conditions are defined where in three of them the baseline rate of the optical background and in two of them the number of active OMs is changed being representatives for the working and environmental conditions of the ANTARES detector. With the simulation of point-like neutrino emission of the five Milagro sources applying these detector conditions and the neutrino flux parametrization of [4], named kappes, as this is available for all five Milagro sources, the analysis was repeated and the sensitivity for these conditions is determined. To derive an estimate of the uncertainty on the sensitivity due to different detector conditions an approximation of the trend of the sensitivity is made for the baseline rate and number of active OMs separately and an interval which contains 68% of all runs around the median of the distributions is taken as basis for the error calculation. The error on the sensitivity is more effected by the number of active OMs than the baseline rate, as the chosen value for standard detector conditions represents the median of the distribution well there, whereas the spread of the number of active OMs is wider. The sensitivity of the sources for the kappes neutrino flux assumption for each source with uncertainties could be determined to

- MGRO J1908+06: $\Phi_{3\sigma} = 9.72^{+10.75}_{-2.41} \cdot 10^{-10}\,\mathrm{TeV^{-1}cm^{-2}s^{-1}}$
- MGRO J2019+37: $\Phi_{3\sigma} = 9.60^{+13.06}_{-3.79} \cdot 10^{-10}\,\mathrm{TeV^{-1}cm^{-2}s^{-1}}$
- MGRO J2031+41: $\Phi_{3\sigma} = 8.03^{+12.01}_{-4.85} \cdot 10^{-10}\,\mathrm{TeV^{-1}cm^{-2}s^{-1}}$
- C1: $\Phi_{3\sigma} = 9.15^{+12.22}_{-4.38} \cdot 10^{-10}\,\mathrm{TeV^{-1}cm^{-2}s^{-1}}$
- C2: $\Phi_{3\sigma} = 6.68^{+8.68}_{-2.39} \cdot 10^{-10}\,\mathrm{TeV^{-1}cm^{-2}s^{-1}}$.

The relative errors are with a mean relative upper uncertainty of 131% and lower uncertainty of 41% quite high. So it is indispensable to take a better selection of runs for the point-like rbr simulation process or do the fully rbr simulation, which takes several CPU-years per source and is very time consuming, before unblinding the data. As the number of expected events is of the order of 10^{-3} and the sensitivity for these sources is about a factor of 10^2 higher than the expected neutrino flux, it is unlikely to measure a significant amount of neutrinos from these sources for a 3σ discovery with ANTARES. To not constrain further analyses with a trial factor the data are not unblinded. In the future it could be studied if another combination of additional or different cut parameters mentioned above could give better results. It would be useful adding an energy-dependent component in the signal PDF, which was shown to lower the number of necessary events [8] maybe combined with a stacking analysis approach and add to the information of KrakeFit events reconstructed with AAFit to enhance the amount of neutrinos. These efforts could lead to a better upper limit in the area of the neutrino flux prediction, but it won't lead to high probability to a discovery of neutrino emission from these sources with ANTARES.

In this thesis it is shown that ANTARES is not able to detect or constrain physically motivated neutrino flux predictions, derived from gamma ray measurements of these sources, which are better motivated than an E^{-2} energy spectrum. By adjusting the method used in this thesis, the upper limit could be improved, but a detection of these Milagro sources with the selected model with ANTARES is unlikely.

Chapter 2

High-energy neutrino sources

When talking about neutrino astronomy, neutrinos with energies in the range of several GeV up to PeV are studied which are assumed being of cosmic origin. IceCube detected cascade-like neutrino events, which is the first evidence of a diffuse cosmic neutrino flux, but they could not identify point-like neutrino sources up to now [1].

In this thesis cosmic neutrino emission from point-like sources with the ANTARES detector are studied. Therefore high-energy track-like muon-neutrino events are selected and analysed. When talking about high-energy neutrino events, neutrinos with energies above several TeV are meant. In this chapter the purpose of neutrino astronomy is motivated starting with the composition and measured spectrum of cosmic rays. Some potentially cosmic neutrino source candidates are shown grouped in galactic and extragalactic sources. Possible cosmic ray acceleration processes and neutrino flux predictions are described with respect to the candidate source selection of this thesis.

2.1 Cosmic rays

The primary cosmic rays[1] (for energies above 100 GeV) consist of charged particles mainly of hadrons (ionized atoms) and only a tiny amount of electrons [9]. Their composition is splitted into ca. 85% protons, 12% α-particles and ca. 3% heavier nuclei [10]. A lot of experiments measured the cosmic ray composition and fluxes depending on energy in direct measurements like satellite or balloon measurements and for energies above 100 TeV in measurements of their secondary particles in air shower experiments. The flux is summarized in Fig. 2.1. The flux is multiplied with energy to make structures visible. The points where the spectral index γ of the energy spectrum ($dN/dE \propto E^{-\gamma}$) changes are called knee, second knee and ankle. For energies up to the PeV range $\gamma = 2.7$, than it steepens to $\gamma = 3.1$ (knee), and it steepens again at the second knee and flattens again at the ankle [11]. As the cosmic ray flux decreases rapidly with energy huge detectors are necessary to measure a significant amount of them within several years for large energies.

To study the mechanisms of acceleration and learn about the sources of cosmic rays, charged particles are not a good messenger as magnetic fields can deflect them during their flight to Earth. A possible method is using electrically neutral particles as gamma rays and neutrinos.

---

[1] Primary cosmic rays are particles originally produced by the cosmic ray source, whereas secondary cosmic rays are particles produced by the interaction of primary cosmic rays with molecules or atoms of the interstellar medium or the atmosphere of the Earth.

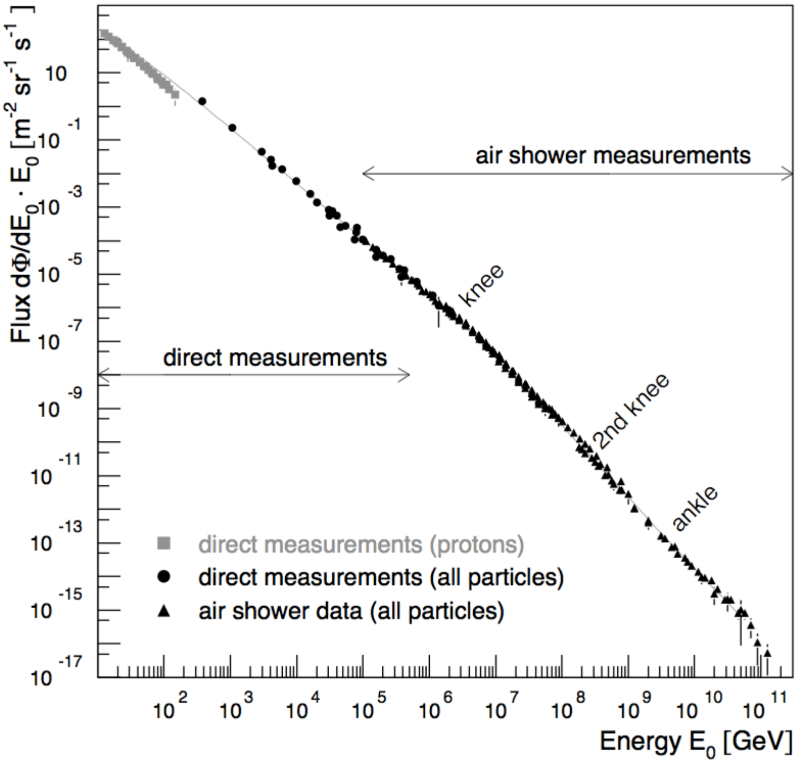


Figure 2.1: Energy spectrum of primary cosmic ray particles measured directly with detectors above the atmosphere and with air shower detectors. Figure taken from [11].

These are produced in a sequence of reactions almost exclusively starting with accelerated cosmic ray protons which can interact with protons, gammas or other nuclei producing hadrons, mostly neutral and charged pions. The neutral pions decay rapidly with a lifetime of $8.4 \cdot 10^{-17}$ s almost exclusively into two gammas [10]

$$\pi^0 \to \gamma + \gamma. \tag{2.1}$$

The charged pion decays into muons and muon-neutrinos with a lifetime of 26 ns [10]

$$\pi^+ \to \mu^+ + \nu_\mu, \quad \pi^- \to \mu^- + \bar{\nu}_\mu. \tag{2.2}$$

In a following process the produced muons can decay to electrons and electron-neutrinos:

$$\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu, \quad \mu^- \to e^- + \bar{\nu}_e + \nu_\mu. \tag{2.3}$$

While gamma rays can also be produced in electromagnetic processes as inverse Compton scattering or Bremsstrahlung, neutrinos can only be produced in the hadronic case. Therefore measuring both neutrinos and gamma rays from cosmic sources solves the problem of the unknown composition of leptonic and hadronic processes of cosmic accelerators. Additionally to their electrically neutral property the small interaction cross-section of neutrinos leads to the fact that they are not absorbed by molecular dust or interstellar media and that they can even pass the Earth without interacting (for energies below 1 PeV). In summary they point back to their sources and can also be emitted from inner parts of the sources which allows to study also inner effects. This nature is used by high-energy neutrino observatories like ANTARES [12] and IceCube [13].

2.2 Cosmic neutrino sources

Possible cosmic neutrino sources can be divided into two groups: galactic sources and extra-galactic sources. A short overview of source candidate classes from both groups are given here. More information about neutrino astronomy and its link to study also particle physics and exotic phenomena is given in the review [9].

2.2.1 Galactic sources

Galactic sources are sources that are located within our galaxy. They are assumed being the primary source for particles with energies up to the knee in the cosmic ray energy spectrum (see Fig. 2.1). The source classes include supernova remnants (SNRs), pulsar wind nebular (PWNs) and star-forming regions.

SNRs are the rest of the stars from supernova explosions. Fermi postulated a mechanism called shock acceleration or first order Fermi acceleration that is assumed to accelerate hadrons in those sources. It is based on the fact that charged particles gain energy by multiple scattering between two shock fronts formed by the ejection from the supernova explosion [10]. Taking this as basis one can derive that the energy spectrum for neutrinos from such sources follows roughly an E^{-2} power law [9].

The class pulsar wind nebula is a special type of supernova remnant, where a pulsar is present and ejecting material in a nebula. The most studied candidate is the crab nebula.

Molecular clouds and magnetic clouds are possible cosmic ray accelerators following Fermi second order acceleration also known as Fermi mechanism [10]. If cosmic ray particles interact with gas clouds moving with a certain velocity $\vec{u}$, they can gain energy by multiple reflection with the cloud. Depending on the relative direction of motion of the particle and the cloud, the particle can gain or loose some energy per reflection process. On average the energy gain is proportional to u^2. As the velocity of such clouds is small, it takes long time for a certain energy gain. Therefore magnetic clouds are assumed being the main collision partners as the interaction probability is larger because of the higher gas density [10].

The detection probability of neutrinos from sources located in regions, where for example a SNR is combined with dense media as molecular clouds, is expected to be enlarged as the probability for particle interaction in such regions (and therefore neutrino production) is higher then in less dense media. This happens for example in so-called star-forming regions, where the star-formation rate is enlarged. 2007 the Milagro experiment measured gamma ray emission from galactic sources, where some of them are located in the Cygnus star-forming region [3]. They are assumed being PeVatrons, accelerators of cosmic rays up to energies of several PeV [14], that produce neutrinos with an energy spectra up to several TeV. Those Milagro sources that can be detected with the ANTARES neutrino telescope have been selected for this study (see chapter 6).

Prediction of neutrino fluxes for galactic sources

As no direct primary cosmic ray measurements are available for the studied sources, neutrino fluxes can be predicted using high-energy photon measurements of the respective sources. This is based on the fact that in hadronic processes neutrino and gamma ray fluxes are correlated as the energy is distributed evenly between the three different types of pions in hadronic collisions [15]. The measured gamma ray flux is transformed using models of pion production in proton-proton collisions and models of their secondary particles to estimate the proton flux of the sources and transformed back to estimate the neutrino fluxes. This is done for example in [16] and [15]. This procedure is complex and underlies various assumptions and uncertainties. If no gamma ray flux measurements and therefore no neutrino flux predictions are available for the source candidates under study, an overall E^{-2} energy spectrum is used motivated by Fermi first order acceleration. Because of this and because of the fact that there are many different theoretical models of the neutrino flux predicted per source and per source class, an overall E^{-2} spectrum is used for comparison purposes and to optimize the direction reconstruction strategies for cosmic neutrinos.

2.2.2 Extragalactic sources

Extragalactic sources are assumed to constitute the primary component of cosmic rays beyond the ankle in the cosmic ray spectrum. Candidate source classes are gamma ray bursts (GRBs), active galactic nuclei (AGNs) and starburst galaxies. The predictions of their neutrino flux show a wider range. They are modeled according to different theoretical models as it is done for GRBs for example in [17].

GRBs are short gamma ray emissions that occur suddenly and without prediction with a duration of seconds (short gamma ray bursts) up to minutes (long gamma ray bursts) [10]. They are studied for example in [17] and [18], but no evidence for neutrinos from such sources could be found with significance.

AGNs consist of a super massive black hole in their center surrounded by a rotating accretion disc which transports material into the center. Perpendicular to the accretion disc jets are formed from particles accelerated near the black hole or the nucleus of a compact galaxy and injected into the radiation field of the source [10]. Shock fronts, formed in the jet, are supposed to accelerate protons and electrons according to Fermi shock acceleration, which initiate hadronic respectively electromagnetic cascades, producing neutrinos and gamma rays according to the processes mentioned above. If a neutrino telescope looks into the direction of the jet, neutrinos from such AGNs could be observed. This is studied in [19] for ANTARES, but no evidence for neutrino emission from these sources could be observed.

Starburst galaxies are galaxies where the rate of star-formation and also death is high. The high rate of supernova explosions in their center leads to a "galactic-scale wind" [9]. According to the fact that supernova remnants are possible neutrino source candidates, as already discussed in the section of galactic sources, this high concentration of supernova explosions and their resulting cosmic ray particle rate together with a high matter density, are good reasons regarding them as possible neutrino sources, which is not studied so far.

There are no dedicated neutrino flux predictions available for starburst galaxies in contrast to the Milagro sources, where the neutrino flux assumption is derived from the measured gamma ray spectrum for each source separately for example in [4] and [5] (as can be seen in section 6.2). Instead of starburst galaxies these sources are located in the galactic plane which is one of the main visible parts of the sky observed by ANTARES. As stated in [4], IceCube might detect neutrinos from these sources with significance within several years of data taking. So a complementary analysis with ANTARES data is useful as this is not done so far. Therefore those Milagro sources visible with ANTARES are chosen for this study.

In the following an overview of the detection principle of high-energy neutrinos with the ANTARES detector is given. Further a new muon track reconstruction algorithm is developed where an overall E^{-2} spectrum is used to optimize the strategy and compare it with other track reconstruction algorithms (see chapter 4). A dedicated data and Monte Carlo simulation (MC) comparison of main reconstruction observables is done in chapter 5 to ensure that the fit works as expected and that the observables of the fit for data are described well by them of the MC. With this new strategy a point-like neutrino source search for the selected Milagro sources is performed using an maximum likelihood-based analysis method (see chapter 7), where the results for the different neutrino flux assumptions per Milagro source are discussed in chapter 8.

Chapter 3

Neutrino detection with the ANTARES detector

Neutrinos can be measured indirectly by their weak interaction products producing light in transparent media like water or ice using the Cherenkov-effect. The technical requirements for using this technique depends on the studied energy range as for example the expected neutrino flux decreases exponentially with energy. Measuring solar neutrinos with energies of several MeV can be done by underground experiments (like Super-Kamiokande [20] or SNO [21]). To measure neutrinos with energies above several GeV up to PeV large scale neutrino detectors are necessary. ANTARES, for example, with an energy threshold of 10 GeV, has an instrumented volume of about 0.01 km^3. Those detectors can only be build in natural water or ice reservoirs as for example the Mediterranean Sea for ANTARES or the South Pole for IceCube. In this chapter the detection principles with the neutrino interaction processes, the Cherenkov-effect and the light propagation are described, where the focus lies on neutrino detectors using water as transparent media as ANTARES data are analysed in this thesis. Further it concentrates on the measurement of muon-neutrinos with energies above 1 TeV producing muons interacting in the charged-current reaction, as those events are regarded for the point source analysis later on. An overview of the ANTARES detector layout and data taking and the influence of environmental conditions are given.

3.1 Detection principle

3.1.1 Neutrino interaction processes

Neutrinos ν_l(or anti-neutrinos $\bar{\nu}_\mathrm{l}$) can only interact via the weak interaction in so-called charged-current or neutral-current interactions with nucleons N by exchanging an Z- or $W^{\pm}$-boson [22]. The following scheme shows these interactions where l is a lepton (e,μ,τ) and X is the hadronic end state.

- Neutral-current reaction (NC):

$$\nu_\mathrm{l} + \mathrm{N} \rightarrow \nu_\mathrm{l} + \mathrm{X} \tag{3.1}$$

$$\bar{\nu}_\mathrm{l} + \mathrm{N} \rightarrow \bar{\nu}_\mathrm{l} + \mathrm{X} \tag{3.2}$$

- Charged-current reaction (CC):

$$\nu_{\mathrm{l}} + \mathrm{N} \rightarrow \mathrm{l}^{-} + \mathrm{X} \tag{3.3}$$

$$\bar{\nu}_{\mathrm{l}} + \mathrm{N} \rightarrow \mathrm{l}^{+} + \mathrm{X}. \tag{3.4}$$

Both reactions end up in shower events which are studied in detail for example in [23], whose precision of the direction of the shower is not as good as for the produced lepton in the CC-reaction. As for the detection of point-like neutrino sources the neutrino direction has to be known with high precision, this thesis concentrates on the charged-current interaction, which produces high-energy muon events (see section 3.1.4) as for neutrino astronomy purposes the direction of the neutrino has to be known with high precision. The mean space angle β between neutrino and produced muon direction is given by [24]

$$\bar{\beta} = \frac{0.7^{\circ}}{(E_{\nu}/\mathrm{TeV})^{0.6}}. \tag{3.5}$$

As the angle between neutrino and lepton direction decreases with neutrino energy, for energies above several TeV it is negligible. Measuring the direction of the lepton, for energies above several TeV, can therefore be used to determine the direction of the initial neutrino.

3.1.2 Cherenkov-effect

The measurement of the direction of this lepton depends on the measurement of the Cherenkov-light produced when charged particles move in a transparent media faster than the light in this media. This effect is known as Cherenkov-effect [25]. The radiation is emitted with an angle Θ_c typical for the transparent media for example water or ice. For the relevant energies $\Theta_c = 42^{\circ}$ for sea water with an refrective index of $n = 1.35$ and a wavelength of 450 nm [24]. Fig. 3.1 shows a drawing of the emission of Cherenkov-light for a muon produced in a CC-interaction of a muon-neutrino, where the light front forms a cone with opening angle $\gamma = 90^{\circ} - \Theta_c$. As the light propagation has this geometrical dependency relative to the direction of motion, it is possible to reconstruct the direction of the lepton by measuring the position and arrival time of single photons of the light cone detected at photon detection units, which are illustrated as grey circles in Fig. 3.1.

3.1.3 Propagation of light in seawater

The emitted Cherenkov-light can be scattered or absorbed at the molecules of water or ice. So the intensity of light is reduced by processes which are described by the absorption length λ_{abs} and the scattering length λ_{scat}. These are summarized in the attenuation length λ_{att}, which is defined as

$$\frac{1}{\lambda_{\mathrm{att}}} = \frac{1}{\lambda_{\mathrm{abs}}} + \frac{1}{\lambda_{\mathrm{scat}}} \tag{3.6}$$

and describes the length where the intensity $I(d)$ is reduced to $1/e$ of the original intensity I_0. The attenuation effect of light in seawater is determined to be smallest in the regime of blue light at wavelengths between 460 nm and 470 nm [26]. The absorption length, determined by a data to Monte Carlo simulation (MC) comparison for data collected in in situ measurements of ANTARES, is in the range between 55 and 65 m [24].

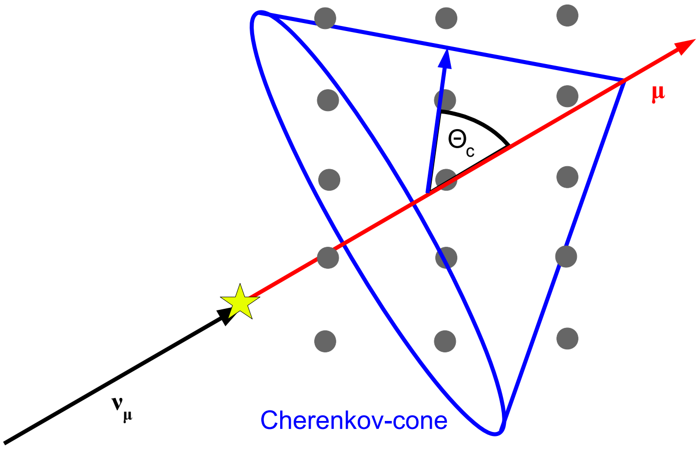


Figure 3.1: Scheme of the Cherenkov-cone shown for a muon μ produced in a CC-interaction of a muon-neutrino ν_μ. The photon detection units are illustrated as grey circles.

The number of detected photons is a measure of the intensity of light and can be calculated in dependence of the distance to the emission point d as

$$I(d) = I_0 \cdot \exp\left(-\frac{d}{\lambda_{\text{att}}}\right). \tag{3.7}$$

Therefore the number of photons is reduced and their time residual, which is the time difference between the theoretical expected time assuming no scattering and the measured time, is higher for larger distances. These two points have to be kept in mind for the reconstruction of the lepton direction (see chapter 4).

3.1.4 Types of neutrinos

Three different neutrino flavours (ν_e, ν_μ, ν_τ) exist. All of them can interact via the CC-channel and produce their corresponding lepton which can in principle be detected by the ANTARES detector using the emission of Cherenkov-radiation. For the following reasons, the main neutrino flavour this thesis concentrates on is the muon-neutrino. Electron-neutrinos produce electrons which have a short absorption length and produce a short cascade in water, which means the detection volume is small (for an energy of 1 TeV the path length is ca. 5 m see Fig. 3.2). They can be used for electromagnetic-shower studies. Tau-neutrinos which produce taus have a short lifetime and transform to muons and muon-neutrinos. These cannot be distinguished from muons produced in the CC-interaction of muon-neutrinos and represent background for this reaction. For large-scale neutrino detectors (as for example IceCube [13]) it might be possible to discriminate between tau-neutrinos and muon-neutrinos as in the tau-neutrino case, they might measure two interactions in their detector caused by the tau-neutrino interaction and the interaction of the produced muon-neutrino within their detection volume. For a neutrino detector with a size of 0.01 km^3 as ANTARES this kind of study is nearly impossible.

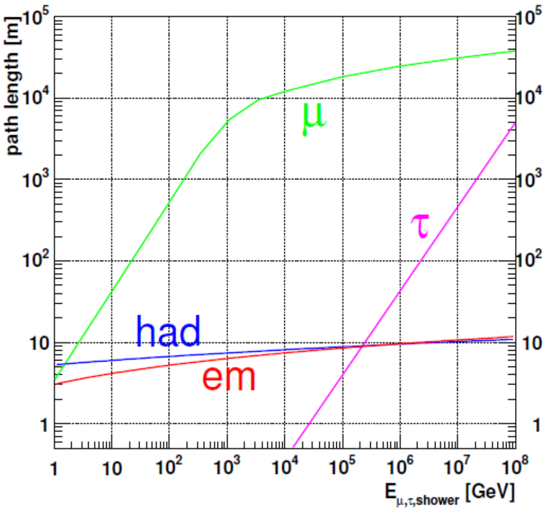


Figure 3.2: Path length in water for muons (green), taus (pink) and electromagnetic (red) and hadronic (blue) showers. Figure taken from [26].

The advantage of muon-neutrinos is that the produced muon can travel kilometers straight through water (see Fig. 3.2) which is good for two reasons. The first is that most muon trajectories have long paths inside the detector, which makes an accurate direction reconstruction possible. And secondly the point of interaction can be far away outside the detector and the muon could still reach it. The detection volume is therefore much larger then the geometric volume of the detector. For the search of point-like neutrino sources, where the direction of the neutrino has to be known with high precision, muon-neutrinos interacting in the CC-channel are the 'golden channel'.

3.2 The ANTARES detector

3.2.1 Detector layout

The ANTARES neutrino telescope is located in the Mediterranean Sea at a water depth of 2475 m ca. 20 km from the coast of south France near Toulon (see Fig. 3.3). It consists of a three-dimensional matrix of 885 photon detection units, the so-called optical modules (OMs), which consist of a pressure resistant glass sphere with a photomultiplier tube (PMT) inside. Always three of these OMs are grouped into a storey with a space angle of 120° between them, with PMTs looking downwards with an angle of 45° to be more sensitive to upgoing neutrino tracks and in order to suppress the atmospheric muon background from above (see section 3.2.4). Fig. 3.4 shows a schematically view of the detector with a picture of one storey. The three-dimensional matrix is fixed by 12 vertical lines, 11 of them containing 25 storeys, one containing 20 storeys and 3 acoustic detection units[1]. The lines are horizontally separated by 60 to 70 m at the seabed and are hold straight by an underwater buoy on top of each line. One line

---

[1] As part of the ANTARES detector an acoustic neutrino detection test system called AMADEUS is installed, see [27] for details.

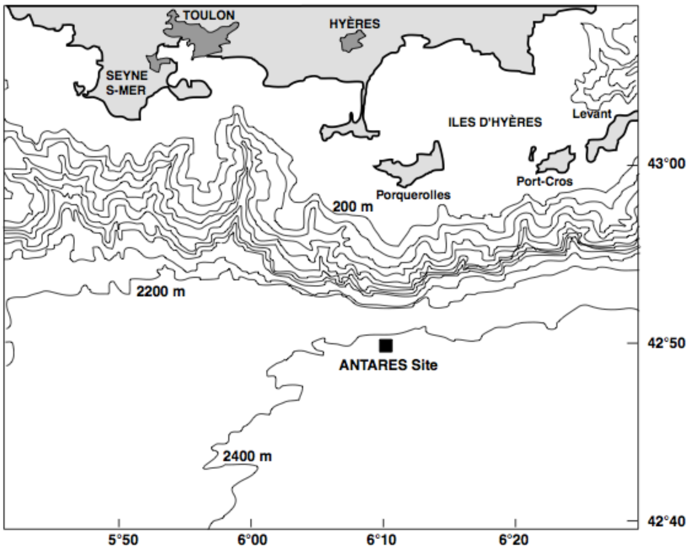


Figure 3.3: Location of the ANTARES detector near the coast of south France in a water depth of 2475 m.

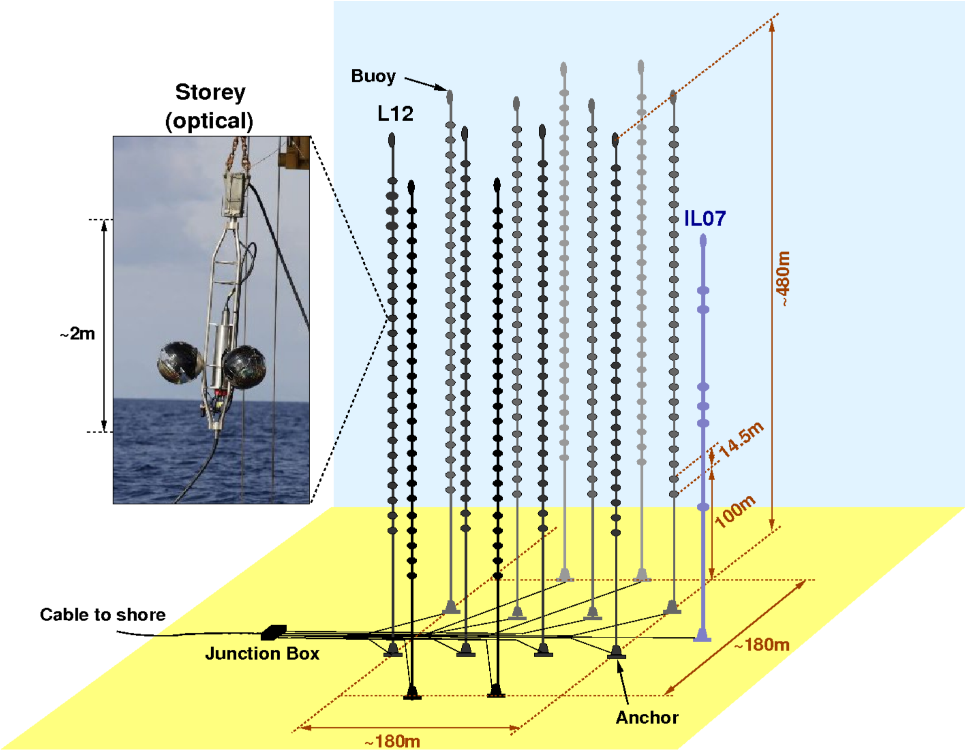


Figure 3.4: A schematical drawing of the ANTARES detector with a picture of one storey.

has a length of ca. 480 m on which 25 storeys are placed at a distance of 14.5 m along the line, the first 100 m above the seabed are not instrumented. In total the detector has an instrumented volume of 0.01 km^3. A detailed description can be found in [28].

3.2.2 Data acquisition and triggering

The measured signals at the PMTs are digitized already at the sea and sent to shore following the all-data-to-shore principle for further processing. The details of these are described in the following.

The electronics of the storey are installed in the LCM (Local Control Module) in the middle between the three OMs. The lines get the power from the 'Junction Box' over an electro-optical cable which has a length of ca. 40 km from the coast station. This cable has also the function of data transmission from the detector to the coast station, where the data are processed and filtered, and sending control data to the detector.

The PMT readout has a fixed threshold which is set to ca. 0.3 pe (photo electrons). If light hits the PMT and its pulse height exceeds the threshold, the pulse is integrated within the integration time of 33 ns[2] [29]. Each PMT has two ARSs (Analogue Ring Sampler) for the readout, which is needed to decrease the acquisition dead time. After the integration time of the first ARS its dead times starts and the second ARS takes over with a short time delay. The ARS transforms the analogue signal to a digital signal by integrating the charge within the integration time, which is called hit. Each hit is stored with a detected arrival time, charge and OM identifier. The charge is correlated with the number of detected photoelectrons whereas the time is synchronized over the whole detector with a precision of 1 ns. All these continuously collected data are send to shore, where a computer farm with filter algorithms (triggers) processes the data current in real time. If the trigger condition of one of the triggers is fulfilled, all data detected within the time window $[t_f - 2.2\,\mathrm{ns}, t_l + 2.2\,\mathrm{ns}]$ are stored on tape as one event, where t_f is the time of the first triggered hit and t_l is the time of the last triggered hit. To run these triggering algorithms in real time a simplified detector configuration of totally straight lines and a non-final calibration, so-called online calibration, is assumed. So the selected data have to be reprocessed with the correct detector geometry and calibration values offline for further analyses as those values can only be determined later on.

Different trigger algorithms are active onshore, they reduce the amount of data by a factor of about 10^4 [28]. The triggers can be divided into different groups: general muon triggers, directional muon triggers for special research purposes, external alert and minimum bias triggers [28]. The general muon triggers are based on a causality criterion (see section 4.1.1) of highly or locally clustered hits or on the coincidence of these hits in next or next-to-next storeys. Some triggeres that are considered for multi-messenger-approach take the position of the source into account. The minimum bias trigger is used to monitor the current detector status and the data quality conditions.

In this analysis only data recorded with general muon triggers are used. As those are independent of the muon direction, the causality of all possible hit combinations is checked. To reduce the computing time for this procedure a preselection of hits is necessary. The hits with either a high charge (typical greater then 3 pe) or hits that are detected on the same storey within a time difference of 20 ns are selected. These hits are called L1 hits (level 1 hits) which is the first trigger level and reduces random background hits substantially. In the case of several coincident hits only the first hit is taken. The general trigger algorithms run with these L1 hits.

---

[2] The integration time is set to 25 ns after crossing the threshold plus 8 ns before.

If the trigger criterion is fulfilled the corresponding hits are the so-called L2 hits (level 2 hits) and an event containing all triggered hits as well as a snapshot of all hits within the time window mentioned above are recorded.

3.2.3 Influence of environmental conditions

As the Mediterranean Sea is a living environment whose environmental conditions can change in time scales in the order of hours up to seasonal changings, the detector and its performance is influenced by these conditions.

Detector geometry

To determine the muon trajectory with high precision the timing resolution must be in the order of a few nanoseconds and the position of each OM has to be known with a precision of a few tens of centimeters. As the detector lines are flexible (only hold straight by a buoy and fixed at the seabed) they can bend correlated with the velocity of the sea current, so that they differ from the geometry of a straight line. In addition the storeys can rotate around the cable of the line. To determine the actual position and orientation of each OM a hybrid positioning system is used. The acoustic positioning system consists of emitters at the seabed and at the anchor of each line and hydrophones on every fifth storey of each line that detect the acoustic signals. Through trilateration the position of each storey can be reconstructed with a precision better then 10 cm. Fig. 3.5 shows the reconstructed position of the storeys of one line for different values of the velocity of the sea current, a typical value is 5 cm/s. The compass-tiltmeter system installed on every storey determines the orientation and inclination of the storey. A detailed description can be found in [30].

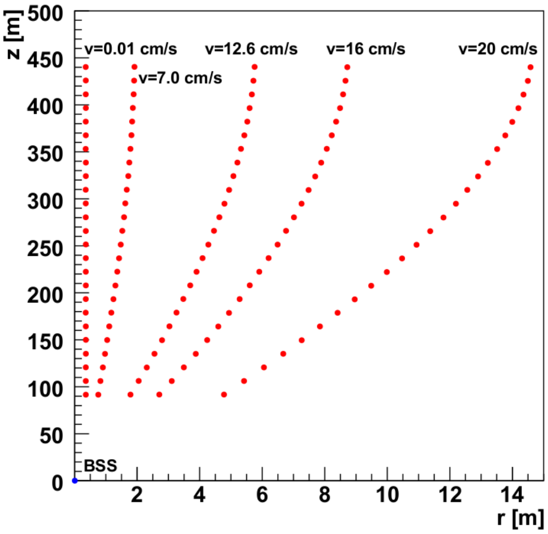


Figure 3.5: Reconstructed position of the storeys of one line with the acoustic positioning system for different sea currents. Figure taken from [30].

Optical background conditions

There are two different types of optical background. One is the decay of ^{40}K which is solved in the seawater and the other is bioluminescence, light emitted by living organisms such as bacteria, pyrosoma and plankton or some fish species [31]. The part of the optical background rate caused by ^{40}K is almost constant over time as the concentration of ^{40}K is constant (only slight changes due to exchange of water). But the contamination of the background rate caused by bioluminescence depends on the season and the sea water current. The first is due to the fact that for example in spring new feed flows from the Alps into the sea and animates the deep-sea fauna to emit light. On smaller time scales, the optical background rate changes due to the sea current as the animals are stimulated to emit light by hitting a detector structure, where the probability of hitting an OM rises with the turbulence around the structure. Typical background baseline rates are in the range from 50 to 100 kHz, whereas some short and local burst can have rates up to 1 MHz (see Fig. 3.6). As these rates vary with time and sea current the minimum bias data are taken to determine the status of the background conditions. The optical background are hits caused by these effects. To take these varying conditions into account a dedicated run-wise MC simulation is done for each run separately (see section 6.3).

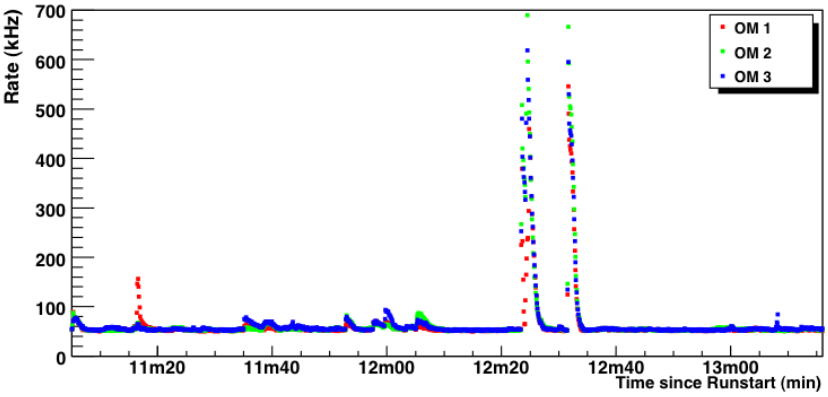


Figure 3.6: Rate of detected photons at the different OMs of one storey. Figure taken from [32].

3.2.4 Signal and background events

The ANTARES detector can measure signatures caused by muons traversing or passing the detector, which is called a signal. These muons are atmospheric muons or muons generated by cosmic or atmospheric neutrinos (neutrino signal). The different event types are shown in Fig. 3.7 and are described in the following. When talking about signal neutrinos, cosmic neutrinos are meant.

Cosmic and atmospheric neutrinos

The generation of both cosmic and atmospheric neutrinos follows the same interaction processes starting with the decay of a charged pion interacting with atoms in the interstellar medium or respectively in the atmosphere and producing muon-neutrinos and muons, which themselves can decay in a second process into electrons and electron-neutrinos as already discussed in chapter 2 see eqn. (2.2) and eqn. (2.3).

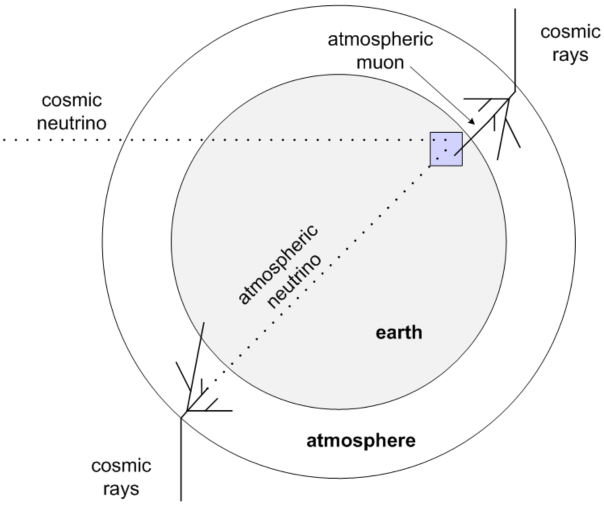


Figure 3.7: Schematic view of possible event sources detected with the ANTARES detector. Figure taken from [32]

If these neutrinos are created in the interaction of cosmic rays with the atmosphere they are called atmospheric neutrinos which are expected to follow an $E^{-3.7}$ spectrum. Cosmic neutrinos created by cosmic sources explained in chapter 2 are expected to generally follow an E^{-2} spectrum from Fermi-acceleration if no prediction for the neutrino flux shape per source candidate can be derived. Both cosmic and atmospheric neutrinos can reach the detector from every direction as their cross-section is very low, so that they can also traverse the Earth and reach the detector from below (see Fig. 3.7). Atmospheric neutrinos are isotropically distributed over the full sky. Even if they correspond to a neutrino signal these neutrino events are background events in the view of a search for cosmic neutrinos from point-like sources.

After the detection of a neutrino event one cannot discriminate between cosmic and atmospheric neutrinos. Therefore statistical methods are necessary to distinguish between them. In this thesis a maximum-likelihood-based analysis method using hypotheses testing is used. It takes the hypothesis of only background events (atmospheric neutrinos) as null hypothesis and evaluates if the measured signal is in agreement with this hypothesis or not (see chapter 7).

Atmospheric muons

If muons produced in charged pion decays (eqn. (2.2)) in the atmosphere of the Earth do not decay further, they can reach the Earth and also the detector. As their travel distance in earth, caused by their energy loss, is small, this can only be from above. So they are absorbed in the Earth and one uses the Earth as shield from atmospheric muons by using only upwards going, short upgoing, events. As the rate of atmospheric muons from above is about a factor 10^6 higher then the atmospheric neutrinos (see Fig. 3.8), the reconstructed direction of the muons is the main parameter used to discriminate between neutrinos and atmospheric muons, being aware of loosing neutrino events coming from above. Further methods could be using only starting tracks or contained events within the detector. But for neutrino energies above 1 TeV these new methods are only possible for large scale detectors such as IceCube or KM3NeT.

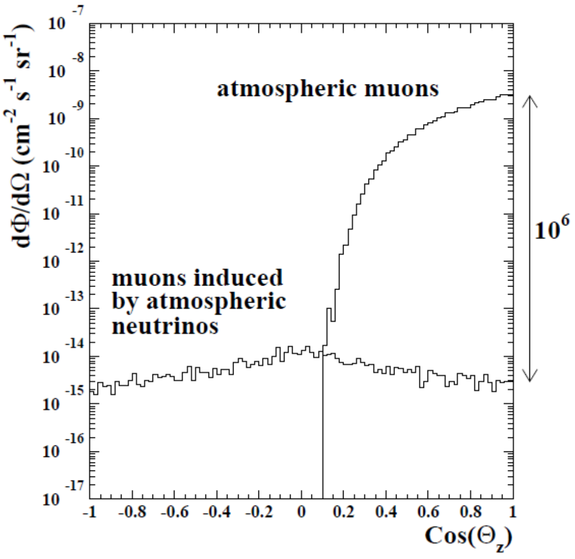


Figure 3.8: Zenith distribution of the muon flux for energies above 1 TeV for atmospheric muons and atmospheric neutrino induced muons in a depth of 2300 m in water. Figure taken from [24].

Chapter 4

Development of a myon track reconstruction algorithm

There are two standard track reconstruction strategies for ANTARES. One of these is a fast track reconstruction algorithm that is also used for online reconstruction of some events, called BBFit [33]. As this reconstruction has to be done in real time it is less accurate using for example the assumption of straight lines and the OMs are located at the center of the storey by ignoring the actual position of the OMs determined by the alignment process (see section 3.2.3). For offline reconstruction with this fit the position of the storeys is taken into account. In principle this strategy is based on a χ^2-fit with a modified χ^2-function. Its resolution for events reconstructed with hits on more than two lines is about 0.94° for an E^{-2} neutrino flux after applying quality cuts [33]. The other reconstruction strategy, called AAFit [34], uses the maximization of a likelihood starting with a robust prefit and a sequence of further prefit steps whose result is used as reference direction for the maximization of a likelihood. It has a resolution of about 0.4° for an E^{-2} neutrino flux after applying quality cuts [35].

A new concept of scanning the whole phase space in Θ and Φ in discrete steps and find the position of such a track fitting best to the hits of the event for one particular direction is established in this thesis. Such a track reconstruction algorithm called FilteringFit was developed and optimized by Claudio Kopper for a proposed geometry of the future KM3NeT detector, which showed promising results [36]. During this thesis this kind of fit was developed for the ANTARES detector to use as a prefit input for an high-resolution maximum likelihood fit. Getting more or additional information for one event of an additional fit to the standard ones, can increase the number of neutrino events for an point source analysis for example.

In this chapter the concept of the FilteringFit track reconstruction algorithm is explained, its adaptation and enhancement to be used as prefit for the ANTARES detector is described and its performance using an additionally maximum likelihood fit as final fit, as compared to the standard track reconstruction strategies of ANTARES, is shown.

4.1 The FilteringFit track reconstruction algorithm

The FilteringFit track reconstruction is based on scanning the whole phase space by generating a discrete number of track hypotheses (Θ, Φ) which are isotropically distributed over the sky. Fig. 4.1 shows an overview of its fitting procedure. For each of these directions an hit selection is performed using the largest number of causally connected hits. For this hit selection the

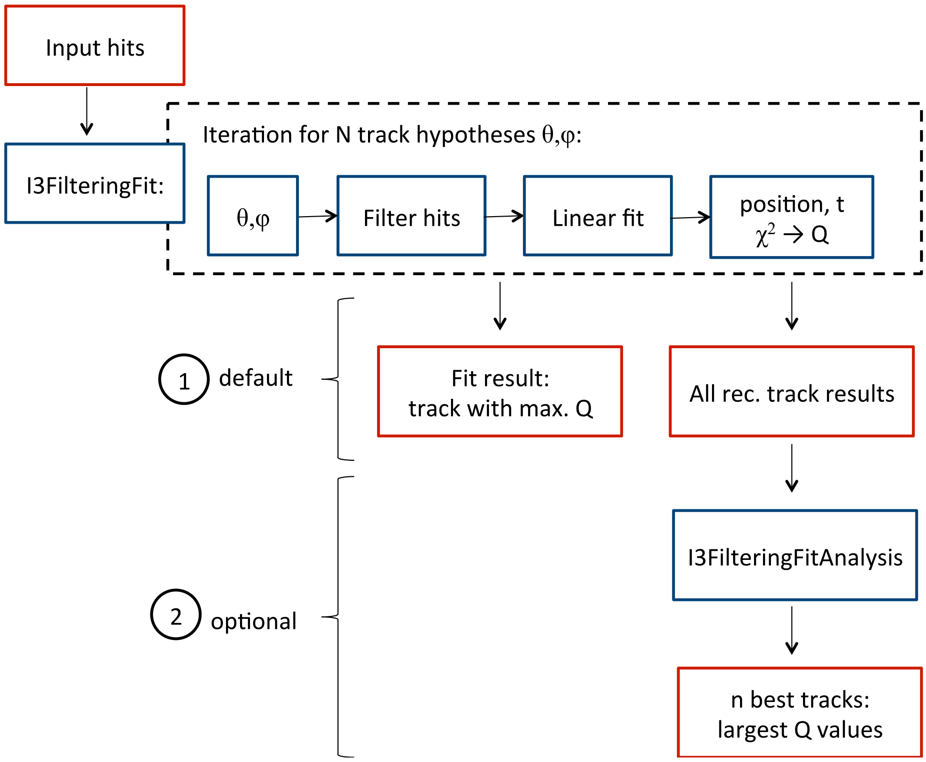


Figure 4.1: Overview of the FilteringFit track reconstruction algorithm. The default path is marked with 1, the optional path is marked with 2 for further reference.

position of the given direction for a point of time is fitted by a linear fit (according to [37]). A quality parameter Q is calculated for each hypothesis in order to determine the best fitted track assumption. The number of tested directions can be changed, changing the space angle between the different directions accordingly, and is therefore correlated with the resolution of the fit. In the following the separate parts of the fitting procedure per track assumption is explained in more detail. To simplify the problem the coordinate system is rotated so that the z-axis is pointing along the muon direction and is transformed back after the fitting process.

4.1.1 Hit selection: filtering of hits

The hit selection is based on a causality criterion where hits that are correlated in time and position are taken, the others are filtered out[1]. This procedure follows the clustering procedure of the so-called 3N trigger described in [38].

To calculate the expected arrival time t_i of an hit i at OM$_i$ at position $(x_i,\ y_i,\ z_i)$ to a assumed muon direction, it is convenient to rotate the coordinate system such that the muon

---

[1] Caused by this filtering of hits, the fit got its name.

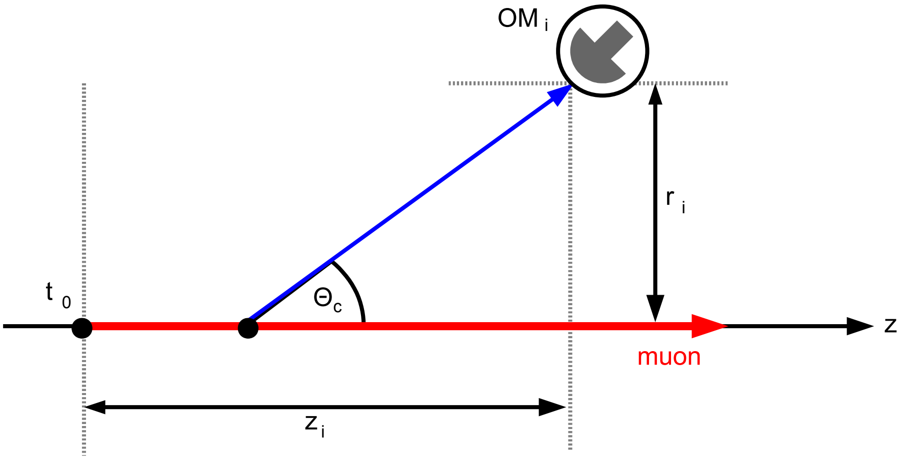


Figure 4.2: Definition of the geometric parameters for one emitted and detected photon at OM$_i$ in the rotated coordinate system, where the muon direction is identical with the z-axis.

direction points along the z-axis. Then t_i is given by

$$t_i = \underbrace{t_0}_{\text{reference time}} + \underbrace{\frac{1}{c}\left(z_i - \frac{r_i}{\tan\Theta_c}\right)}_{\substack{\text{time the muon needs to}\\ \text{travel to the photon emission point}}} + \underbrace{\frac{1}{v_g}\frac{r_i}{\sin\Theta_c}}_{\substack{\text{photon travel time from}\\ \text{the emission point to the OM}}} , \tag{4.1}$$

where z_i is the position along the muon track, r_i the perpenticular distance of the OM from the muon track (see Fig. 4.2), $v_g = c/n_g$ is the group velocity (c is the speed of light, n_g is the refractive index of water) and Θ_c the Cherenkov-angle in water. This formula simplifies to

$$t_i = t_0 + \frac{1}{c}(z_i + \kappa r_i) \text{ with } \kappa = \frac{c}{v_g \sin\Theta_c} - \frac{1}{tan\Theta_c} = \frac{n_g}{\sin\Theta_c} - \frac{1}{tan\Theta_c}. \tag{4.2}$$

So for two hits on two OMs the minimal perpendicular distances to the muon track r_1 and r_2 are unknown as only the direction but not the position of the muon is given. But the maximal distance between them in the xy-plane $R = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ is known. So it follows for the time difference

$$\underbrace{(t_2 - \frac{z_2}{c}) - (t1 - \frac{z_1}{c})}_{\Delta t} \leq \frac{\kappa}{c}R. \tag{4.3}$$

Including an extra time t_{extra} for time calibration uncertainties and scattering the criterion used in the FilteringFit algorithm is

$$\left(\Delta t_{corr} \cdot \frac{c}{\kappa}\right)^2 \leq R^2, \tag{4.4}$$

with $\Delta t_{corr} = |\Delta t| - t_{\text{extra}}$. Additionally to this criterion one can apply an maximal allowed distance between two hits $R < d_{\max}$.

The first step of the fitting procedure after rotating the coordinate system into the muon track direction is the filtering of the hits due to the criterion from eqn. (4.4). All hit pair combinations are tested to this criterion iteratively starting with one of the hits and finding the hits causally connected to it, which defines a cluster. The cluster of hits with the largest number of hits N_{hits} is taken for the actual fitting of the position of the muon track in the next step.

4.1.2 Linear fit

In the rotated coordinate system, fixing the direction (Θ, Φ) of the track, the original 5-dimensional problem becomes a 3-dimensional problem: finding the position of the muon track along the z-axis (x,y,t). This can be done with a linear fit described in [37]. In principle the task is to solve a linear system of equations $H\vec{p} = \vec{y}$, where $\vec{p} = (x_0, y_0, t_0)$ is the position of the muon track and $\vec{y}$ is a vector containing all hits (x_i, y_i, t_i).

$$H = \begin{pmatrix} 2(x_2 - x_1) & 2(y_2 - y_1) & -2(t_2 - t_1) \\ 2(x_3 - x_2) & 2(y_3 - y_2) & -2(t_3 - t_2) \\ . & . & . \\ . & . & . \\ 2(x_1 - x_n) & 2(y_1 - y_n) & -2(t_1 - t_n) \end{pmatrix} \tag{4.5}$$

$$\vec{y} = \begin{pmatrix} x_2^2 - x_1^2 + y_2^2 - y_1^2 - t_2^2 + t_1^2 \\ x_3^2 - x_2^2 + y_3^2 - y_2^2 - t_3^2 + t_2^2 \\ . \\ . \\ x_1^2 - x_n^2 + y_1^2 - y_n^2 - t_1^2 + t_n^2 \end{pmatrix} \tag{4.6}$$

This can be derived from eqn. (4.1), that can be rewritten as

$$t_i = t_0 + \frac{z_i - z_0}{c} + \tan(\Theta_c)\frac{r_i}{c}, \tag{4.7}$$

with the perpendicular distance to the track $r_i = \sqrt{(x_i - x_0)^2 + (x_i - x_0)^2}$. By taking the following substitutions $t'_0 = t_0 c/\tan(\Theta_c)$ and $t'_i = t_i c/\tan(\Theta_c) - (z_i - z_0)/\tan(\Theta_c)$ one can convert eqn. (4.7) to

$$(t'_i - t'_0)^2 = (x_i - x_0)^2 + (x_i - x_0)^2. \tag{4.8}$$

Calculating the difference of eqn. (4.8) for two different hits one gets

$$t'^2_j - t'^2_i - 2(t'_j - t'_i)'_0 = x_j^2 - x_i^2 - 2(x_j - x_i)x_0 + y_j^2 - y_i^2 - 2(y_j - y_i)y_0, \tag{4.9}$$

which is one of the linear equations that has to be solved. This can be formulated as $H\vec{p} = \vec{y}$ with the matrix H defined in eqn. (4.5) and $\vec{p}$ defined in eqn. (4.6) taking the respective substitutions for t' instead of the time t.

As the matrix H combines the time and position differences of two consecutive hits per row the ordering of the hits is important. They are therefore ordered before inserting in eqn. (4.5) and eqn. (4.6) by the polar angle with respect to their center of gravity and interleaved to have the maximal difference between them [39]. The $\vec{p}$ of equation $H\vec{p} = \vec{y}$ is the fitted muon position for this particular direction (Θ, Φ) for which the quality parameter χ^2 is calculated as

$$\chi^2 = \frac{\sum_{i=1}^{N_\text{hits}} t_{\text{res},i}^2}{\sigma^2} = \frac{\sum_{i=1}^{N_\text{hits}} (t_{i,\text{theo}} - t_{i,\text{meas}})^2}{\sigma^2}, \tag{4.10}$$

where $t_{\text{res},i}$ is called the time residual calculated by the difference of the expected arrival time $t_{i,\text{theo}}$ (see eqn. (4.1)) and the measured time of the hit $t_{i,\text{meas}}$, and $\sigma = 2\,\text{ns}$ which is a measure of the time resolution of the PMT.

4.1.3 Quality parameter

A modified quality parameter Q is calculated using the resulting χ^2 as

$$Q = N_\text{hits} - w \cdot \frac{\chi^2}{\text{NDoF}}, \tag{4.11}$$

where N_hits is the number of hits used for the fit, $\text{NDoF} = N_\text{hits} - 3$ is the number degrees of freedom for 3 free parameters and w is a weighting factor which can be set by the user. As the hits used for fitting are causally correlated with respect to the track direction, a track hypothesis with a large number of hits used for fitting should fit the event better, than a track hypothesis with less causally connected hits. Therefore this is included in the quality parameter Q, which assigns a track assumption with many hits a better quality than an assumption with less hits and a good χ^2 value. On the other hand as N_hits is an integer number, a low χ^2 helps to decide between direction with a similar number of hits as fine tuning. The parameter w can be used to fine tune the influence of the χ^2.

4.1.4 FilteringFit result

For each track hypothesis the steps of the hit selection, linear fit and the calculation of Q is performed. The fit provides the possibility to use several outputs, shown in Fig. 4.1 with a default path (path 1) and an optional path (path 2). As default output the track hypothesis with the largest Q value is taken as fit result, which is named as 'maxQ' in the following. In addition to the fit result the parameters of all reconstructed track hypotheses are stored in a vector class, which allows for the analysis of all performed fitting steps. This is done with the SeaTray[2] module 'I3FilteringFitAnalysis' as shown in Fig. 4.1, optional path. One option of this module is to select the fitting results of the tracks with the n best Q values. A minimal allowed distance between the tracks, considered for this selection, can be set to ensure that the n best tracks are spacial separated by at least that angle.

4.2 The high-resolution muon track reconstruction: KrakeFit

The aim in developing KrakeFit was to adapt the FilteringFit reconstruction strategy to be used as a prefit for a final fit which uses a maximum likelihood-based method. The resolution of the prefit should be about 1° as studies of an analytical PDF (probability density function) give this number as the maximum angular distance where this PDF can converge. For this purpose the

---

[2] SeaTray is the ANTARES software framework [40] adapted from the icetray software framework of IceCube [41].

FilteringFit module was extended in this study and compared to BBFit [33] for a small neutrino sample. Furthermore, a maximum-likelihood-based fit was applied to reach a higher resolution. In this section the different steps of the prefit and the final fit are explained, which is in total called 'KrakeFit'.

Definitions for plots

For this study a small sample[3] of upgoing neutrino events that are triggered by at least one of the standard ANTARES triggers (3N and/or 2T3 [42], [43], [44]) were used. An optical background of 60 kHz random noise and a fully working detector were assumed. The neutrinos have energies in the range between 1 TeV and 1 PeV and are weighted according to an E^{-2} energy spectrum. The angle between the true and the reconstructed muon direction is called $\Delta\alpha$ in the following. When talking about well-reconstructed events, those with an angular error $\Delta\alpha < 6^\circ$ are meant. With efficiency the efficiency of well-reconstructed events is meant, if not stated differently, it is defined as follows

$$\text{efficiency} = \frac{\text{reconstructed events with } \Delta\alpha < 6^\circ}{\text{all triggered events}}. \tag{4.12}$$

4.2.1 Performance of FilteringFit for ANTARES

As explained in the previous section 4.1, there are several parameters that can be tuned in FilteringFit and they have to be adapted for the ANTARES detector. For this purpose each of these parameters was optimized with respect to the median of the angular error, by holding the other parameters fixed. By this procedure the following parameters are found to work best for ANTARES.

As input hits for starting the hit selection not all detected hits are used as the cluster algorithm is quite time consuming, so the time of fitting scales factorial with number of input hits [39]. Further without a preselection of hits, the number of background hits is high (at a background rate of 60 kHz the mean number of background hits is about 260) which could fulfill the causality criterion just by chance. That minimizes the chance of finding the best cluster of hits fulfilling this criterion per direction. It turns out that the strict L1 hit criteria is a good choice as input for the hit selection for the studied high-energy regime as the number of signal hits is still high[4] (for low-energy events it might be different [32]). The minimum number of hits in a cluster for each direction is set to 6 hits to start the linear fit process, as in the non-linearized problem 5 parameters have to be fitted and at least 6 hits are needed to solve this problem. This number is not changed even though only 3 parameters have to be fitted once the direction (Θ, Φ) is fixed. The parameters $t_{\text{extra}} = 5\,\text{ns}$ and $d_{\text{max}} = 300\,\text{m}$ have been chosen, the latter is practically no limit on the transverse distance as the OMs are installed at about 350 m along one line (see Fig. 3.4 in section 3.2). The number of directions used in the fit (NumIterations) work best at 60000, as the distance between two track assumptions is smaller than 1°. For the calculation of the Q value, the weight w is set to 0.5. For the optional path the parameters were set that 10 tracks with the best Q values are selected that have a minimal distance of 1°

---

[3] 10^{10} neutrino events were generated for an energy range between 10 GeV and 10^7 GeV.

[4] The L1 hit selection collects roughly 50% of the signal hits with a purity better than 90% for an background rate of 60 kHz; the mean number of signal hits included is about 20 (70) for neutrinos with energies between 1 TeV to 10 TeV (100 TeV to 10^3 TeV).

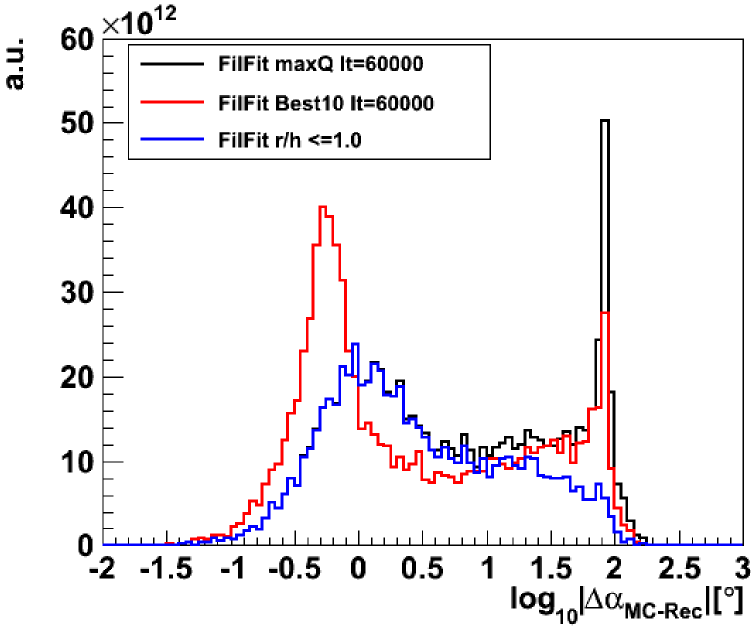


Figure 4.3: Distribution of $\log_{10}\Delta\alpha$ for the FilteringFit result maxQ (black), the best of the 10 best tracks (red) and for the FilteringFit result maxQ without tracks where $r/h > 1.0$ (blue), events are weighted according to an E^{-2} spectrum.

between them. If in the plots the label 'Best$n$' is used it means that the track with the minimal $\Delta\alpha$ out of the n best FilteringFit tracks is used for plotting[5].

The resulting angular error $\Delta\alpha$ of the FilteringFit is shown in Fig. 4.3 (black curve) exhibiting a strong peak at 84°. The assumption is that this is attributed to mirror solutions, which can appear if the pattern of hits used to fit the track forms a symmetry plane. These symmetries can be identified if the rank of the tensor of inertia calculated for the used hits is smaller than 3. It turned out that only 50% of the entries of the peak can be identified by this criterion. So a new or additional criterion is needed to identify badly reconstructed tracks, which will be discussed in section 4.2.2.

For the red curve in Fig. 4.3 and Fig. 4.4 the track with the minimal $\Delta\alpha$ within the 10 best tracks was used. The main peak of the angular error distribution in Fig. 4.3 (red curve) is shifted to lower angular errors compared to the standard FilteringFit result (black curve). This means that within the 10 best tracks there is most of the time at least one track that is nearer to the true muon direction than the standard FilteringFit result 'maxQ'. Therefore one could gain in precision by using the 10 best tracks instead of the final fit result of the FilteringFit for the next step. This requires a criterion to identify the track with the minimal angular error within these 10 tracks, which is discussed in section 4.2.2.

---

[5] It is worth to mention that here MC information is used to determine the track with the minimal $\Delta\alpha$ out of the n best tracks only for the plotting purpose. It is not used in the fitting procedure itself.

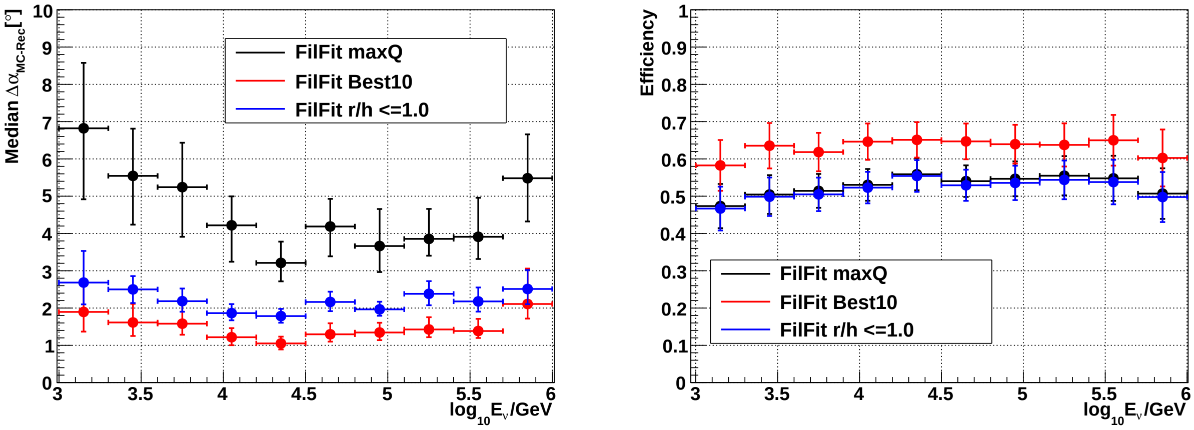


Figure 4.4: Left: Median of the angular error $\Delta\alpha$ for different neutrino energies E_ν/GeV; right: efficiency of reconstructed events with $\Delta\alpha < 6°$ for different neutrino energies E_ν/GeV. Both for FilteringFit result maxQ (black), the best of the 10 best tracks (red) and for the FilteringFit result maxQ without tracks where $r/h > 1.0$ (blue).

4.2.2 Extension of the FilteringFit reconstruction chain

Criterion to identify badly reconstructed tracks: cylinder criterion r/h

The following criterion showed to be effective indentifying badly reconstructed events: For every reconstructed track direction L1 hits are selected with time residuals between $-5\text{ns} \leq t_{\text{res}} \leq 20\text{ns}$. The time residual t_{res} is the difference of the measured time of the hit and its theoretical time t_{theo} which can be calculated with eqn. (4.1). One can define a cylinder around each track with minimal radius r and minimal height h containing all the selected hits. The assumption for neutrinos with energies above 1 TeV is that these hits lie along the track within a small distance. The reason is that high-energy neutrinos produce muons with long travel distances through the detector; the larger the distance of the detected photons to the muon track, the higher is the probability of scattering. Therefore hits with small time residuals lie along the track with relatively small distance compared to the long track lengths. Reconstructed muon tracks that have a larger cylinder radius than height are expected to be misreconstructed (Fig. 4.5 shows two different track assumptions for the same hit pattern and the corresponding cylinder containing these hits). Therefore tracks with cylinder criterion $r/h > 1.0$ are disregarded. In Fig. 4.3 the distribution of the angular error is shown in black without r/h cut and in blue with r/h cut. One can see that most of the badly reconstructed tracks are rejected without losing well-reconstructed tracks. For tracks reconstructed with angular errors smaller than 5° 98.8% of events remain, but tracks reconstructed with angular errors larger than 5° (60°) only 56.5% (20.5%) survive this cut. This can also be seen in Fig. 4.4 in the right plot, where the efficiency of well-reconstructed tracks is shown. There is almost no difference between the black and the blue curve. Further in the left plot of Fig. 4.4 the median of the blue line is (as expected) better and gets closer to the median of the red curve, which defines the optimum obtainable by the strategy.

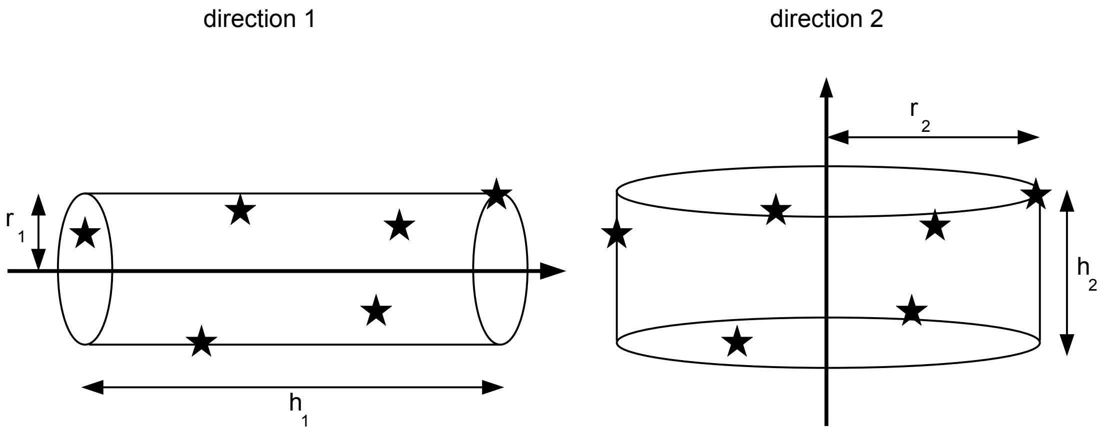


Figure 4.5: Schematic drawing of cylinders with radius r_i and height h_i calculated around two different track hypotheses with the same hit pattern.

Criterion to identify a well-reconstructed track within the n best tracks: surface density criterion (sD)

The cylinder criterion can be used as a preselection of tracks out of the n best tracks. The next step is to have a criterion to find the best track within the selected ones. This is performed using the surface density (sD) which is defined as

$$sD = \sum_{i=1}^{3} \frac{N^2_{\mathrm{hits,r_i}}}{\sum\limits_{\mathrm{OMs}} AngAcc_{r_i}}, \tag{4.13}$$

where $N_{\mathrm{hits,r_i}}$ is the number of hits and $\sum\limits_{\mathrm{OMs}} AngAcc_{r_i}$ is the sum of the angular acceptance of the OMs within in a cylinder with radius r_i around the track. The chosen values for the three radii determined by optimization are $r_1 = 2.0 \cdot 55.0\,\mathrm{m} \cdot \sin\theta_c$, $r_2 = 1.5 \cdot r1$, and $r_3 = 2.0 \cdot r1$, motivated by the absorption of light in seawater (see section 3.1.3). Eqn. (4.13) is adapted from [45] where it was established for low-energy events. The underlying idea is the same as for the cylinder criterion. The number of hits around the track decreases exponentially with the distance to the track. The fraction of detected hits over the sum of OMs that could have detected light should be higher for inner cylinders than for outer cylinders. Eqn. (4.13) takes this into account. Instead of using the number of OMs within the cylinder the angular acceptance of the OMs, which is the detection probability of light depending on the angle of impact at the PMT (ranges from 0 to 1), assuming direct photons from the track, is used as OMs hit for example from the back should not be taken into account. If one track assumption doesn't fit the true muon direction well, this fraction should be small and therefore its sD value will be small compared to a track assumption which is closer to the truth. The sD value is calculated for each track and the one with the largest sD value is selected.

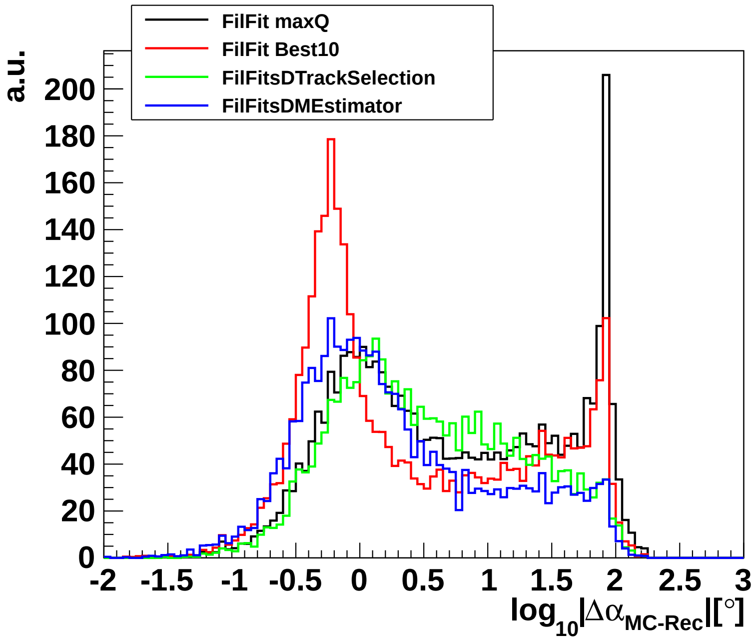


Figure 4.6: Distributions of $\log_{10}\Delta\alpha$ for the FilteringFit result maxQ (black), the best of the 10 best tracks (red), the FilteringFitsDTrackSelection result (green) and FilteringFitsDMEstimator result (blue).

Combination of r/h and sD: FilteringFitsDTrackSelection

After some tests it turned out that the combination of the r/h and the sD criteria gives the best results. This combination is implemented as a SeaTray module called 'FilteringFitsDTrackSelection' in the filteringfit project. The n best tracks are first filtered so that only tracks that fulfill the r/h criterion are passed to the calculation of the sD value, where the track with the maximum sD is taken as result. This means that events where none of the n best tracks fulfills the cylinder criterion are lost. The combination with the sD criterion enhances the performance compared to using only the maxQ tracks that pass the r/h cut, because the total efficiency is increased[6]. The distribution of $\Delta\alpha$ for this step is shown in Fig. 4.6 (green curve).

Improving the resolution: MEstimator fit

In the AAFit [34] and the BBFit [33] reconstruction algorithms, applying an MEstimator fit results in an improvement in angular resolution. The MEstimator downweights outliers in the fit and with this improves the angular resolution by maximizing $\sum_{\mathrm{i}=1}^{N_{\mathrm{hits}}} g(t_{\mathrm{res},i})$ with

$$g(t_{\mathrm{res},i}) = -2\sqrt{1+\frac{t_{\mathrm{res},i}^2}{2}}+2, \tag{4.14}$$

---

[6] As an example if the maxQ track doesn't fulfill the r/h criterion, the event would be rejected. But if the n best tracks are used, there is still the chance, that one of the other $n-1$ best tracks of this event fulfills this criterion and so it won't be filtered out.

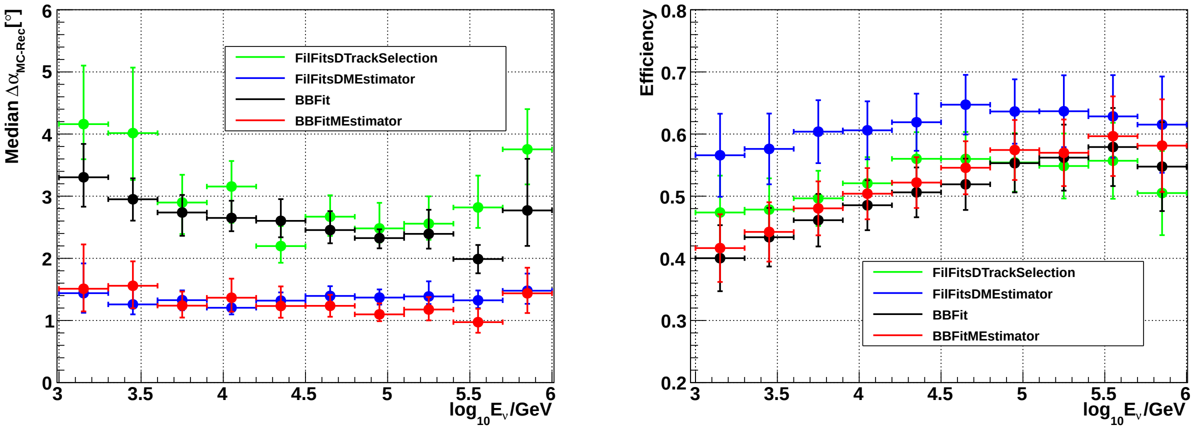


Figure 4.7: Left: Median of the angular error $\Delta\alpha$ for different neutrino energies E_ν/GeV; right: efficiency of reconstructed events with $\Delta\alpha < 6°$ for different neutrino energies E_ν/GeV. Both for FilteringFitsDTrackSelection (green), FilteringFitsDMEstimator (blue), BBFit (black) and BBFitMEstimator (red).

where $t_{\text{res},i}$ (in ns) is the time residual of the hit. This function is linear in t_{res} for large value of t_{res}, but quadratic for small values of t_{res}. Therefore hits with larger time residuals have less influence in the result of the fit than hits with small time residuals. As the AartStrategy [34] is implemented as a chain of single modules in SeaTray and it is possible to use its individual parts also in other reconstruction chains, its hit selection and MEstimator are used here.

The output of the n best tracks after applying the r/h and sD track selection (FilteringFitsDTrackSelection) was used as a reference direction for the hit selection and the MEstimator fit. As can be seen in Fig. 4.6, the use of the MEstimator (blue curve) moves the main peak to lower angular errors and thus the resolution improves by using this fit. In Fig. 4.7, the median angular error is shown for all reconstructed events on the left and the efficiency of well-reconstructed events (events reconstructed with angular errors less then 6°) are shown on the right. In comparison to Fig. 4.3 and Fig. 4.4, one can see an improvement in the resolution as well as the efficiency for well-reconstructed events. In the plots the BBFit and BBFitMEstimator results are also shown. The FilteringFitsDMEstimator gives comparable results for the median of the angular error and improves the efficiency compared to BBFitMEstimator, where no quality cuts are applied[7]. One can see an improvement of the FilteringFit reconstruction chain for high-energy neutrino events with this extension of the reconstruction chain, whose result will be called 'FilteringFitsDMEstimator' and has a resolution for neutrino energies above 1 TeV between 1.2° and 1.5°. It is the prefit of an high-resolution track reconstruction chain, which is shown in the following section.

---

[7] BBFit single-line events are disregarded here as only the zenith but not the azimuth could be determined for such events with that strategy. So the space angle between reconstructed track and the muon direction cannot be calculated.

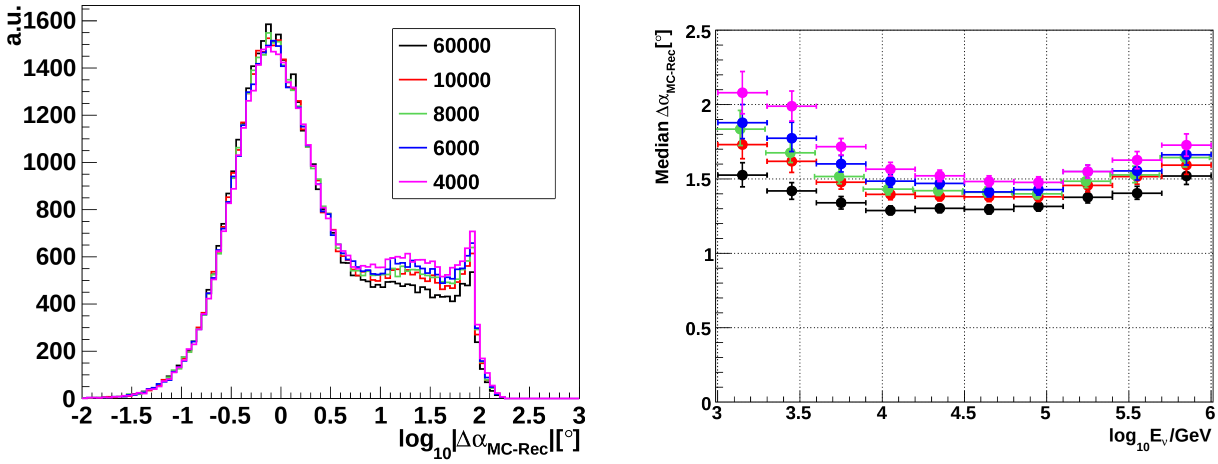


Figure 4.8: Left: Distribution of $\log_{10}\Delta\alpha$; right: median of the angular error $\Delta\alpha$ for different neutrino energies E_ν/GeV. Both for the FilteringFitsDMEstimator result after applying different numbers of track directions 60000 (black), 10000 (red) 8000 (green), 6000 (blue) and 4000 (pink) in the FilteringFit step.

4.2.3 Overview of the total track reconstruction algorithm

As shown in the previous section, the prefit starting with the FilteringFit module and ending with the MEstimator fit shows good results. The fitting procedure is time consuming, where the computing time is dominated by the FilteringFit step. The CPU time per event depends on the number of directions used for the fit as for each of them the hit selection and linear fit has to be performed. To decrease the computing time, this number was reduced to find an optimal value while still having a good resolution after the MEstimator step. Fig. 4.8 (left) shows the distribution of $\log_{10}\Delta\alpha$ for the 'FilteringFitsDMEstimator' step for different numbers of iterations set in the FilteringFit part. On the right plot the median of the space angle error is plotted versus the neutrino energy. Combining these two plots one can deduce that the lower the number of iterations, the worse the resolution gets. A good compromise between resolution and computing time is using 10000 iterations instead of 60000 per event. The mean computing time per event for $E_\mu = 1\,\mathrm{TeV}$ is then approximately 0.3 s and for $E_\mu = 10\,\mathrm{TeV}$ is approximately 1.0 s (see Fig. 4.9). These numbers are calculated for events that passed the whole fitting chain.

The final parameters for the prefit are given here:

- FilteringFit
  - InputHits: L1 hits
  - NumIterations: $10 \cdot 10^3$
  - $t_{\mathrm{extra}} = 5\,\mathrm{ns}$
  - $d_{\max} = 300\,\mathrm{m}$
  - $w = 0.5$
  - Minimum Cluster Size $N_{\mathrm{hits}} \geq 6$

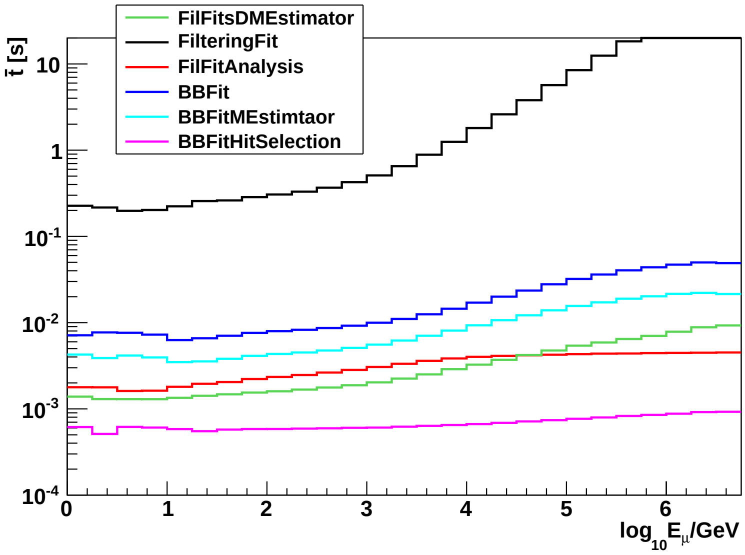


Figure 4.9: Mean computing time per event as a function of the muon energy. Only events were considered that passed the whole fitting chain.

  - $n = 10$
  - MinAngularDistanceForNewBestPoint $= 1°$
- FilteringFitsDTrackSelection
  - InputHits: L1 hits
  - InputTracks: 10 best FilteringFit tracks
  - HitSelection: $-5\,\mathrm{ns} \leq t_{\mathrm{res}} \leq 20\,\mathrm{ns}$
- MEstimatorFit
  - InputHits: L0 hits
  - InputTrack: FilteringFitsD track
  - HitSelection: $-150\,\mathrm{ns} \leq t_{\mathrm{res}} \leq 150\,\mathrm{ns}$, $d_{\mathrm{max}} = 100\,\mathrm{m}$

All steps up to now constitute a prefit, so a good approximation for the neutrino direction but not for the purpose of a point source search where a resolution of less then 1° is needed. Therefore a maximum-likelihood-based final fit is used to establish an high-resolution track reconstruction strategy. The idea is to find that trajectory that maximizes the likelihood for the given hit pattern $L = \prod_{i=1}^{N_{\mathrm{hits}}} P_i(x_1, x_2, ...)$ or minimize the negative logarithm of it

$$-\log L = -\log \prod_{i=1}^{N_{\mathrm{hits}}} P_i(x_1, x_2, ...) = -\sum_{i=1}^{N_{\mathrm{hits}}} \log P_i(x_1, x_2, ...), \tag{4.15}$$

where $(x_1, x_2, ...)$ are the parameters of the probability function. The better the fit is the higher the likelihood should be and therefore the smaller the minus log-likelihood should be. So an quality parameter, if the fit converges well, is the reduced log-likelihood value (rlogL), which is

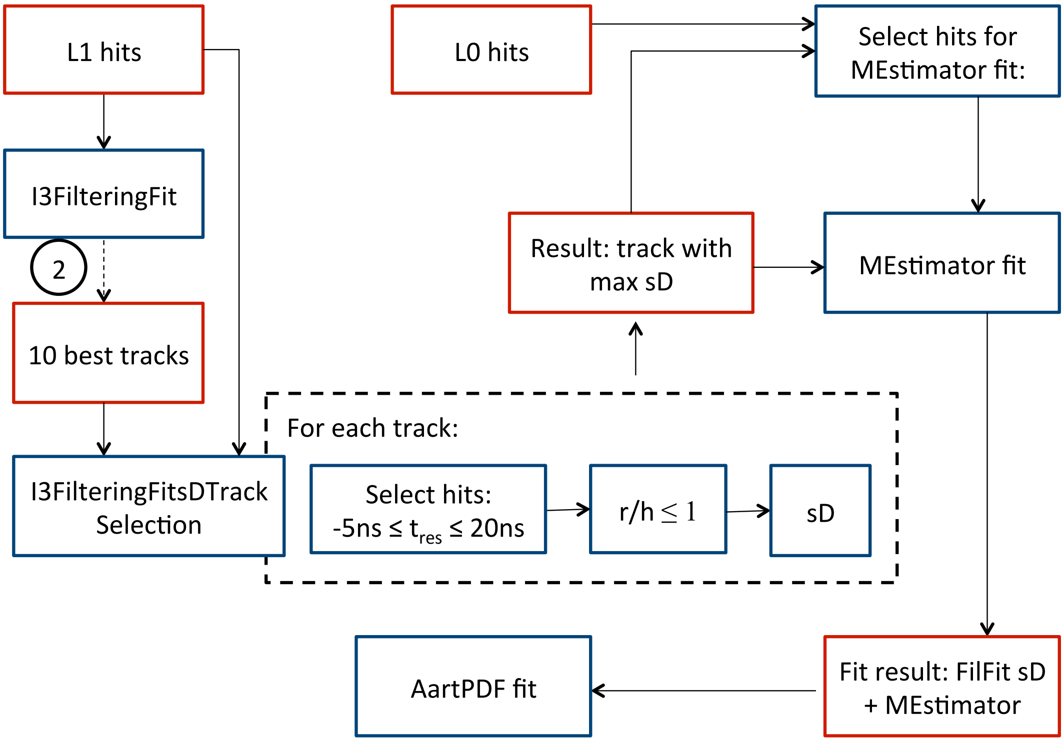


Figure 4.10: Overview of the enhancement of the FilteringFit track reconstruction chain (called 'FilteringFitsDMEstimator'). With the AartPDF fit as final fit it is called 'KrakeFit'. The '2' marks the optional path shown in Fig. 4.1. The cylinder criterion r/h and the surface density criterion sD is explained in section 4.2.2.

defined as rlogL $= \frac{\log L}{NDoF}$. Here the probability density function AartPDF developed in [34] was used, where the underlying hit time probability was extracted from simulations. For the final fit step an hit selection starting with all detected hits of the event (L0 hits) is performed, where hits are selected with small time residuals and small distances due to the FilteringFitsDMEstimator track. The maximal allowed distance of hits to the track is set to 300 m and the time residual limits are set to $t_{\text{res,min}} = -150\,\text{ns}$ and $t_{\text{res,max}} = 150\text{ns}$. Only the timing information and the position of the hits are used, their charge is not taken into account.

The whole fitting chain is called KrakeFit in the following. Its performance as compared to the standard ANTARES track reconstruction chains is studied in the following section.

4.3 Performance of KrakeFit

For the performance study $8 \cdot 10^{11}$ upgoing (anti-)neutrino events that were either triggered by the ANTARES 3N and/or 2T3 trigger algorithms ([42], [43], [44]) were used. In the following an optical background rate of 60 kHz random noise and a fully working detector were chosen if

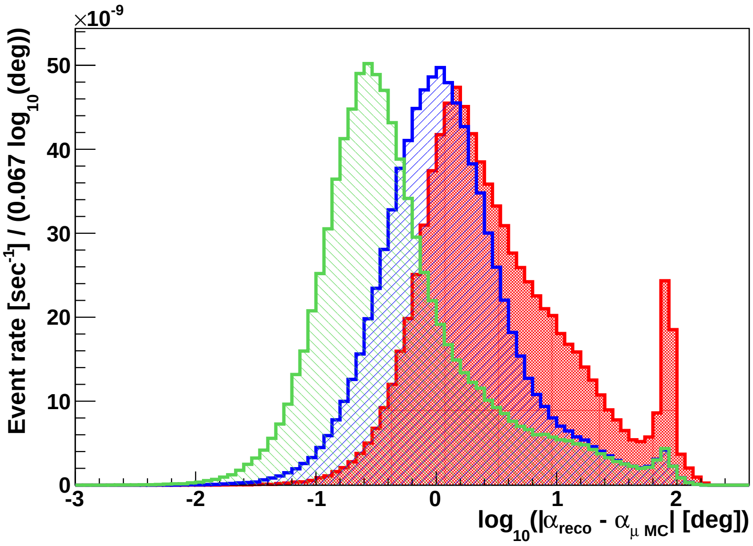


Figure 4.11: Distribution of the angular error $\Delta\alpha$ for the default output of FilteringFit (red), the result of the FilteringFitsDMEstimator (blue) and the KrakeFit result (green) with the cuts rdf=1, $s > 0.9$, rlogL < 5.6 for events reconstructed as upgoing. Events are weighted according to an E^{-2} spectrum.

not stated otherwise. The neutrinos have energies in the range between 10 GeV and 10 PeV and the events are weighted according to an E^{-2} spectrum. For the study of atmospheric muons the same detector conditions and trigger conditions were applied. Atmospheric muons are simulated corresponding to a live-time of 30 days with the parametrization of the muon flux described in [46].

The distribution of the angular error of the KrakeFit track reconstruction steps is shown in Fig. 4.11 for a combination of cuts obtained in section 4.3.1. To compare the results with other track reconstruction strategies it is important that the rate of misreconstructed atmospheric muons is nearly the same. Therefore an optimization of the cut parameters is done with respect to the rate of misreconstructed atmospheric muons.

4.3.1 Cut optimization

The reduced log-likelihood value of the final fit can be used as quality criterion of the fit. This value can also be used to distinguish between misreconstructed atmospheric muons and upgoing neutrinos which is shown in Fig. 4.12, where the cumulative distribution of rlogL is shown for upgoing reconstructed atmospheric neutrinos (dashed black curve) and misreconstructed atmospheric muons (continuous black curve). To get a reasonable misreconstruction rate of atmospheric muons a cut on rlogL < 5.5 is needed. So the question is, is it possible to use another parameter or value in order to get more neutrinos but the same misreconstruction rate? For this purpose an up-down-classification tool using random decision forest [47] is examined.

The use of the random decision forest (rdf) [47] is to identify upgoing events with high purity and high efficiency. The tool is available as the SeaTray module 'I3RDFClassify' [48]. In principle it has input parameters with several decision trees built on them that decide if the

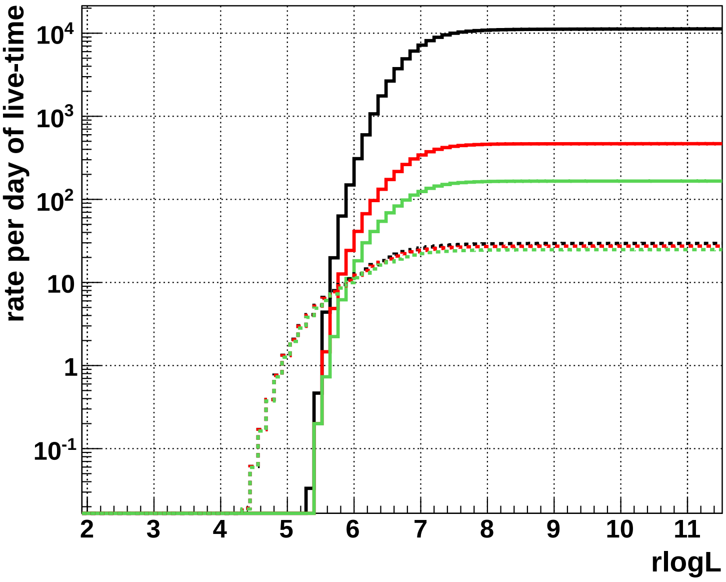


Figure 4.12: Cumulative distribution of the event rate per day of live-time depending on rlogL for upward reconstructed events with KrakeFit. Solid lines for atmospheric muons, dashed lines for atmospheric (anti-)neutrinos. The colors identify the cuts: black without cuts, red with rdf=1 and $s > 0.8$ and green rdf=1 and $s > 0.9$.

event is up- or down-going. The final decision is defined as the class where more than 50% of the decision trees agree. This is stored in the RDFClass (rdf) value which indicates if it is upgoing (rdf=1) or downgoing (rdf=0). Additionally the RDFSafety (s) value is stored which is the fraction of the trees which gave this decision. The higher this RDFSafety is, the more certain is the classification. Fig. 4.12 shows also the effect of the usage of rdf for two different safety (s) values, where the cumulative distribution of rlogL is shown for upward reconstructed events with KrakeFit (continuous lines for atmospheric muons, dashed lines for atmospheric neutrinos). From this plot one can see the effect of the rdf compared to the line without cut by looking at the rlogL value of 10, where the cumulative distribution is flat. The usage of rdf with $s > 0.8$ reduces the number of misreconstructed atmospheric muons by a factor of 22.4 and with $s > 0.9$ by a factor of 62.2. The number of upgoing reconstructed neutrinos is only slightly reduced, 93.4% (85.1%) of the neutrinos remain for rdf=1 with $s > 0.8$ ($s > 0.9$).

Another question is: does the rdf cut affect the performance of KrakeFit? As it shall only be used to distinguish between misreconstructed atmospheric muons and upgoing neutrinos it should not affect the resolution of the fit for neutrinos or the efficiency. A comparison is performed for neutrinos with an E^{-2} energy spectrum for KrakeFit without cut (blue line) and with rdf=1 and $s > 0.8$ (green line), which is shown in Fig. 4.13. By comparing the blue and the green curve in the upper left plot, one can see that the efficiency is slightly reduced for neutrino energies above 1 TeV, the difference of the efficiency for these curves is 1.3% (1.8%) for 1 TeV (100 TeV). The median of the angular error (lower left plot) for 1 TeV (100 TeV) is about 0.54° (0.18°) smaller. In the upper right plot the efficiency depending on the zenith angle is shown, one can see a reduction for zenith angles $< 100°$. For nearly horizontal upgoing neutrino events the rdf has difficulties identifying them as upgoing due to low training statistics in this zenith area [49].

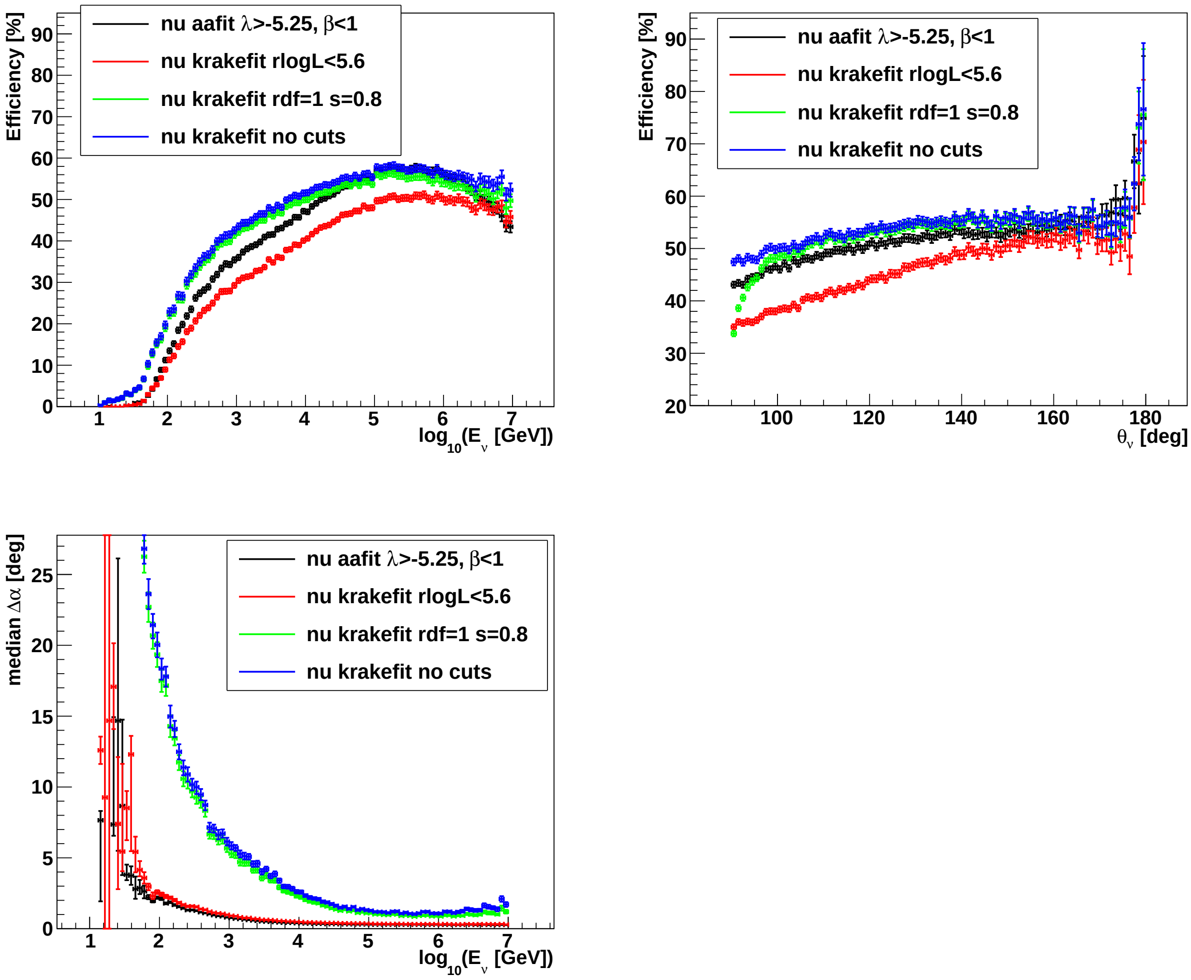


Figure 4.13: Top: Efficiency of reconstructed events with $\Delta\alpha < 6°$ for different neutrino energies (left) and depending on the zenith angle (right); bottom: median of the angular error $\Delta\alpha$ for different neutrino energies. Plots generated for events reconstructed as upgoing with KrakeFit without cuts (blue), with rdf=1 $s > 0.8$ (green) and with rlogL < 5.6 (red) and with AAFit with $\lambda > -5.25$ and $\beta < 1°$ (black).

The conclusion is that despite the strong effect in the efficiency for nearly horizontal neutrino events the rdf has not much influence on the performance of KrakeFit. For finding an optimal cut of rlogL, using additionally the rdf, the relevant part of the cumulative distribution of Fig. 4.12 which is 4.0 < rlogL < 6.5 is shown in Fig. 4.14.

To get a comparable rate of misreconstructed atmospheric muons with AAFit and BBFit, the misreconstruction rate should be around 0.35 to 0.50 atmospheric muons per day. Without rdf this would lead to rlogL < 5.52, with rdf=1 and $s > 0.9$ one can loosen the rlogL cut up

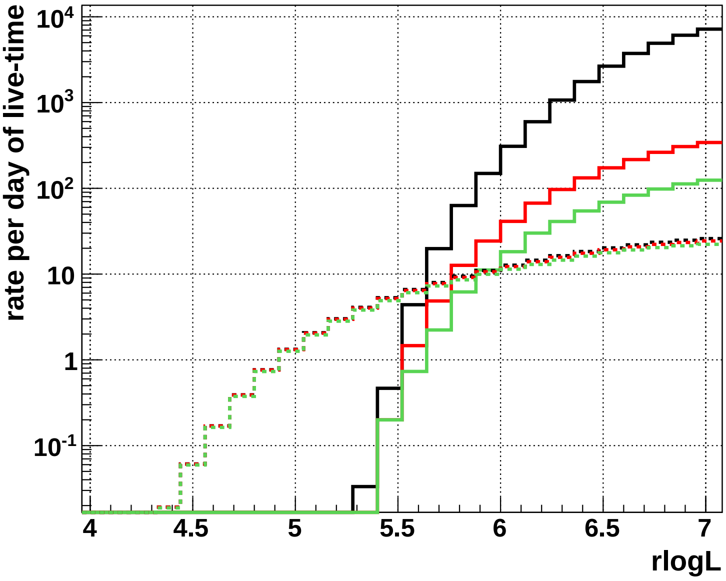


Figure 4.14: Cumulative distribution of the event rate per day of live-time depending on rlogL in the range 4.0 < rlogL < 6.5 for upward reconstructed events with KrakeFit. Solid lines for atmospheric muons, dashed lines for atmospheric (anti-)neutrinos. The colors identify the cuts: black without cuts, red with rdf=1 and $s > 0.8$ and green rdf=1 and $s > 0.9$.

to 5.64 to have nearly the same misreconstruction rate and gain in the number of atmospheric neutrinos. So two different rlogL cuts in this rlogL-bin were tested and give the following rate values per day:

- KrakeFit rdf=1, $s > 0.9$, rlogL < 5.55:
  - 0.37 atmospheric muons per day
  - 5.18 atmospheric neutrinos per day
- KrakeFit rdf=1, $s > 0.9$, rlogL < 5.6:
  - 0.49 atmospheric muons per day
  - 5.65 atmospheric neutrinos per day.

These misreconstruction rates are similar to those from AAFit and BBFit with the following cuts and can therefore be used for a comparison with these two reconstruction strategies, where typical cuts for these strategies are applied [35][33]:

- AAFit $\lambda > -5.25$, $\beta < 1°$:
  - 0.20 atmospheric muons per day
  - 7.72 atmospheric neutrinos per day

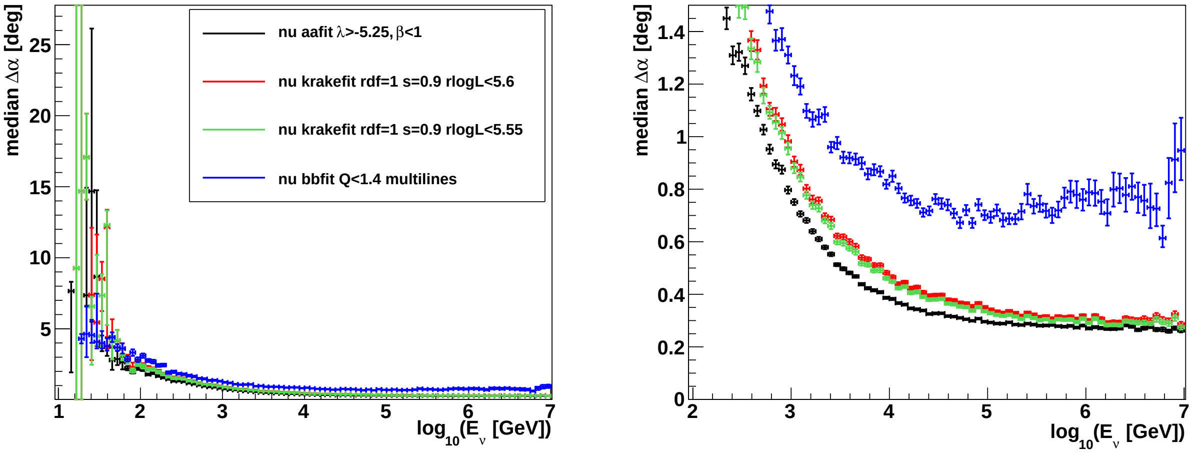


Figure 4.15: Median angular error $\Delta\alpha$ for neutrino events reconstructed as upgoing as a function of the neutrino energy for AAFit ($\lambda > -5.25$, $\beta < 1°$) (black), BBFit ($Q < 1.4$, NLines> 1) (blue), KrakeFit (rdf=1, $s > 0.9$, rlogL < 5.6) (red) and (rdf=1, $s > 0.9$, rlogL < 5.55) (green) for an E^{-2} energy spectrum. Right: Median angular error for neutrino energies above 100 GeV (same color coding).

- BBFit $Q < 1.4$, NLines> 1:
  - 0.30 atmospheric muons per day
  - 5.90 atmospheric neutrinos per day.

4.3.2 Comparison with standard ANTARES track reconstruction strategies

As an example, Fig. 4.11 shows the distribution of the space angle error of the different KrakeFit steps for events that survive the cuts rdf=1, $s > 0.9$, rlogL < 5.6. The various steps shift the maximum of the distribution to smaller angular errors. The median angular error for KrakeFit with the two cuts is shown in comparison to AAFit ($\lambda > -5.25$, $\beta < 1°$) and BBFit ($Q < 1.4$ NLines> 1) in Fig. 4.15. The right plot is a zoom into the relevant part where the median of the space angle error of KrakeFit is slightly worse than AAFit. Table 4.1 compares the median angular error for AAFit and KrakeFit for different neutrino energies. The resolution of KrakeFit at high energies ($E_\nu > 10$ TeV) is comparable to the resolution of AAFit. Below 50 GeV BBFit has a better resolution than KrakeFit, but the tendency changes for higher neutrino energies (see Tab. 4.2 and Fig. 4.15). The efficiency of KrakeFit is shown in Fig. 4.16 on the left versus

E_ν [TeV]	AAFit($\lambda > -5.25$, $\beta < 1°$)	KrakeFit(rdf=1, $s > 0.9$, rlogL < 5.55)
1	0.75°	0.88°
10	0.38°	0.45°
100	0.29°	0.33°

Table 4.1: Median angular error of AAFit and KrakeFit for different neutrino energies.

the neutrino energy and on the right versus the zenith angle of the muon track. The efficiency of well-reconstructed tracks is much higher than for BBFit but lower than for AAFit. At a neutrino energy of 10 TeV there is a difference of about 10%. For large zenith angles (which means vertical upgoing tracks) the efficiency of KrakeFit and AAFit are the same. However the efficiency of KrakeFit drops faster for smaller zenith angles and for nearly horizontal tracks there is a lack of efficiency which is due to the rdf, already discussed above. This lack in efficiency is also visible in the effective area plots in Fig. 4.17. The zoom on the right-hand figure shows the reduction in efficiency for energies above 1 TeV. From Fig. 4.18, which shows the ratio of the effective areas of KrakeFit and AAFit, one observes that the reduction is of the order of 30% and independent of energy.

E_ν [TeV]	BBFit($Q < 1.4$ NLines$>$ 1)	KrakeFit(rdf=1, $s > 0.9$, rlogL < 5.55)
0.05	3.7°	4.2°
0.1	3.1°	2.4°
1	1.3°	0.88°

Table 4.2: Median angular error of BBFit and KrakeFit for different neutrino energies.

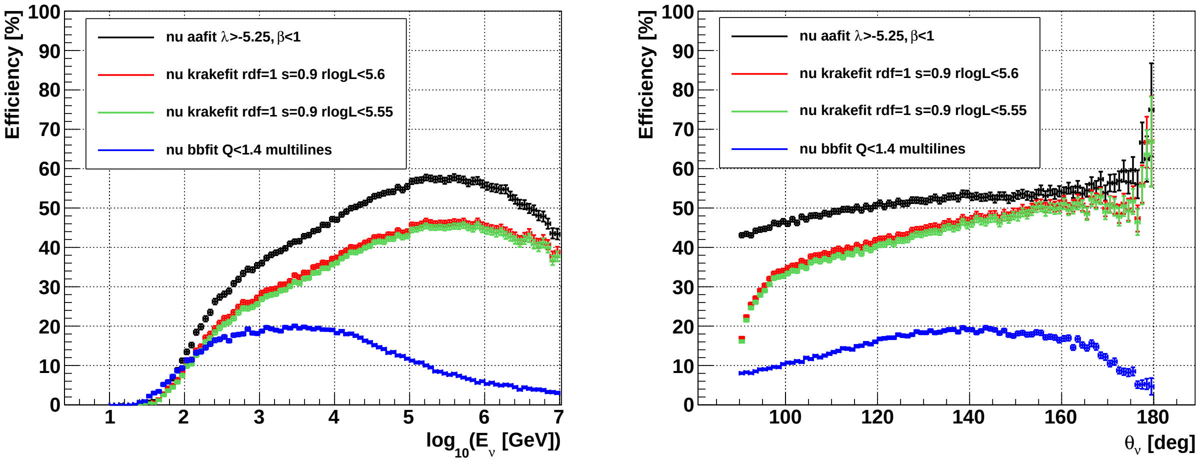


Figure 4.16: Efficiency for neutrino events reconstructed as upgoing with $\Delta\alpha < 6°$ for AAFit ($\lambda > -5.25$, $\beta < 1°$) (black), BBFit ($Q < 1.4$, NLines$>$ 1 (blue), KrakeFit (rdf=1, $s > 0.9$, rlogL < 5.6) (red) and (rdf=1, $s > 0.9$, rlogL < 5.55) (green) depending on the neutrino energy (left) and the zenith angle (right). Plots generated with events weighted according to an E^{-2} spectrum.

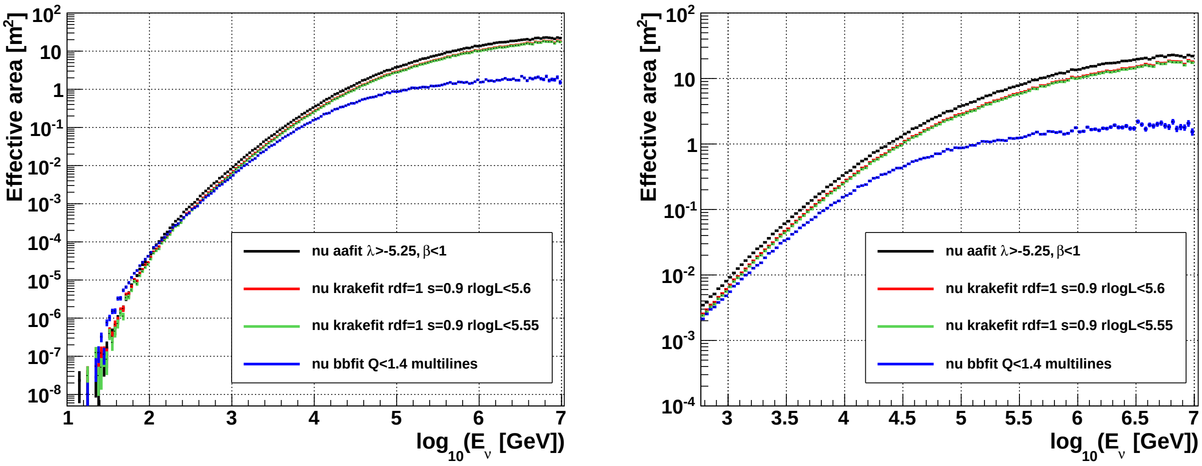


Figure 4.17: Effective area depending on neutrino energy for events reconstructed as upgoing with AAFit ($\lambda > -5.25$, $\beta < 1°$) (black), BBFit ($Q < 1.4$, NLines$>$ 1) (blue), KrakeFit (rdf=1, $s > 0.9$, rlogL < 5.6) (red) and (rdf=1, $s > 0.9$, rlogL < 5.55) (green) (left) for energies above 1 TeV (right).

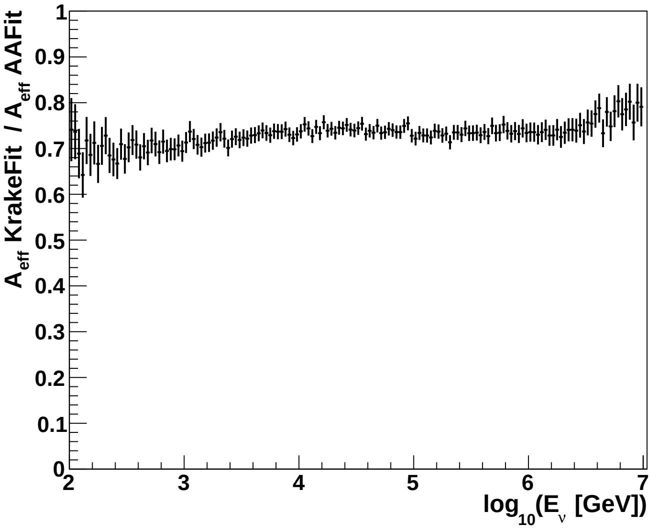


Figure 4.18: Ratio of the effective area of KrakeFit (rdf=1, $s > 0.9$, rlogL$L < 5.55$) and AAFit ($\lambda > -5.25$, $\beta < 1°$).

Conclusion

KrakeFit is a reconstruction strategy which has an high resolution and a good efficiency for the reconstruction of (high-energy) neutrinos and can therefore be used for a point source analysis approach. In terms of resolution and efficiency it performs slightly worse than AAFit (the standard ANTARES high-resolution track reconstruction strategy). As these two track reconstruction strategies start with different prefits and select therefore different hits according to them for the final fitting step, they reconstruct different tracks and events. So KrakeFit can be used for example as a cross-check for the muon track direction reconstructed with AAFit. For this purpose upgoing neutrino events were selected that were reconstructed with AAFit ($\lambda > -5.25$, $\beta < 1°$) and KrakeFit (rdf=1, $s > 0.9$, rlogL < 5.55). 35.9% of all triggered events were reconstructed with both strategies, 20.8% were only reconstructed with AAFit and 5.6% were additionally reconstructed with KrakeFit. Fig. 4.19 shows a comparison of the space angle error for events reconstructed with both strategies. The majority of the events are reconstructed with similar reconstruction errors. In Fig. 4.20 the distribution of the space angle between these two reconstruction strategies for the stated cuts is shown. The mean of this distribution is 1.4°.

It has to be studied if one could gain from the additional information obtained by KrakeFit. For example it could be determined if one could gain in sensitivity from the additional 5% of neutrinos. Or for example it can be studied if the overlap of events reconstructed with both strategies can reduce the angular uncertainty or reduce the number of atmospheric muons in the sample, which could lead to a looser cut and gain also in sensitivity. An approach of combining different track reconstruction results in a multivariate data analysis is studied in an ongoing phd thesis [50]. These optimization processes have to be done for each analysis separately.

In this thesis it is studied which sensitivity could be obtained with the KrakeFit track reconstruction strategy for a set of potential high-energy neutrino point sources, for which a cosmic neutrino flux expectation per source is available derived from gamma ray measurements (see chapter 6 section 6.2). First a data-MC-comparison is performed in chapter 5 for the observables determined by KrakeFit and possible cut parameters for the optimization process of a maximum likelihood point source analysis are studied.

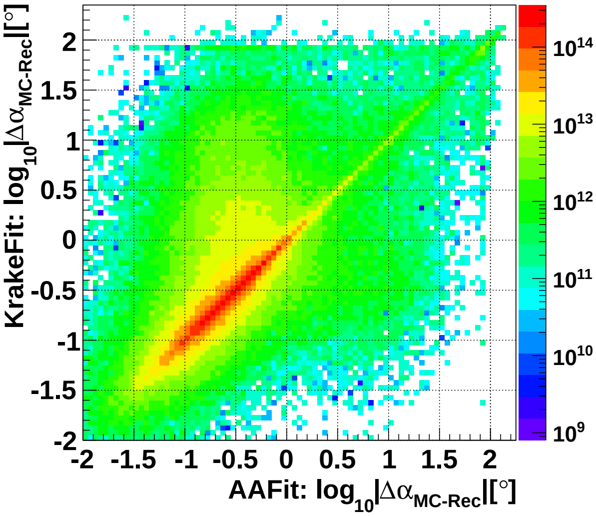


Figure 4.19: Comparison of the angular error for neutrino events reconstructed with AAFit ($\lambda > -5.25$, $\beta < 1°$) and KrakeFit (rdf=1, $s > 0.9$, rlogL < 5.55), events are weighted according for an E^{-2} spectrum.

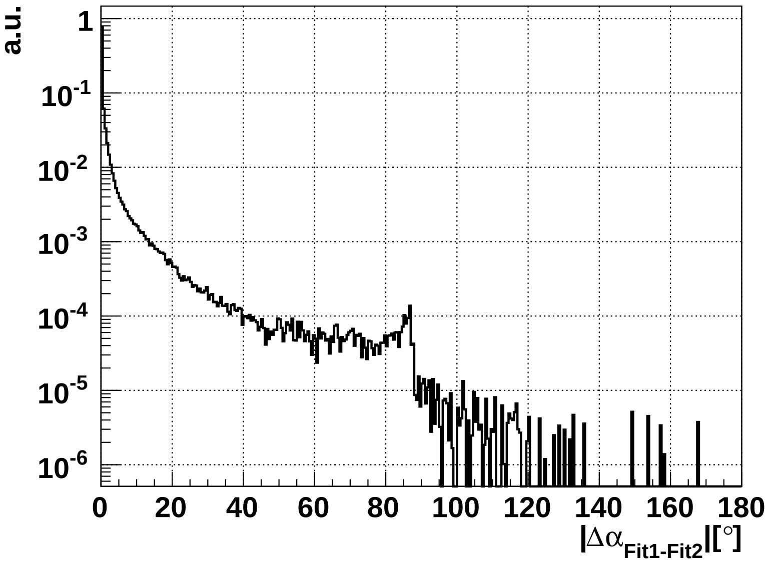


Figure 4.20: Distribution of the space angle between the AAFit track ($\lambda > -5.25$, $\beta < 1°$) and the KrakeFit track (rdf=1, $s > 0.9$, rlogL < 5.55) for neutrinos weighted according to an E^{-2} spectrum.

Chapter 5

Data-Monte-Carlo-comparison

5.1 Selection of runs

The construction of the ANTARES detector started 2006 with the installation of the first line [28] and lasted about 2.5 years in which the detector was constructed stepwise. During the construction phase, data were taken with a non-complete detector configuration. The 12-line detector was complete in spring 2008 and is taking data with the full detector configuration since May 2008. As the track reconstruction is optimized for a 12-line detector, only runs are considered that were taken in the 12-line period starting from May 2008.

As the quality of the recorded data varies from run to run, a preselection of runs with good conditions is used, which were preselected from the IFIC group for their point source analysis [51], which ranges from 2007 to 2012. The underlying criteria were

- physics runs: no calibration or acoustic (AMADEUS[1]) runs
- no scan runs: the scan flag is set for runs which are in principle used for preliminary tests on new calibrations, e.g. tunings of the PMT high voltage
- Quality flag ≥ 1: minimum requirement for a run to be selected. The flag is set by the ANTARES data quality group, which performs quality checks, for example that the trigger rates are within acceptable limits
- no sparking run: 'sparking' flags the runs with known issues of sparking PMTs.

Additionally I applied more requirements to the runs. For every run a valid MC simulation of atmospheric muons and (anti-) neutrinos must be available. The run-wise MC simulation (rbr), available at the starting point of this analysis, ended with runs taken in November 2012. So only runs taken between May 2008 and November 2012 are taken into account. Additionally runs with known issues qualified by the data quality group and not covered by the criteria above are not taken into account. There are 11007 runs in the 12-line period from 2008-2012, this number reduces to 7603 with a live-time of 744.68 days requiring the stated criteria and that KrakeFit has been executed on these runs[2]. Tab. 5.1 summarizes the number of runs and the corresponding live-time, it contains the total (default) run selection called 'totRuns' and a subsample of it '0Runs', which contains only runs whose run number ends with 0. This corresponds to roughly

---

[1] AMADEUS is an acoustic neutrino detection test system of ANTARES [27], which takes separate runs.

[2] Because of technical issues KrakeFit was not applied to a small number of runs.

name	number of runs	live-time
totRuns	7603	744.68 days
0Runs	717	70.67 days

Table 5.1: Summary of the number of selected runs and the corresponding live-time for the 12-line-data-taking period from 2008 to 2012.

10% of the totRuns selection and can be used for systematic studies on a small data sample representing the whole data well.

5.2 Data-MC-comparison of KrakeFit

For the official data production the output per reconstruction strategy has to be reduced to the minimal necessary information used for the different analyses as the storage space is limited. So not the whole information, e.g. hits, per fitting step of KrakeFit is available for all runs. To study the agreement between data and MC for the different prefit steps one run was chosen. Without applying quality cuts, the sample is dominated by atmospheric muons as the number of expected neutrinos for the live-time of one run (of the order of 1-2 hours) is negligible. So only atmospheric muons are considered in this comparison. The comparison plots are shown in the appendix B.1 where a good agreement between data and MC can be observed.

For the final fit result the data-MC-comparison is given here for the run selection stated above. If not stated differently only events reconstructed as upgoing are shown for the totRuns selection. The color code is in all plots the same: for data black, for atmospheric muons red, for atmospheric neutrinos green, for the sum of atmospheric neutrinos and muons blue. If the ratio of data to MC is calculated, the MC is represented by the sum of atmospheric neutrinos and muons.

5.2.1 General observables

For studying the data-MC-agreement for KrakeFit some general parameters concerning the fit result are compared. These parameters are the zenith angle and azimuth angle, which represent the sensitivity of the fit to local directions, the number of hits and number of lines used for the reconstruction with KrakeFit and the quality parameter rlogL. These parameters are used to inspect if the event reconstruction works as expected and if the MC describes the data well. As the quality parameter rlogL is the main parameter, after the requirement of upgoing tracks, to distinguish between atmospheric muons and neutrinos, it is essential that the expectation of the atmospheric muon background in data can be derived from the MC, and therefore a good agreement between data and MC in the relevant cut region is necessary.

Fig. 5.1 left shows the distribution of the zenith angle for all reconstructed events. The ratio between data and MC is in relativ good agreement (ratio ≈ 1.1) for zenith angles below $60°$, above the disagreement rises. In this case no quality cut on the reconstructed events is applied. Therefore the distribution of rlogL is shown only for upgoing events (zenith $> 90°$) in Fig. 5.1 right. Below an rlogL value of about 5.4 the neutrinos are dominant, above this value the distribution changes and the atmospheric muons dominate. The ratio between data and MC for this distribution fluctuates around 1 below rlogL value of 6.4. This is a good agreement between data and MC up to these rlogL values.

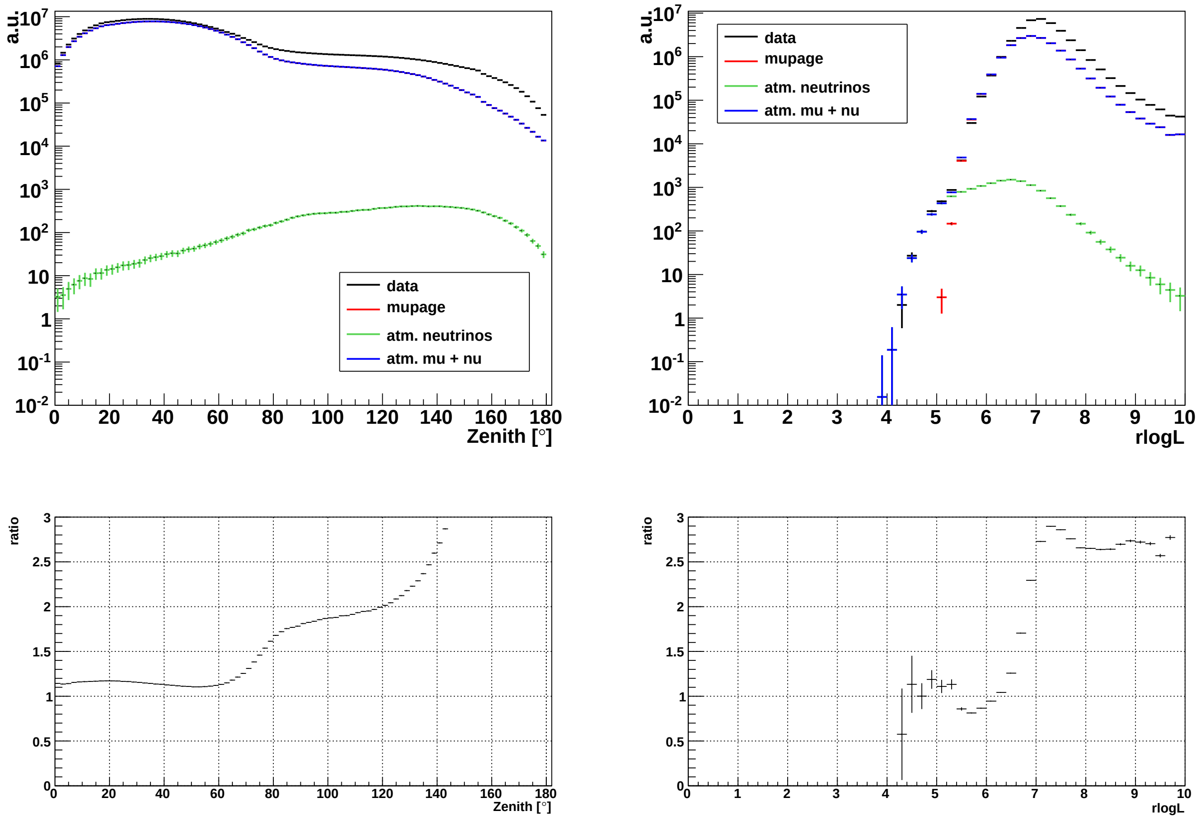


Figure 5.1: Distribution of reconstructed zenith angle for all reconstructed events (upper left) and rlogL for upgoing reconstructed events (upper right) with KrakeFit for data (black), atmospheric muons (red), atmospheric neutrinos (green) and the sum of atmospheric muons and neutrinos (blue). The bottom plots show the ratio between data and MC for the distributions above.

The following data-MC-comparison plots are done using events passing the cuts for upgoing events (except for the distribution of the zenith angle) and rlogL < 5.5 to show a general data-MC-comparison for a quality cut value that could be used in a point source analysis because of the good agreement of data and MC for that value and because of a possible fraction of atmospheric muon contamination in the neutrino sample. Fig. 5.2 shows four different distributions and their respective data-MC-ratio directly below the distribution. On the left upper plot the distribution of the zenith angle for rlogL < 5.5 is shown, which is now in perfect agreement between data and MC for upgoing events. The distribution of the azimuth angle (see Fig. 5.2 top right) shows only slight variations which can therefore be regarded as flat distributed, which is also in good agreement for data and MC. In the lower plots of Fig. 5.2 the distributions of the number of lines and number of hits used in the reconstruction for upgoing events with rlogL < 5.5 are shown. The MC curve follows perfectly the data curve, which can be seen at a ratio of 1 for the number of lines. For the number of hits the distribution is in better agreement for less hits and differs more at $N_{\text{hits}} > 130$ where also the statistics is low and the errorbars get large. To summarize the data and MC is in good agreement applying quality cuts on the direction of the event and

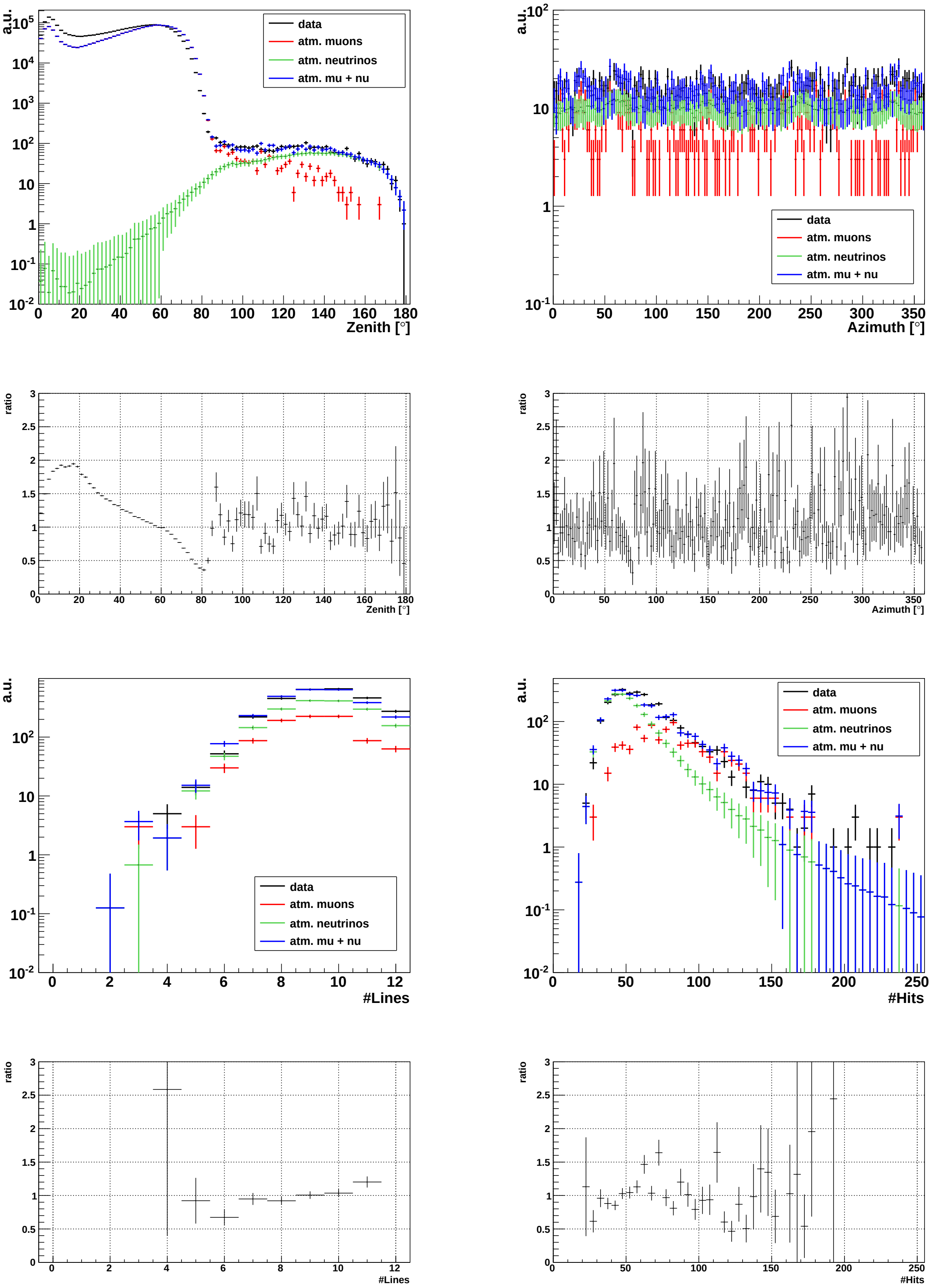


Figure 5.2: Distribution of reconstructed zenith and azimuth angle, number of lines and hits used in the reconstruction with KrakeFit with rlogL < 5.5. The respective bottom plots show the ratio between data and MC for the distributions above.

the log-likelihood value rlogL of the direction reconstruction in reasonable borders up to about rlogL ~ 6. I refer to the appendix B.2 Fig. B.4, where the data MC comparison is shown for rlogL < 6 for the 0Runs selection.

5.2.2 Derived observables

In the section above some general parameters that are used for a data-MC-comparison are shown to demonstrate that the event reconstruction works as expected and the MC describes the data well. In this section the focus lies on the search for further derived observables that could be used for example as pre-quality cuts to suppress the atmospheric muon background further.

For this purpose an hit sample defined as 'direct hits' are used. It is a subsample of the hits used in the final reconstruction step (maximum-likelihood-based fit) that fulfill the time residual criteria $-5\,\mathrm{ns} < t_{\mathrm{res}} < 20\,\mathrm{ns}$, which is calculated with respect to the fitted track. The number of direct hits and their distribution along the track determined by perpendicular projection on the track (length along track) are shown in the upper plots of Fig. 5.3 for upgoing reconstructed events with rlogL < 5.5. The number of direct hits is in good agreement between data and MC, which is also the case for their length along the track. In this case one can see that for lengths above approximately 170 m the neutrinos dominate the distribution and is nearly constant whereas the curve for atmospheric muons slightly decreases. This confirms the idea used in the cylinder selection criteria of section 4.2.2 that identifies misreconstructed events.

The nearest distance of the track to the detector center is calculated in the xy-plane (horizontal distance to detector center) and along the z-axis (vertical distance to the detector center). Both distributions are shown in the lower plots of Fig. 5.3 for events reconstructed as upgoing with rlogL < 5.5. The vertical distance shows a more or less constant distribution for atmospheric muons for distances up to 140 m. For neutrinos the vertical distance to the detector center decreases slowly.

As already mentioned in section 4.3.2 the space angle between the AAFit and the KrakeFit reconstructed track might be useful as cut criteria. It is calculated in the case that both strategies fitted the event. Without applying any quality cut no benefit can be seen (Fig. 5.4 left). Applying a cut on the rlogL value as shown in the right plot of Fig. 5.4 (upgoing reconstructed events with KrakeFit with rlogL < 5.5), indeed leads to an excess of neutrino events of the order of one magnitude in the first bin for low space angles between these two track reconstructions. So this parameter has the potential being a second cut which has to be optimized for each rlogL value separately, which requires a dedicated study of this effect on the sensitivity of a point source analysis.

In the previous chapter the use of the rdf was applied to get comparable misreconstruction rates of atmospheric muons compared to the AAFit und BBFit track reconstruction strategies. The rdf classification tool decides if the event is upgoing (rdfclass or short rdf=1) or downgoing (rdf=0) and gives the rdf safety parameter which indicates how many decision trees decided for this classification. The rdf safety value, which ranges from 0.5 to 1.0 per rdf class (as it is only a two-component decision)[3], is shown in Fig. 5.5. The data-MC-comparison is in good agreement and one can clearly identify the excess in neutrinos for rdf safety > 0.9.

---

[3] The minimal requirement for one decision is that at least 50% of the decision trees have the same result. So it ranges from 0.5 to 1.0 per identification e.g. rdf=1.

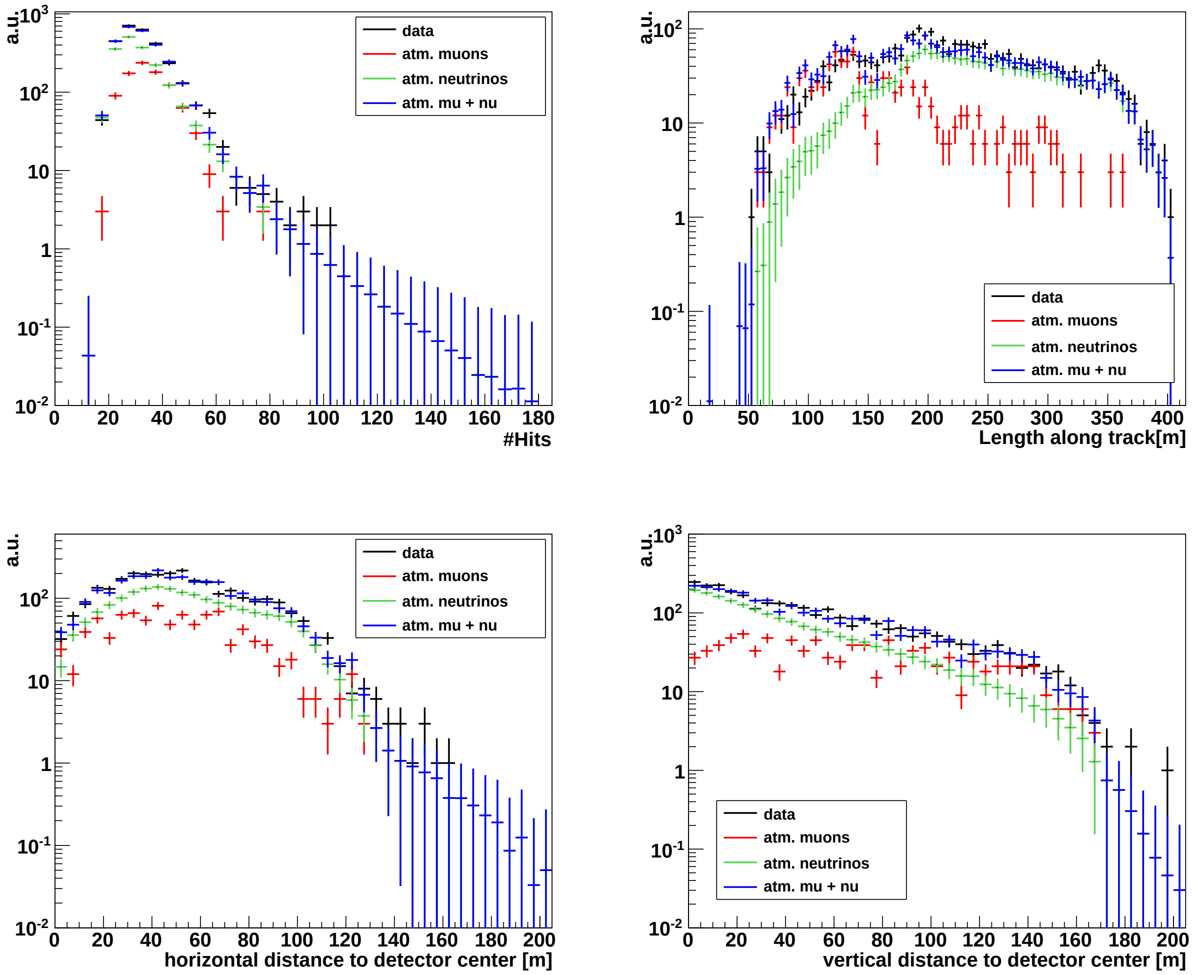


Figure 5.3: Distribution of number of hits and length along the track for 'direct hits' and horizontal and vertical distance to the detector center for events reconstructed as upgoing with KrakeFit with rlogL < 5.5 for data (black), atm. muons (red), atm. neutrinos (green) and the sum of atm. muons and neutrinos (blue).

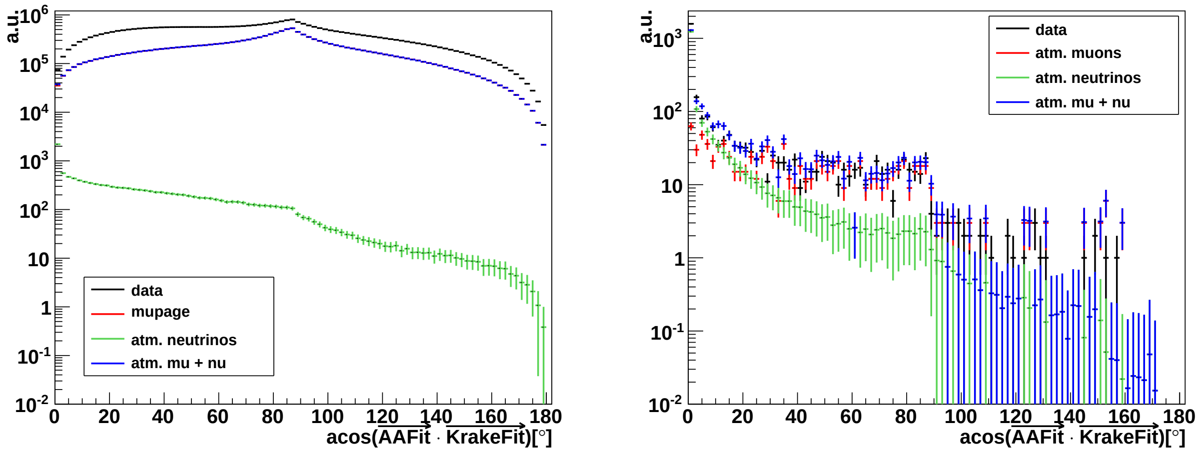


Figure 5.4: Distribution of the space angle between KrakeFit and AAFit for events reconstructed as upgoing for KrakeFit without quality cut (left) and with cut rlogL < 5.5 (right). Both for data (black), atm. muons (red), atm. neutrinos (green) and the sum of atm. muons and neutrinos (blue).

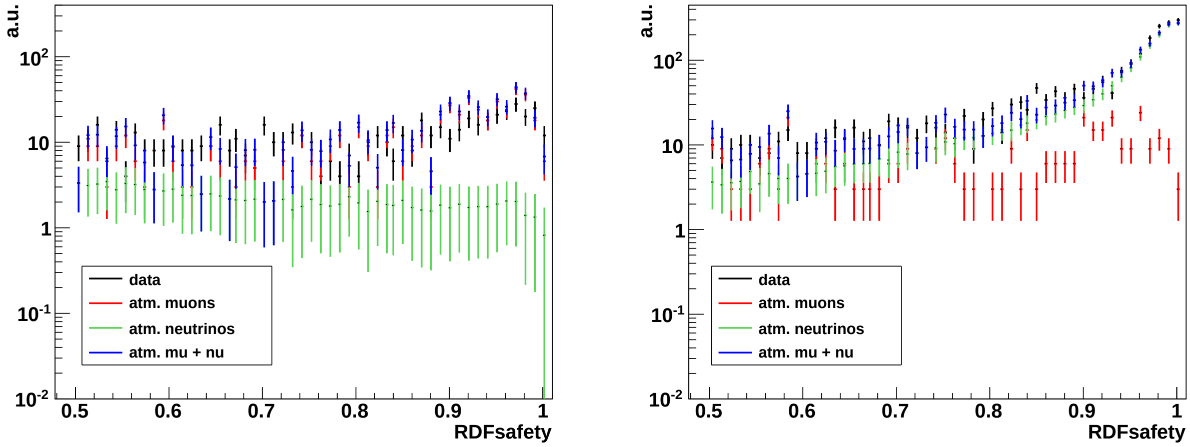


Figure 5.5: Distribution of rdf safety for rdf=0 (left) and for rdf=1 (right) for events reconstructed as upgoing with KrakeFit with rlogL < 5.5. Both for data (black), atm. muons (red), atm. neutrinos (green) and the sum of atm. muons and neutrinos (blue).

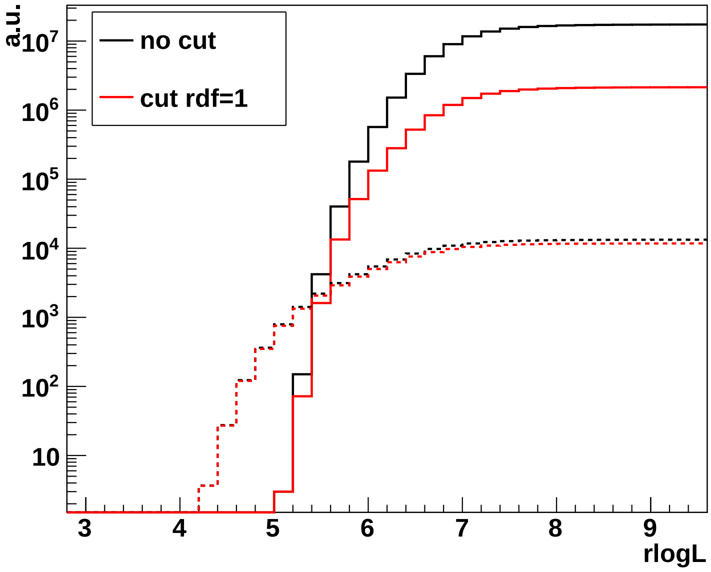


Figure 5.6: Cumulative distribution of the rlogL value for events reconstructed as upgoing with KrakeFit for atmospheric muons (continuous lines) and atmospheric neutrinos (dotted lines) without (black) and with (red) rdf=1 cut.

For a test how efficient the rdf cut is for more realistic detector conditions the cumulative distribution of the rlogL value is shown in Fig. 5.6 once for atmospheric muons (continuous lines) reconstructed as upgoing and once for atmospheric neutrinos (dotted lines) reconstructed as upgoing with (red line) and without (black line) rdf cut. The minimal requirement of the rdf=1 classification with safety value of at least 0.5 is used here. 88.3% of atmospheric neutrinos are kept and 87.7% of atmospheric muons are rejected, these numbers are calculated from the distributions for the totRuns selection without any further cut except that the events are reconstructed as upgoing. It has to be studied in detail if the rdf cut gives good and efficient results for a point source analysis or not. For completeness the data-MC-comparison plots for the general parameters applying additionally to the rlogL cut the rdf=1 cut can be found in the appendix B.2 Fig. B.5 where a good agreement between data and MC is observed.

To conclude there are some derived parameters that might be good parameters to separate between neutrinos and atmospheric muons. I see potential in the distribution of direct hits along the track (length of direct hits), the space angle between the two reconstruction strategies and the rdf classification with its safety value. A dedicated study is needed to give a final answer to its potential for the different kinds of analyses. It might be useful to combine the parameters and use a self learning algorithm to find the optimum. Different combinations of these parameters might be needed depending on the kind of analysis done, for example if it is a point source analysis, extended-source analysis, if the sources are stacked or not, or if one just looks for an overall atmospheric neutrino flux measurement. An example of the optimization for an point source analysis is shown in chapter 7, where the rlogL parameter is optimized and it is tested if one gains from an additional rdf=1 cut.

Chapter 6

The five strong Milagro sources

The Milagro Gamma Ray Observatory is a large-area water Cherenkov detector located at the Fenton Hill site of the Los Alamos National Laboratory. To detect TeV gamma rays of the entire observable sky, a water reservoir ($60\,\mathrm{m} \times 80\,\mathrm{m} \times 8\,\mathrm{m}$) is instrumented with more than 700 PMTs installed in two layers ([52],[53]). The upper layer ('shower layer') is used to reconstruct the direction of the detected air shower with an average accuracy of about 0.75°. The bottom layer ('muon layer') distinguishes between hadronic air showers and air showers initiated by gamma rays [53].

The Milagro detector was build - amongst other goals - to measure the gamma ray emission from the galactic plane [52]. 2007 the collaboration published the measurement of eight galactic TeV gamma ray candidate sources ([2],[3]) (see Tab. 6.1), where four sources (Crab, MGRO J2019+37, MGRO J2031+41 and MGRO J1908+06) exceed a significance of 4.5σ. Three of the eight sources can be associated to pulsar wind nebular (PWN). Four of the others are localed in the Cygnus region, one is near the Galactic center (see Fig. 6.1). The Cygnus region is a so-called star-forming region with a high rate of star-formation. With the combination of supernova remnants (SNR) and molecular clouds, it has a high potential of emitting high-energetic neutrinos. As cosmic rays interacting with atoms in the interstellar medium produce pions that decay into photons and neutrinos as already explained in section 3.2.4, the combination of SNR with the high density of atoms in molecular clouds increases the probability of such an interaction compared to less dense parts of the interstellar medium and increases therefore also the probability to detect neutrinos from such sources. As stated in [4] neutrinos from these sources might be detected within several years of data taking with IceCube. As they are also in the field of view of the ANTARES detector, the Milagro sources MGRO J2019+37, MGRO J2031+41, MGRO J1908+06, C1 and C2 are selected as source candidates for a point source analysis. In the following we will have a closer look at the sources, the prediction of their neutrino fluxes and the neutrino simulation.

6.1 The Milagro sources

Since 2007 newer measurements of the gamma ray fluxes of these sources are available and so their locations are also updated due to a larger data sample. In [54] the photon spectrum is reported in the range from 1 TeV to 100 TeV. In the following the five sources are regarded separately and the updated source positions used in this analysis are given.

Object	Location (l, b) (deg)	Error radius	Flux at 20TeV	extent diameter (deg)	possible counterparts
Crab	184.5,-5.7	0.11	10.9 ± 1.2	n/a	Crab
MGRO J2019+37	75.0, 0.2	0.19	8.7 ± 1.4	1.1 ± 0.5	GeV J2020+3658 PWN G75.2+0.1
MGRO J1908+06	40.4,-1.0	0.24	8.8 ± 2.4	$< 2.6°$ (90%CL)	GEV J1907+0557 SN G40.5-0.5
MGRO J2031+41	80.3,1.1	0.47	9.8 ± 2.9	3.0 ± 0.9	GeV J2035+4214 TeV J2031+413
C1	77.5, -3.9	0.24	3.1 ± 0.6	< 2.0 (90%CL)	n/a
C2	76.1,-1.7	n/a	3.4 ± 0.8	n/a	n/a
C3	195.7,4.1	0.40	6.9 ± 1.6	2.8 ± 0.8	Geminga
C4	105.8,2.0	0.52	4.0 ± 1.3	3.4 ± 1.7	GeV J2227+6106 SNR G106.6+2.9

Table 6.1: Galactic sources and source candidates measured by the Milagro observatory, table extracted from [3]. The flux is given in $10^{-15}\,\mathrm{TeV}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

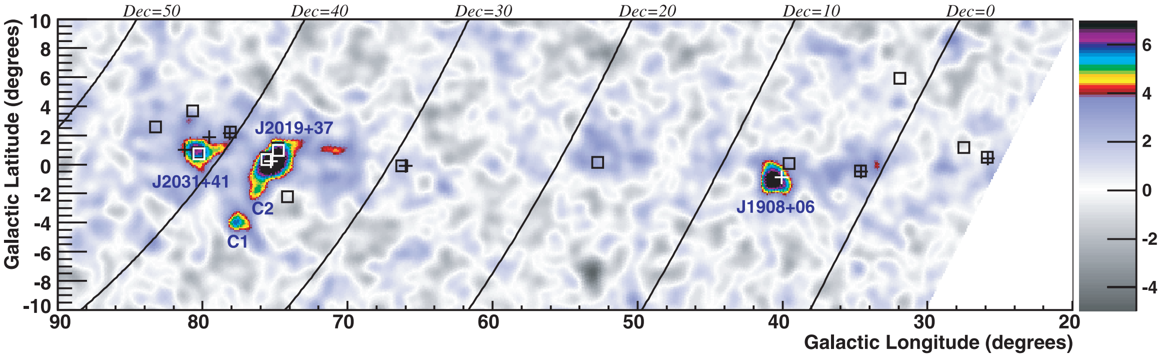


Figure 6.1: Section of the significance map of the Galactic Plane showing the sources MGRO J2019+37, MGRO J2031+41, MGRO J1908+06, C1 and C2 (taken from [3]). The color code shows the pretrial significance, the boxes (crosses) indicate the location of the EGRET 3G (GeV) sources.

It is necessary to mention that some of these sources are considered as extended in gamma ray astronomy even if their diameters are updated to be less than 1°. But for a neutrino telescope with an angular resolution of ca. 0.3°, sources with such an extention can be regarded as point sources.

MGRO J2019+37

MGRO J2019+37 is the brightest source in the Milagro sample after the Crab. It is associated with the EGRET sources 3EG J2016+3657 and 3EG J2021+3716 [2]. A two-dimensional Gaussian fit gives an extention of $\sigma = 0.32^\circ \pm 0.12^\circ$ [2]. It is associated with the young pulsar OFGL J2020.8+3649 [55]. In the analysis [54] the spectrum is fitted well with a power-law with an exponential cut-off (see below), where the normalization of the photon spectrum at 10 TeV is $7 \cdot 10^{-10}\,\mathrm{s}^{-1}\,\mathrm{m}^{-2}\,\mathrm{TeV}^{-1}$ with a spectral index of 2.0 and a cut-off energy of 29 TeV. In the same analysis also the coordinates were updated as listed in Tab. 6.2.

MGRO J2031+41

This source is located in an area of the Cygnus region with the largest concentration of atomic gas and has an extension of $3.0^\circ \pm 0.9^\circ$ [3]. Because of the reported extension and the consistency of the location with the source TeV J2031+413, whose measured flux is only about one-third of the flux reported by Milagro, it is assumed that MGRO J2031+41 consists of more than one TeV source [3]. This source could be associated with the pulsar OFGL J2032.2+4122 [55]. In the newer analysis from 2012 the coordinates (see Tab. 6.2) and photon flux parameters are updated. The best fit parameters are found for a power-law without cut-off with a normalization $2.1 \cdot 10^{-10}\,\mathrm{s}^{-1}\,\mathrm{m}^{-2}\,\mathrm{TeV}^{-1}$ at 10 TeV and spectral index of 3.22 [54].

MGRO J1908+06

The Milagro data for this source are consistent with both a point source and an extended source in the gamma ray picture. HESS could confirm this source and set an diameter of 0.34° [56]. The coordinates are updated to

$$l = 40^\circ 23^{'} 9.2^{''} \pm 2.4^{'}\ ,\ b = -0^\circ 47^{'} 10.1^{''} \pm 2.4^{'} [56].$$

These galactic coordinates are transformed into equatorial coordinates which can be found in Tab. 6.2. The source is still unidentified [57], but ARGO-YBJ supports that it is the pulsar wind nebular of PSR J1907+0602 [58].

C1 and C2

The candidate sources C1 and C2 have significance less then 4.5σ, but they are located in the Cygnus star-forming region, which renders them possible neutrino sources. Both cannot be associated with PWN. Milagro states that C2 could be an extension of MGRO J2019+37, but it is seperated by 2.2° [3]. The coordinates of these two sources are given in Tab. 6.2, where the galactic coordinates given in Tab. 6.1 are transformed into equatorial coordinates, which are necessary for the point source analysis and the simulation of neutrinos from these point sources.

source	R.A. (deg)	Decl. (deg)
MGRO J1908+06	286.97	6.26
MGRO J2019+37	304.63	36.88
MGRO J2031+41	307.18	41.31
C1	310.97	36.30
C2	307.74	36.51

Table 6.2: Summary of the equatorial coordiantes of the sources MGRO J1908+06 (taken from [56]), MGRO J2019+37 (taken from [54]), MGRO J2031+41 (taken from [54]), C1 and C2 (taken from [3]).

source	k_γ $[\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	α_γ	$E_{cut,\gamma}$[TeV]
MGRO J1908+06	$3.4 \cdot 10^{-12}$	2.01	300
MGRO J2019+37	$3.5 \cdot 10^{-12}$	2.00	300
MGRO J2031+41	$2.4 \cdot 10^{-12}$	2.00	300
C1 (MGRO J2043+36)	$1.7 \cdot 10^{-12}$	2.00	300
C2 (MGRO J2032+37)	$1.3 \cdot 10^{-12}$	2.00	300

Table 6.3: Gamma ray flux parametrization of the Milagro sources determined in [4], exact values recieved from [59].

6.2 Neutrino flux predictions

The neutrino flux for a given gamma ray spectrum can be calculated following the method explained in [16]. This method is used in [4] and [5] to determine the neutrino flux assumptions for the five strong Milagro sources. The result of this application is given in the following.

The gamma ray flux is parametrized as a power-law with spectral index α_γ, a normalization factor k_γ and an exponential cut-off energy $E_{cut,\gamma}$

$$\frac{dN_\gamma(E_\gamma)}{dE_\gamma} = k_\gamma \left(\frac{E_\gamma}{\mathrm{TeV}}\right)^{-\alpha_\gamma} \exp\left(-\sqrt{\frac{E_\gamma}{E_{cut,\gamma}}}\right), \tag{6.1}$$

for a purely power-law without cut-off energy, one can approximate $E_{cut,\gamma} \to \infty$. Then the muon neutrino flux on earth after oscillation can be derived as

$$\frac{dN_\nu(E_\nu)}{dE_\nu} = k_\nu \left(\frac{E_\nu}{\mathrm{TeV}}\right)^{-\alpha_\nu} \exp\left(-\sqrt{\frac{E_\nu}{E_{cut,\nu}}}\right), \tag{6.2}$$

where the neutrino flux parameters k_ν, α_ν and $E_{cut,\nu}$ are linked to the gamma ray flux parameters as (from [5])

- $k_\nu = (0.694 - 0.16\alpha_\gamma)k_\gamma$
- $\alpha_\nu = \alpha_\gamma$
- $E_{cut,\nu} = 0.59 E_{cut,\gamma}$.

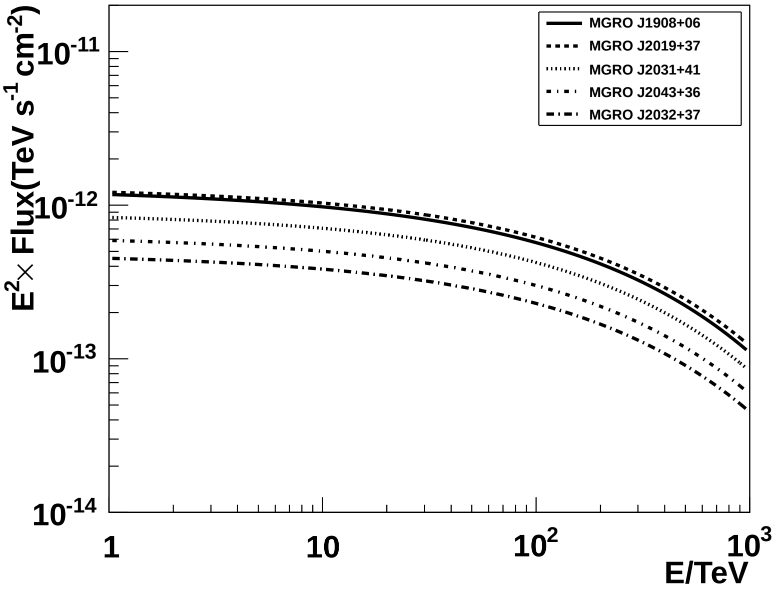


Figure 6.2: Neutrino flux of five Milagro sources with the parametrization of [4] (see Tab. 6.3).

source	K [TeV^{-1}cm^{-2}s^{-1}]	E_{norm} [TeV]	$\alpha_{\gamma,i}$	$E_{cut,\gamma,i}$[TeV]
MGRO J2019+37	$7\cdot10^{-14}$	10	1.8, 2.0, 2.2	30, 100, 300
MGRO J1908+06	$6.1\cdot10^{-13}$	4	1.9, 2.1, 2.3	30, 100, 300
MGRO J2031+41	$2.1\cdot10^{-14}$	10	2.6, 2.8, 3.0	30, 100, 300

Table 6.4: Gamma ray flux parametrization of the Milagro sources determined in [5], where $k_\gamma = \frac{K}{E_{norm}^{-\alpha_\gamma}}$.

The parameters k_γ, α_γ and $E_{cut,\gamma}$ for the Milagro sources and source candidates considered in [4] are summarized in Tab. 6.3. With this parameters the neutrino fluxes are calculated using eqn. (6.2) and are shown in Fig. 6.2. In the following these neutrino flux assumptions are named kappes to distinguish them from the assumptions made in [5].

As already explained in the previous section, there are newer measurements of the gamma ray spectrum available for these three Milagro sources: MGRO J2019+37, MGRO J1908+06 and MGRO J2031+41 [54],[58]. As they are using a different parametrization of the gamma ray flux the transformed parameters for eqn. (6.1) are calculated in [5]. The new parameters are given in Tab. 6.4 extracted from Tab. 2 of [5] with $k_\gamma = \frac{K}{E_{norm}^{-\alpha_\gamma}}$. The calculated neutrino fluxes for three different spectral indices $\alpha_{\gamma,i}$ around the best fit value and three cut-off energies $E_{cut,\gamma,i}$ is shown in Fig. 6.3, for a comparison also the spectra determined in [4] for these sources is plotted (solid line). As there are several neutrino flux parametrizations available for the sources MGRO J2031+41, MGRO J1908+06 and MGRO J2019+37 the nine different parametrizations from Tab. 6.4 are named halzen1 to halzen9. A detailed mapping of names to parameters can be found in the appendix A.1.

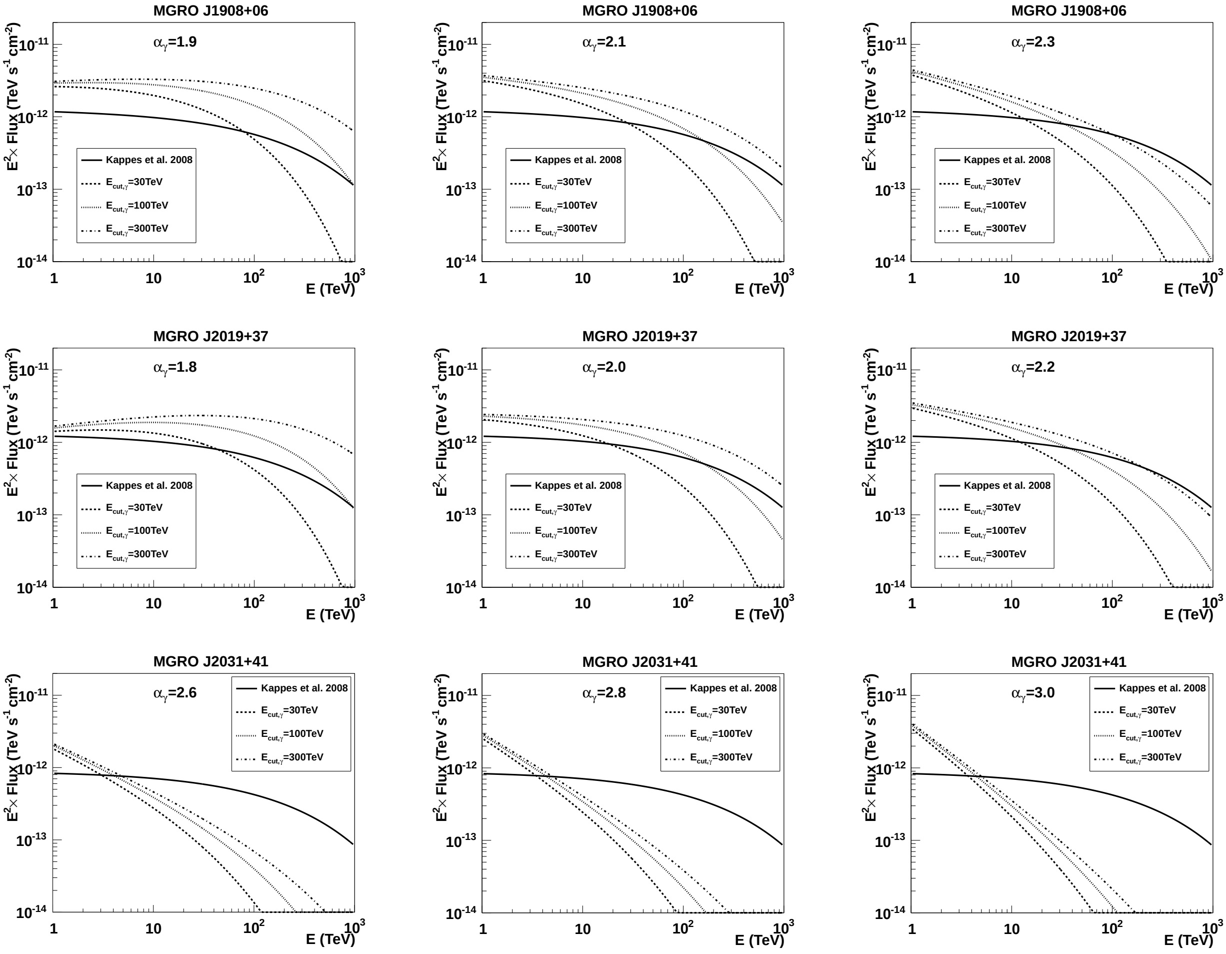


Figure 6.3: Neutrino fluxes calculated with the values of Tab. 6.4 for the sources MGRO J1908+06, MGRO J2019+37 and MGRO J2031+41 (top, middle, bottom row respectively) are shown from left to right for different values of α_γ and for different energies $E_{cut,\gamma}$. As solid line the neutrino flux parametrization of [4] (from Tab. 6.3) is shown for these sources.

6.3 Simulation of neutrinos for each Milagro source

In ANTARES neutrinos are simulated using an so called run-by-run simulation scheme (rbr) as the optical background rates and detector conditions can change in small time scales (of the order of hours) due to the living environment the Mediterranean Sea and caused by environmental factors as the sea current for example. This means that for every data taking run neutrinos distributed over the whole sky are simulated. The optical background and data taking conditions for that particular run are taken into account for the simulation of the detector response. So the single rates of each OM and its working status can be mapped in the simulation process. Weighting the resulting neutrino events afterwards according to a cosmic or atmospheric neutrino flux gives the number of expected events.

For the simulation of neutrinos coming from one single point source in the sky the declination has to be fixed in the first step of the simulation software. This step is not time consuming, but simulating all possible detector conditions in the simulation process of the detector response for this single source would take about 10 CPU-years. Therefore the official rbr scripts have been adapted. The idea is to take a small subset of runs, that describe the global data taking conditions well and use them for the simulation of the detector conditions.

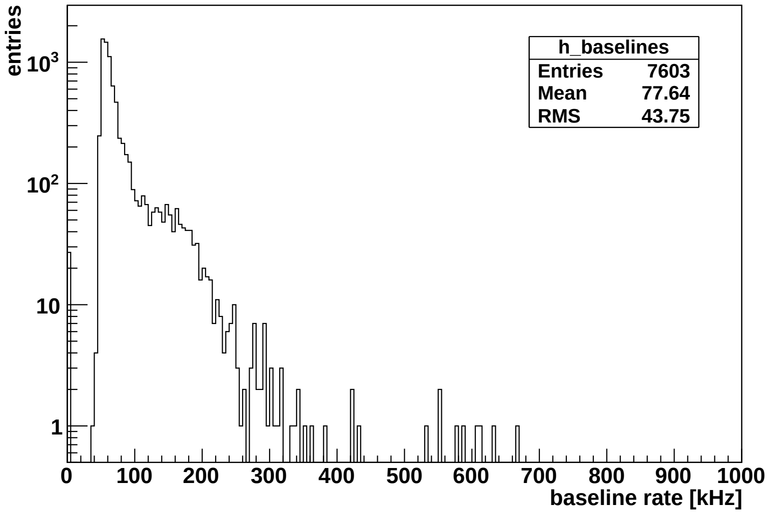


Figure 6.4: Distribution of the baseline rates for good runs from 12-line-data-taking period from 2008 to 2012. For runs where the basline rate could not be determined the baseline rate is set to 0 kHz.

Possible criteria are the run duration, the mean baseline rate of the OMs, which is a quantity describing the optical background conditions, and the number of working OMs describing the status of the detector. The distribution of the latter two quantities are shown in Fig. 6.4 and Fig. 6.5 for the 12-line-data-taking period from 2008 to 2012, for runs qualified as good (see section 5.1). From Fig. 6.4 one can extract that the mean value of the baseline rate 77.6 kHz lies in the main peak of the distribution. So this value with an uncertainty of 10% can be taken as one criteria. The distribution of the number of active OMs (Fig. 6.5) shows a big dip at about 650 OMs, which is the mean of this distribution. In this case the mean value is not a

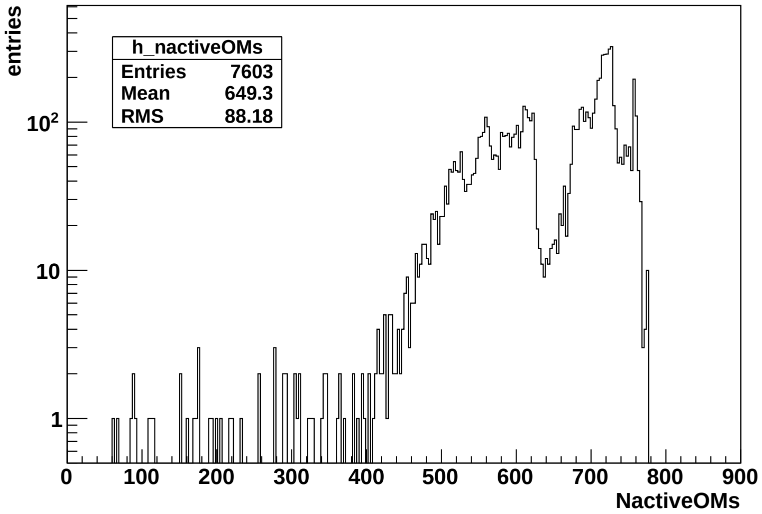


Figure 6.5: Distribution of the number of active OMs for good runs from 12-line-data-taking period from 2008 to 2012.

good choice as selection criteria. So the decision is to take runs with 700 active OMs with an uncertainty of again 10% which roughly describes the right peak of the distribution and stands for good detector conditions with 12 working lines. So in the end the following parameters were used to identify possible runs for the background and detector conditions for the simulation of the detector response:

- baseline rate $(77,6 \pm 10\%)\,\text{kHz}$
- nactiveOMs $700 \pm 10\%$
- run duration > 100 min.

All runs fulfilling these criteria were taken 2012, where all 12 lines were installed and running stable after single outages of lines (representing the dip in Fig. 6.5) and their respective repair the years before. From these runs 10 runs are selected and defined as standard conditions. To study the systematic effect of changing the detector conditions additionally five conditions are defined (see appendix A.3 for the run numbers), where one time the baseline rate changes and the number of active OMs is constant and the other way around:

- Different baseline rates:
  - cond1: baseline rate $(100 \pm 10\%)\,\text{kHz}$, nactiveOMs $700 \pm 10\%$
  - cond2: baseline rate $(150 \pm 10\%)\,\text{kHz}$, nactiveOMs $700 \pm 10\%$
  - cond3: baseline rate $(200 \pm 10\%)\,\text{kHz}$, nactiveOMs $700 \pm 10\%$

- Different nactiveOMs:
    - cond4: baseline rate $(77,6 \pm 10\%)$ kHz, nactiveOMs $600 \pm 10\%$
    - cond5: baseline rate $(77,6 \pm 10\%)$ kHz, nactiveOMs $500 \pm 10\%$

For each of the five Milagro sources neutrinos and anti-neutrinos are simulated in separate processes with the official rbr simulation software using the 10 runs defined for the specified conditions. For each condition and source $100 \cdot 10^{10}$ neutrinos respectively anti-neutrinos are generated in the energy range 10 GeV to 10^8 GeV. All simulated neutrino events are then processed with the SeaTray framework using the same calibrations and muon track and energy reconstructions as in the official data and rbr production to be consistent. The detector response and the visibility of the source is included by applying a cut of only as upgoing reconstructed events.

Weighting of events according to the neutrino flux assumption

In the simulation process neutrino events are simulated according to a relatively flat neutrino spectrum (in this case $E^{-1.4}$), so that the statistics in the high energy range is high enough. Every simulated neutrino is stored with a parameter $w2$, which can be used to re-weight the events according to a chosen neutrino flux $dN_\nu/dE_\nu(E_\nu)$ such as $w = \frac{w2}{\text{nGenEvts}} \cdot dN_\nu/dE_\nu(E_\nu)$, where nGenEvts is the total number of generated events. As seen in section 6.2 the neutrino flux parametrization is given in the dimensions $\text{TeV}^{-1}\text{cm}^{-2}\text{s}^{-1}$. The weight $w2$ has the dimensions $\text{GeV}\,\text{m}^2\,\text{s}\,\text{year}^{-1}$ in point source mode (fixed declination). So the weight per event w is then given by

$$w = \frac{w2}{\text{nGenEvts}} \cdot 10.0 \cdot 0.5 \cdot \frac{dN_\nu}{dE_\nu}(E_\nu/\text{TeV}) \tag{6.3}$$

and has the dimension year^{-1}. The factor 10.0 comes from the transformation of the dimensions of $dN_\nu/dE_\nu(E_\nu)$ from

$$\text{TeV}^{-1}\text{cm}^{-2}\text{s}^{-1} = 10^{-3}\text{GeV}^{-1} \cdot 10^{4}\text{m}^{-2}\text{s}^{-1} = 10.0 \cdot \text{GeV}^{-1}\text{m}^{-2}\text{s}^{-1}. \tag{6.4}$$

The factor 0.5 is needed because $dN_\nu/dE_\nu(E_\nu)$ is the total neutrino flux combining neutrinos and anti-neutrinos. As neutrinos and anti-neutrinos are simulated separately and added up afterwards, the flux that has to be used for the weighting is only half of the total neutrino flux. To determine the number of expected events the weights w from eqn. (6.3) have to be added up for all neutrino events surviving some chosen cuts and multiply this with the live-time of the data examined.

Chapter 7

Maximum-likelihood-based point source analysis method

After the cut of upgoing reconstructed events one cannot distinguish between cosmic neutrinos, atmospheric neutrinos or misreconstructed atmospheric muons. If we want to study neutrinos from point-like sources in the sky we have to apply a known statistical concept: hypotheses testing. The null hypothesis is that the measured data only consists of background events (atmospheric neutrinos and misreconstructed atmospheric muons) and compares this with the signal hypothesis of cosmic neutrinos. For this purpose a likelihood function with a signal and a background component is defined and maximized in order to find the number of signal events contained in that data sample. In this chapter the maximum likelihood point source analysis method is explained in detail. This method is in the following used to study the sensitivity for the five strong Milagro sources described in the previous chapter.

7.1 Maximum likelihood function and test statistic calculation

After measuring N events in the data, the task is to find the number of signal events n_s present within this set of events that are connected to the point source. As each event has a certain probability being signal- or background-like the total likelihood L is defined as product of the single probabilities for all events under the assumption having detected n_s signal events $L = \prod_{i=1}^{N} P_i(n_s)$. The assumption is that the likelihood is maximal for the n_s signal events present in the data sample. So the value of n_s that maximizes the likelihood is the best choice. This method is known as the maximum likelihood method [60]. Technically it is easier to maximize the logarithm of likelihood instead of the likelihood itself, which transforms then to

$$\log L(n_s) = \log \prod_{i=1}^{N} P_i(n_s) = \sum_{i=1}^{N} \log(P_i(n_s)) = \sum_{i=1}^{N} \log \left(\frac{n_s}{N} S_i + \left(1 - \frac{n_s}{N}\right) B_i \right). \tag{7.1}$$

S_i is the signal PDF (probability density function), B_i is the background PDF, N is the total number of events, n_s is the estimated number of signal events. The overall integral of the probability function has to be equal to 1 [60]. This is the case if both $\int S_i(x)\mathrm{d}x = 1$ and

$\int B_i(x)\mathrm{d}x = 1$, as shown here

$$\begin{aligned}\int P_i(n_s)\mathrm{d}x &= \int \frac{n_s}{N} S_i(x) + \left(1 - \frac{n_s}{N}\right) B_i(x)\mathrm{d}x \\ &= \frac{n_s}{N} \underbrace{\int S_i(x)\mathrm{d}x}_{=1} + \left(1 - \frac{n_s}{N}\right) \underbrace{\int B_i(x)\mathrm{d}x}_{=1} = 1.\end{aligned} \tag{7.2}$$

Therefore both signal and background PDF have to be normalized before starting the maximization process. In section 7.2 it is described which signal and background PDF are used in this analysis and it is explained how they are determined.

To decide if the measured events are signal- or background-like, the hypotheses testing method is used. The null hypothesis is that the data sample consists of only background events (with the likelihood $L(n_s = 0)$). The alternative hypothesis is that, additional to the background events, signal events are present in the data sample. A test statistic Q is then the fraction of the likelihood for signal and background divided by the likelihood for only background. Using the logarithm, it can again be transformed to the difference of the likelihood with signal assumption minus the likelihood for the background only case, as derived here

$$Q = \log\left(\frac{\max(L(n_s))}{L(n_s = 0)}\right) = \max(\log L(n_s)) - \underbrace{\log L(n_s = 0)}_{=\text{const.}} = \max(\log L(n_s) - \log L(n_s = 0)). \tag{7.3}$$

The smaller the value of the test statistic Q the more probable it is that the null hypothesis is true. To determine the necessary values for Q ($Q_{3\sigma}^{threshold}$, $Q_{5\sigma}^{threshold}$) to exclude the null hypothesis with a significance of 3σ and 5σ respectively, pseudo experiments (PEs) are performed for the background only case (see section 7.3 and 7.4.1). Additionally pseudo-experiments are performed for signal- and background-like data samples varying the number of signal events n_{sig} from 1 to 9. This is necessary to determine the value of n_s needed for a 50% chance for a 3σ and a 5σ discovery respectively . With this information the sensitivity for the single Milagro sources can be calculated assuming events detected with ANTARES and reconstructed as upgoing with KrakeFit. The generation of pseudo-experiments are explained in detail in section 7.3 and their statistical interpretation including the optimization process of the cut values for this analysis, the calculation of the sensitivity and a 90% confidence level upper limit are discussed in section 7.4.

source	R.A. (deg)	Decl. (deg)
MGRO J1908+06	286.97	6.26
MGRO J2019+37	304.63	36.88
MGRO J2031+41	307.18	41.31
C1	310.97	36.30
C2	307.74	36.51

Table 7.1: Summary of the equatorial coordinates of the sources MGRO J1908+06 (taken from [56]), MGRO J2019+37 (taken from [54]), MGRO J2031+41 (taken from [54]), C1 and C2 (taken from [3]).

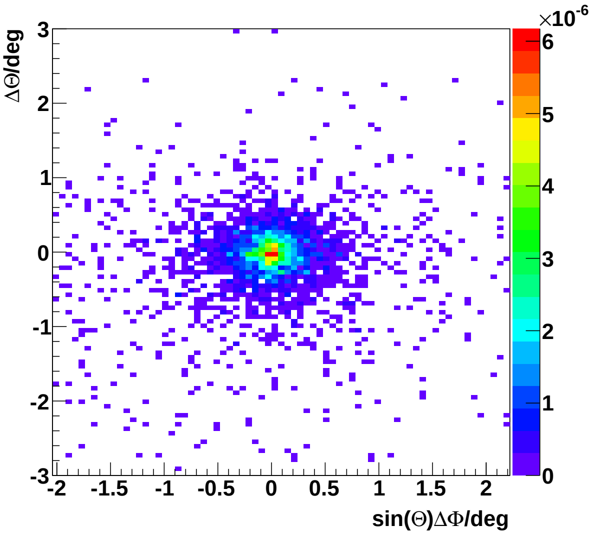


Figure 7.1: Two-dimensional distribution of the error in Θ and Φ for reconstructed muon events around the coordinates of the source C1.

7.2 Signal and background probability density functions

For the search of neutrinos from the point-like five Milagro sources discussed in chapter 6 the signal PDFs for the maximum-likelihood-based analysis have to be determined for each source and flux assumption separately. This has to be done as the resolution of the reconstructed neutrino depends on the location at the sky. Additionally the visibility of the sources (see appendix A.2) have to be included to get the number of expected neutrino events per source for the live-time of the studied data sample. The background rate depends also on the zenith angle. As the source is visible at different zenith angles within the detector, it is better to use a declination dependent background PDF than a zenith dependent one. Only events reconstructed as upgoing with the muon track reconstruction strategy KrakeFit are considered for the generation of the PDFs after different quality cuts.

7.2.1 Signal PDF

The signal PDF has to be determined for each Milagro source, flux assumption and applied quality cuts separately. The calculation is described in the following. Neutrinos are simulated for each source and expected neutrino flux separately with standard detector conditions as explained in section 6.3. A summary of the sources and their locations used in this analysis can be found in Tab. 7.1. The different flux predictions are named from halzen1 to halzen9 (if available) and kappes. A detailed mapping from names to parameters can be found in appendix A.1.

The point spread function (PSF) is used as signal PDF. To determine the PSF, the distribution of the event rate per year $N(\beta)$ is determined for the simulated neutrino events that pass the applied cuts and weighted according to their expected neutrino flux spectrum (see eqn. (6.3)). β is the space angle between reconstructed muon and source location in local detector coordinates. To check if the reconstructed muon events are distributed around the source following a circular symmetry a two-dimensional distribution of the error in Θ and Φ in local detector coordinates is shown for C1 in Fig. 7.1 which shows the expected behaviour.

As the event rate $dN/d\beta$ per space angle bin has to be transformed to $dN/d\Omega$ the following transformation has to be used to get the PSF $dN/d\Omega$:

$$\frac{dN}{d\Omega} = \frac{dN}{d\beta}\frac{d\beta}{d\Omega} = \frac{dN}{d\beta}\frac{1}{2\pi\sin(\beta)}. \tag{7.4}$$

Technically the enclosed solid angle Ω is determined by integrating over the bins, where r_1 and r_2 are the lower and the higher bin boundaries

$$\Omega = \int_{r_1}^{r_2} \mathrm{d}\beta \sin\beta = \cos r_1 - \cos r_2. \tag{7.5}$$

For each bin of the distribution $N(\beta)$ one has to divide N per bin by $2\pi \cdot (\cos r_1 - \cos r_2)$ to get the histogram of the PSF. To get a better resolution for very small angular errors the PSF histogram is determined as a function of $\log_{10}(\beta)$. Motivated by the parametrization of the PSF used in [17] the function fitted to the histogram is defined as

$$\log_{10} S(x) = \log_{10}\frac{dN}{d\Omega}(x) = \begin{cases} A & x < x_0 \\ A + B\left(-1 + \exp\left(\frac{(x-x_0)^2}{2\sigma^2}\right)\right) & x \geq x_0, \end{cases} \tag{7.6}$$

with $x = \log_{10}\beta$ and x_0 is the threshold for which the PSF is expected to be constant. An example for a PSF histogram with fitted function can be found in Fig. 7.2 left. The PSF function $S(\beta)$ is then derived from eqn. (7.6) as

$$S(\beta) = 10^{\log_{10} S(\log_{10}\beta)}. \tag{7.7}$$

To control if the transformation was correct the fitted function was transformed back into the β-distribution, formally written as

$$S'(\beta) = \sin(\beta) \cdot S(\beta) = \sin(\beta) \cdot 10^{\log_{10} S(\log_{10}\beta)}. \tag{7.8}$$

It is compared to the distribution of β (see Fig. 7.2 right), where the integral of $S'(\beta)$ over the corresponding bin is shown. Both distributions fit well within their respective uncertainties with a small systematic change to lower β for $S'(\beta)$.

Normalization of the signal PDF

The signal PDF is defined as $S_i(x) = \frac{f_i(x)}{2\pi\sin x}$ which has to be normalized so that $\int S_i(x)d\Omega = 1$. The integration over Ω gives in the unnormalized function:

$$\begin{aligned} \int S_i(x)\mathrm{d}\Omega = \int \frac{f_i(x)}{2\pi\sin x}\mathrm{d}\Omega &= \int\int \frac{f_i(x)}{2\pi\sin x}\mathrm{d}\phi\sin x\mathrm{d}x \\ &= \frac{1}{2\pi}\cdot 2\pi\int f_i(x)\mathrm{d}x = \int f_i(x)\mathrm{d}x \\ &= N. \end{aligned} \tag{7.9}$$

So it has to be divided by the number of events N.

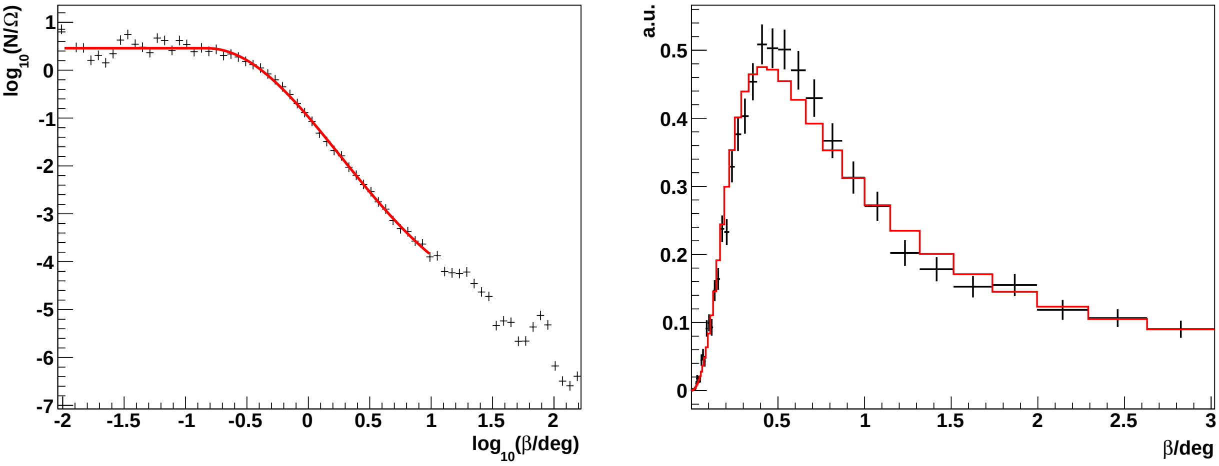


Figure 7.2: Left: PSF for C1 with flux kappes and a rlogL cut of 5.5, the red line is the fitted PSF function defined in eqn. (7.6), both are not normalized. Right: distribution of β (black) and the integral over the transformed PSF function $S'(\beta)$ defined in eqn. (7.8) per bin (red), both are normalized.

7.2.2 Background PDF

The background rate which is the number of atmospheric neutrinos and muons reconstructed as upgoing, after applying quality cuts, depends on the zenith angle. The azimuthal distribution of the background rate is more or less flat. As already mentioned above, the source position in local detector coordinates changes with time, so we assume the background rate depends on the declination δ instead of the zenith angle and use this as background PDF which makes it easier to perform the pseudo-experiments later on. The background rate is defined as $dN/d\sin(\delta)$, where one example is shown in Fig. 7.3 for upgoing events with rlogL < 5.5, which is derived from data assuming no signal[1]. As the data events are blinded[2] and thus the actual position in equatorial coordinates of the event is not known, the declination is calculated by taking the reconstructed coordinates of the event in local detector coordinates and transforming them with a randomly taken time into the equatorial system. Taking this histogram $N(\sin\delta)$ as a basis for the background PDF, a smooth behaviour is necessary for the maximization process of the likelihood. Therefore two spline curves are fitted once to the original histogram (blue curve) and once to the smoothed histogram (red curve), see Fig. 7.3. Because of the smaller fluctuations the red curve is used for the maximization of the likelihood.

To get the background PDF $\frac{dN}{d\Omega} = \frac{1}{2\pi}\frac{dN}{d\sin(\delta)}$ one has to divide the background rate by 2π and normalize the histogram for the maximum likelihood analysis.

---

[1] Being specific, theoretically signal events are included in this histogram, but as the number of signal events is very small compared to the number of background events and as the data events are additionally randomized, these do not contribute significantly to this distribution and are negligible.

[2] The data events are blinded means that the absolute time of the event is not known and scrambled, so that it is not possible to know to exact position of the event at the sky before the optimization of an analysis is finished. After the collaboration decides to 'unblind the data, the absolute time of the event is taken.

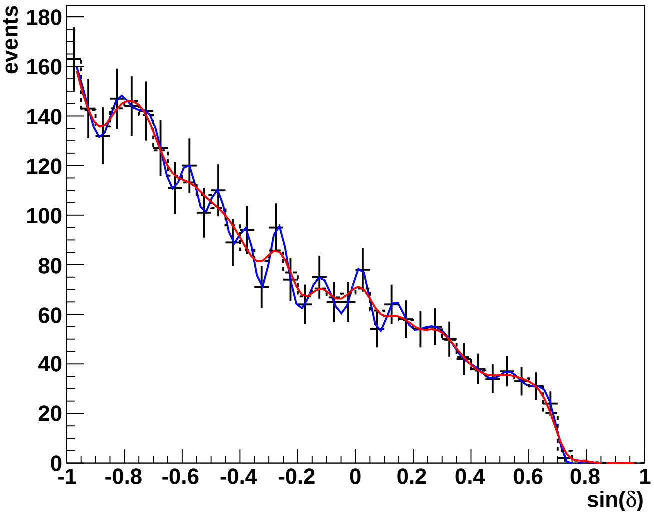


Figure 7.3: Distribution of the number of events per sin(δ) for events reconstructed as upgoing with KrakeFit with rlogL < 5.5 for data (black), spline to the distribution (blue), smoothed distribution (dotted histogram), spline to smoothed distribution (red).

Normalization of the background PDF

For the generation of the pseudo-experiments (see section 7.3) one needs to know the total number of events N_{tot} after cuts, which can be extracted direcly from the histogram. Further for the random generation of background events the distribution itself doesn't have to be normalized. Therefore the background rate histogram is not normalized, but one has to know the normalization factor for the calculation of the likelihood later on. The normalized background PDF gets

$$B(\delta) = \frac{1}{2\pi} \cdot \frac{N(\sin\delta)}{N_{\mathrm{tot}}}, \tag{7.10}$$

which is proven here

$$\begin{aligned} \int B(\delta)\mathrm{d}\Omega &= \int\int \frac{1}{2\pi} \cdot \frac{N(\sin\delta)}{N_{\mathrm{tot}}} \mathrm{d}\phi \mathrm{d}\sin\delta \\ = \frac{2\pi}{2\pi \cdot N_{\mathrm{tot}}} \int N(\sin\delta)\mathrm{d}\sin\delta &= \frac{1}{N_{\mathrm{tot}}} \cdot N_{\mathrm{tot}} = 1. \end{aligned} \tag{7.11}$$

Also if the spline is used as PDF, it is better to use the number of events as normalization rather than the integral computed from the spline, the latter strongly depends on the number of bins used to describe the spline. The total number of data events N_{tot} left after applying different rlogL cuts with and without applying additionally the rdf=1 cut is summarized in Tab. 7.2.

7.2.3 Definition of the likelihood function

The likelihood function defined in eqn. (7.1) is summarized here incorporating the signal and background PDF determined in this section. The point spread function (PSF) is used as signal

rlogL	5.0	5.1	5.2	5.3	5.4	5.5	5.6
without rdf	510	629	889	1282	1761	2786	5923
with rdf=1	403	606	850	1207	1580	2193	3518

Table 7.2: Summary of the total number of data events N_{tot} after applying different rlogL cuts and respectively additionally applying the rdf=1 cut.

PDF depending on the space angle β between reconstructed muon and source location and the parametrization of the background rate depending on the declination δ, precisely $\sin\delta$, is used as background PDF. The likelihood used in this point source analysis with included normalization as derived in this section is

$$\log L(n_s) = \sum_{i=1}^{N} \log\left(\frac{n_s}{N} S_i(\beta) + \left(1 - \frac{n_s}{N}\right) B_i(\delta)\right). \tag{7.12}$$

7.3 Pseudo-experiments

7.3.1 Generating pseudo-experiments

To get a statistical interpretation of the measurement of highly energetic neutrinos from the five Milagro sources pseudo-experiments are necessary, as the counting rate of events is low. The idea is to simulate a fixed number of events $N = n_b + n_{sig}$ where n_b events are simulated according to background and n_{sig} events according to signal distributions and calculate the test statistic Q and determine the corresponding n_s for this set of events, which is defined as one pseudo-experiment. We repeat it N_{PE} times and get distributions of Q and n_s for one set (n_b, n_{sig}) of pseudo-experiments. A special case is the background only case, where only background events are simulated. 10^5 background only PEs and 10^4 signal-like PEs with $(0 < n_{sig} \leq 9)$ are generated. The total number of events simulated in one pseudo-experiment is fixed to the number of data events detected and reconstructed as upgoing $N = N_{tot}$, which depends only on the quality cuts applied (see Tab. 7.2). If signal events are added to the sample, the number of background events is reduced by that number $n_b = N - n_{sig}$.

Background events are drawn from the smooth distribution of $N(\sin\delta)$, whereas the signal events are drawn from the β-distribution $N(\beta)$. The ingredients for the likelihood computation of eqn. (7.12) are the declination δ and the angle between source location and the reconstructed event β. To compute the likelihood function one needs the declination and space angle information for both signal and background events. As only one of these parameters is available directly from the pseudo-experiment the next subsections explain the determination of the missing parameters.

Generation of signal events

The signal events are drawn from the distribution $N(\beta)$ (see Fig. 7.2 right as an example) up to an angle of $\beta < 10°$. The distance β to the source with coordinates (δ_s, α_s) is known, but not the declination of the drawn event. Therefore the coordinates of the event (δ_e, α_e) have to be determined for an angular distance β between source and event.

Assuming a coordinate transformation such that the vector of the source points in z-direction $(0, 0, 1)$, the possible coordinates of the event lie on a cone with opening angle β (see Fig. 7.4).

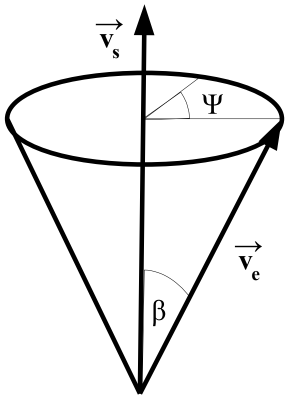


Figure 7.4: Possible event directions $\vec{v}_e$ with angular distance β to the source direction $\vec{v}_s$, Ψ is the azimuthal angle.

The azimuthal angle Ψ is free and is therefore randomly chosen between 0 and 2π. Using cartesian coordinates the event vector is $\vec{v}_e = (\sin\beta\cos\Psi, \sin\beta\sin\Psi, \cos\beta)$, which has to be transformed back by rotating by $90° - \delta_s$ around the y-axis and by rotating by α_s around z-axis. Then the coordinates of the event are ($\delta_e = \sin^{-1} z$, $\alpha_e = \mathrm{atan}\,\frac{y}{x}$). This is done according to the procedure used in [61]. To proof its consistency, the source coordinates itself are taken and transformed back:

$$\begin{aligned}
&\begin{pmatrix} \cos\alpha_s & -\sin\alpha_s & 0 \\ \sin\alpha_s & \cos\alpha_s & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos(90° - \delta_s) & 0 & \sin(90° - \delta_s) \\ 0 & 1 & 0 \\ -\sin(90° - \delta_s) & 0 & \cos(90° - \delta_s) \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\
&\qquad = \begin{pmatrix} \cos\alpha_s & -\sin\alpha_s & 0 \\ \sin\alpha_s & \cos\alpha_s & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \sin(90° - \delta_s) \\ 0 \\ \cos(90° - \delta_s) \end{pmatrix} \\
&= \begin{pmatrix} \cos\alpha_s \sin(90° - \delta_s) \\ \sin\alpha_s \sin(90° - \delta_s) \\ \cos(90° - \delta_s) \end{pmatrix} = \begin{pmatrix} \cos\alpha_s \cos(\delta_s) \\ \sin\alpha_s \cos(\delta_s) \\ \sin(\delta_s) \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}
\end{aligned} \tag{7.13}$$

So the backwards transformed coordinates are

$$\sin^{-1} z' = \sin^{-1}(\sin\delta_s) = \delta_s, \tag{7.14}$$

$$\mathrm{atan}\left(\frac{y'}{x'}\right) = \tan^{-1}\left(\frac{\sin\alpha_s\cos\delta_s}{\cos\alpha_s\cos\delta_s}\right) = \tan^{-1}(\tan\alpha_s) = \alpha_s, \tag{7.15}$$

which are the original coordinates. For another consistency check the angular distance of source and drawn event is recalculated after the coordinate transformation with eqn. (7.16) and is found to be the same.

Generation of background events

The declination δ_e of the event is drawn from the background rate distribution $N(\sin\delta)$, its right ascension α_e is taken randomly between 0 and 2π. The angle between source and event β in equatorial coordinates is then calculated by [62]

$$\beta = \text{acos}\left(\cos(\delta_s)\cdot\cos(\delta_e)\cdot\cos(\alpha_s-\alpha_e)+\sin(\delta_s)\cdot\sin(\delta_e)\right). \tag{7.16}$$

Assumption $\beta_{\text{local}} = \beta_{\text{eq}}$

The signal PDF is calculated by simulating a point source with fixed declination (see sections 6.3 and 7.2.1), so a declination band is used, where the right ascension parameter is free. Each simulated source event has coordinates (δ_s, α_s) that are transformed into local coordinates (Θ_s,Φ_s) and stored only in the local coordinate system per event[3]. As the events are reconstructed in local detector coordinates, the angle β is calculated in local coordinates. The signal events are drawn from that β-distribution by generating the pseudo-experiments, but it is used as the distance in the equatorial coordinate system in the maximization process of the likelihood, which is not perfectly the same. It is studied here if this assumption can be used.

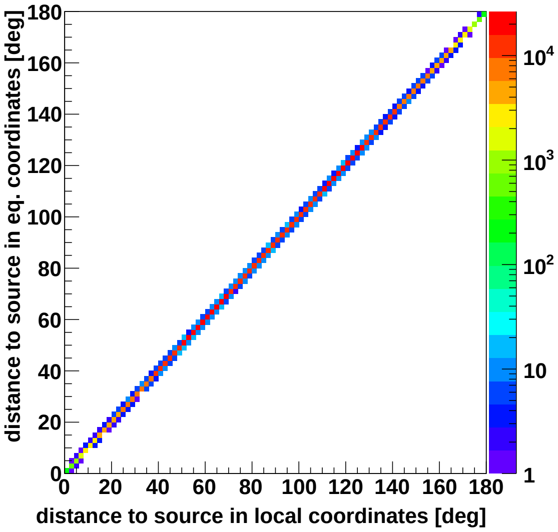


Figure 7.5: Distance to source MGRO J1908+06 for a given time in the local and the equatorial coordinate system for randomly chosen local directions.

For one source the equatorial coordinates are transformed into detector coordinates for a given time t. For that fixed time, track directions are randomly taken over the whole detector ($\Theta \in [0,\pi]$, $\Phi \in [0,2\pi]$). All these directions are transformed back into the equatorial coordinates system. The distance to the source for each direction is calculated before and after the transformation (see Fig. 7.5). The values lie on the bisecting line with small uncertainties. The difference of these distances is much lower than 0.01° independently of the distance in local

---

[3] This is done internally by the simulation software, so it is not possible to recalculate the equatorial coordinates of the event afterwards because of missing information (absolute time of the event).

coordinates. So the assumption that the angular distance in local coordinates is equal to the angular distance in equatorial coordinates is validated with an accuracy better than 0.01°. As the signal PDF is taken as constant for angular distances typically lower than 0.3°, this assumption is valid.

7.3.2 Maximization process of the test statistic

As already mentioned above, N_{tot} events are generated for one pseudo-experiment, where $N_{tot} = n_b + n_{sig}$. All these generated events are used in the computation process of the likelihood, but events farther away than 10° from the source position are regarded as background events ($S_i(\beta > 10^\circ) = 0$) to accelerate the maximization process. This is validated as the fraction of simulated events with $\beta > 10^\circ$ is small after applying quality cuts[4]. Instead of maximizing the likelihood, the test statistic Q is maximized which is equivalent as derived in eqn. (7.3).

$$Q = \max(\log L(n_s) - \log L(n_s = 0)). \tag{7.17}$$

The maximization is done by calculating the Q value for different n_s, where n_s varies in discrete steps between 0 and $n_{s,max} = N_{tot}$ ($n_{s,max} = 30$ if $N_{tot} > 30$), with a maximal injected number of signal events $n_{sig,max} = 9$. The probability to measure a $n_s > 15$ is low (the probability to measure $n_s = 15$ for $n_{sig} = 9$ is about 1.9% due to Poisson statistics), so $n_{s,max} = 2 \cdot 15 = 30$ is taken as upper border to include all possible results of n_s and additionally accelerate the maximization process. The number of steps of the scan is $6 \cdot n_{s,max}$, which gives a step size lower 0.2. Afterwards that n_s value is taken for which Q is maximal. To distinguish between the tested n_s and the result of the maximization process, the parameter μ_s is introduced, which is defined as $\mu_s = n_s$ for which Q is maximal.

To check whether the likelihood computation is correct, an example for the distributions of μ_s for a defined number of injected signal events n_{sig} is shown in Fig. 7.6. The distributions follow roughly a Poisson distribution with mean n_{sig}.

7.4 Statistical interpretation of the pseudo-experiment results

7.4.1 Determining Q-threshold for 3σ and 5σ discovery

An example distribution of Q for the different n_{sig} (called $h_{n_{sig}}(Q)$) is shown in Fig. 7.7. From the normalized distribution of Q for the background only case ($h_0(Q)$), we can extract the Q value for which 3σ, 4σ and 5σ discovery is falsely be claimed, or the null hypothesis is excluded, respectively. The fraction of events above this threshold is defined as p-value $p(n\sigma) = 1 - n\sigma_{\text{fraction}}$. To get to the significance of 5σ with a p-value of $5.7 \cdot 10^{-7}$ (see Tab. 7.3) at least 10^8 pseudo-experiments for background only experiments are necessary for a trial factor (TF) of 1. Technically we are limited by CPU time, so it is not possible to get up to that significance by generating 10^8 background only PEs. The solution is to generate the affordable number of $1.1 \cdot 10^5$ background only PEs ($n_{sig} = 0$) and extrapolate the distribution of $h_0(Q)$, which is done by fitting an exponential decrease to the tail of the distribution which contains 3%[5] of the

---

[4] For MGRO J1908+06 the fraction of events with $\beta > 10^\circ$ is less then 10% for rlogL=5.5 and about 3% for rlogL=5.0.

[5] For higher percentage values the peak at $Q = 0$ would also be fitted by the exponential function, which is not useful for modelling the tail of this distribution.

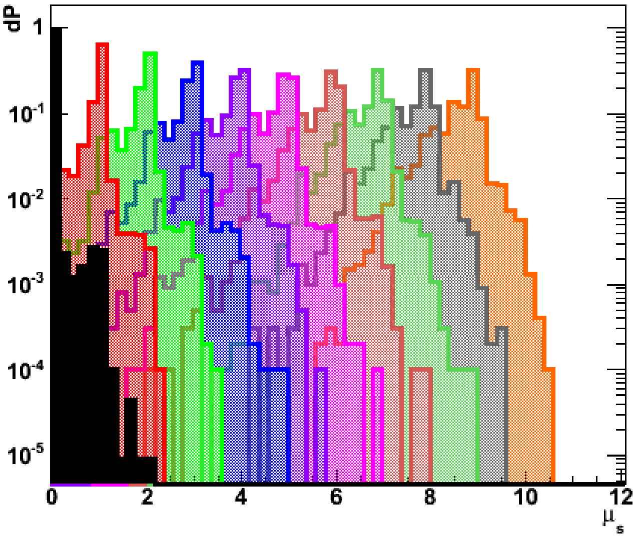


Figure 7.6: Distribution of the estimated number of signals μ_s for 0 (black), 1 (red), 2 (light green), 3 (blue), 4 (purple), 5 (pink), 6 (brown), 7 (green), 8 (grey) and 9 (orange) injected signal events for source C1 applying rlogL < 5.2.

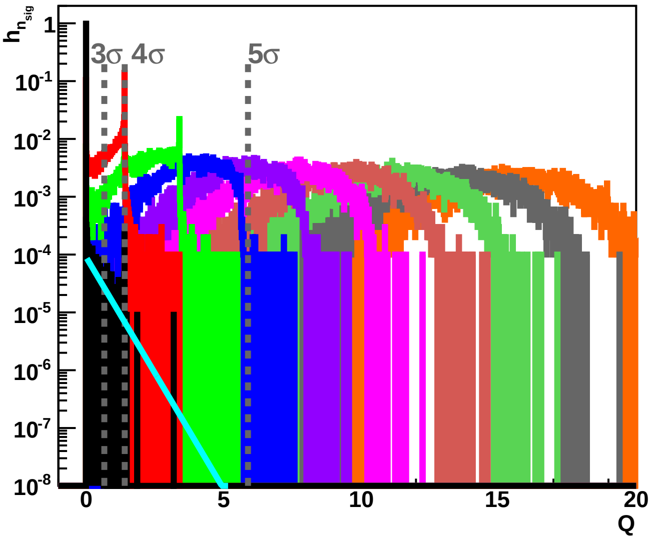


Figure 7.7: Distribution of the test statistic Q ($h_{n_{sig}}(Q)$) for 0 (black), 1 (red), 2 (light green), 3 (blue), 4 (purple), 5 (pink), 6 (brown), 7 (green), 8 (grey) and 9 (orange) injected signal events for C1 applying rlogL < 5.2. In turquoise the fit f(x) (see eqn. (7.18)) of the tail of $h_0(Q)$ is shown.

$n_{sig} = 0$ PEs with the highest Q values. The exponential function is defined as

$$f(x) = a \cdot \exp(-b \cdot x),\ x > x_1 \tag{7.18}$$

where $a > 0$ and $b > 0$. To determine the threshold $Q_{n\sigma}^{\text{threshold}} = x_2$ of Q for a defined value of $n\sigma$, one has to integrate over the distribution of Q. The integral is equal to $1 - p(n\sigma)$.

$$\int_0^{x_2} h_0(x)\mathrm{d}x = \int_0^{x_1} h_0(x)\mathrm{d}x + \int_{x_1}^{x_2} f(x)\mathrm{d}x = 0.97 + \int_{x_1}^{x_2} f(x)\mathrm{d}x = 1 - p(n\sigma) \tag{7.19}$$

The following equation has to be solved to get x_2

$$\int_{x_1}^{x_2} f(x)\mathrm{d}x = 0.03 - p(n\sigma) \overset{!}{=} I, \tag{7.20}$$

which is found to be

$$x_2 = \frac{-1}{b} \cdot \ln\left(\exp(-b \cdot x_1) - \frac{b}{a} \cdot I\right). \tag{7.21}$$

In eqn. (7.20) it is assumed that the total integral of $\int_{x_1}^{\infty} f(x)\mathrm{d}x = 0.03$. To ensure that, the integral $\int_{x_1}^{\infty} f(x)\mathrm{d}x = I'$ is computed and the factor to get the correct normalization of the function is calculated by $c = 0.03/I'$, so one has to substitute a with $a' = c \cdot a$ in eqn. (7.21). An example for the fitted function f(x) to the tail of the distribution of $h_0(Q)$ is shown in Fig. 7.7 as turquoise line.

The probability to measure Q for an estimated μ_s is given by

$$P(Q|\mu_s) = \sum_{n_s=0}^{\infty} \mathcal{P}(n_s|\mu_s) \cdot h_{n_{sig}}(Q), \tag{7.22}$$

where $\mathcal{P}(n_s|\mu_s)$ is the Poisson distribution. For each μ_s and Q value the probability $P(Q|\mu_s)$ to measure Q is calculated using eqn. (7.22) and is plotted in a two-dimensional histogram shown in Fig. 7.8 left as example. It is needed for the calculation of the MDP (model discovery potential) and MRF (model rejection factor).

$n\sigma$	fraction	p-value $p(n\sigma)$
1σ	0.682689492	$3.17 \cdot 10^{-1}$
2σ	0.954499736	$4.55 \cdot 10^{-2}$
3σ	0.997300204	$2.69 \cdot 10^{-3}$
4σ	0.99993666	$6.33 \cdot 10^{-5}$
5σ	0.999999426697	$5.74 \cdot 10^{-7}$

Table 7.3: Summary of the fraction of events included in the $n\sigma$ level and its corresponding p-values.

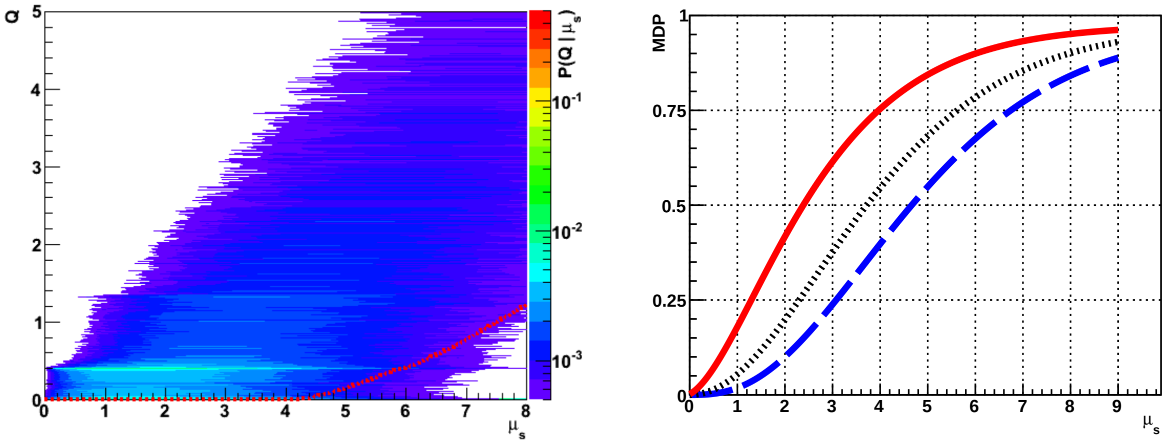


Figure 7.8: Left: two-dimensional distribution of $P(Q|\mu_s)$, the 10% curve (red) is calculated where 10% of PEs are below that line. Right: MDP depending on μ_s for 3σ (red), 4σ (black) and 5σ (blue) for source C1 with rlogL < 5.5.

7.4.2 MDP and MRF

Different cut parameters are established in chapter 5, where as main parameter, which is also the quality parameter, rlogL is used to distinguish between background and signal. In addition to this parameter, the rdf classification (rdf=1, for upgoing) is studied. Therefore several iterations of pseudo-experiments are done varying the rlogL cut from 5.0 to 5.5 (without rdf cut) and 5.0 to 5.6 with rdf=1 cut. This regions are defined as above an rlogL cut of 5.5 the fraction of atmospheric muons in the event sample rises extremely fast, so that these dominate the sample, which can be derived from the cumulative distribution of rlogL shown in Fig. 5.6 in section 5.2. If also the rdf=1 cut is applied, it is possible to loosen the cut, as the turning point is later in the rlogL curve. For example in Tab. 7.2 one can see that at an rlogL cut of 5.6 the total number of data events is a factor ~ 1.7 larger than the number of events with rdf cut. Below a cut of rlogL < 5.0 no atmospheric muons are left in the sample, in data the statistics for this cut is very low, only a few events, so that the background PDF can't be determined properly. Further, such cuts are too restrictive so that its unlikely to detect cosmic neutrinos.

To find the optimal cut values two different methods can be used: the model discovery potential (MDP) and the model rejection factor (MRF). These two concepts are explained in the following.

Model discovery potential MDP

The model discovery potential (MDP), which is the probability making a discovery assuming that the model (here neutrino flux assumption) is correct, can be calculated by the probability to measure a Q value greater then the determined $Q_p^{threshold}$ value for a measured μ_s: $P(Q \geq Q_p^{threshold}|\mu_s)$. The procedure to calculate the $MDP(\mu_s)$ is shown in eqn. (7.23) where an upper bound on the discovery potential is set, as the number of injected signal events is $0 \leq n_{sig} \leq 9$

in the pseudo-experiment generation.

$$\begin{aligned} MDP(\mu_s) = P(Q \geq Q_p^{threshold}|\mu_s) &= \int_{Q_p^{threshold}}^{\infty} P(Q|\mu_s)\mathrm{d}Q \\ &= \sum_{n_s=0}^{\infty} \mathcal{P}(n_s|\mu_s) \cdot \int_{Q_p^{threshold}}^{\infty} h_{n_s}(Q)\mathrm{d}Q \qquad (7.23) \\ \leq \sum_{n_s=0}^{9} \mathcal{P}(n_s|\mu_s) \cdot \int_{Q_p^{threshold}}^{\infty} h_{n_s}(Q)\mathrm{d}Q &+ \left(1 - \sum_{n_s=0}^{9} \mathcal{P}(n_s|\mu_s)\right) \cdot \int_{Q_p^{threshold}}^{\infty} h_9(Q)\mathrm{d}Q \end{aligned}$$

An example for the MDP calculation per μ_s can be found in Fig. 7.8 right, where the MDP is calculated for 3σ, 4σ and 5σ. The best cut combination is found at the cuts, with an maximal MDP or minimal $1/MDP$.

Model rejection factor MRF

Before looking into data an upper limit cannot be set, but an average upper limit on the number of observed events for 90% confidence level (C.L.) $\bar{\mu}_{90}(n_b)$ can be set by weighting all possible observed events by their Poisson probability to occur. Following the method described in [63], the model rejection factor (MRF) is defined as

$$MRF = \frac{\bar{\mu}_{90}}{n_{exp}}. \tag{7.24}$$

So the average upper flux limit gets $\bar{\Phi}(E)_{90} = \Phi(E) \cdot \frac{\bar{\mu}_{90}}{n_{exp}}$. The cut combination to receive the best limit when only background is observed, is found when the MRF is minimal.

In this analysis the MDP optimization for 3σ (and 5σ) is used. The MRF value is additional calculated to set an average upper limit on the expected neutrino flux (see 7.4.4).

7.4.3 Determining the sensitivity

To calculate the sensitivity, the number of expected events n_{exp} and the number of events necessary for a 50% chance to get a $n\sigma$ discovery $n_{n\sigma}$ have to be known. The number of expected events is determined from the β-distributions $N(\beta)$, derived from the point-like MC simulation per source and neutrino flux for each cut values separately. An example for this distribution is shown in Fig. 7.2 right. $n_{3\sigma}$ ($n_{5\sigma}$) is determined following the procedure of eqn. (7.23) by solving $MDP(n_{3\sigma}) = 0.5$ ($MDP(n_{5\sigma}) = 0.5$).

The sensitivity is then calculated by the fraction $n_{3\sigma,5\sigma}/n_{exp}$ times the expected neutrino flux for a 3σ (respectively 5σ) discovery as defined here

$$\text{sensitivity} = \frac{dN_\nu}{dE_\nu} \cdot \frac{n_{3\sigma,5\sigma}}{n_{exp}} = \Phi_{3\sigma,5\sigma} \cdot \left(\frac{E_\nu}{\text{TeV}}\right)^{-\alpha_\nu} \exp\left(-\sqrt{\frac{E_\nu}{E_{cut,\nu}}}\right), \tag{7.25}$$

with the parameters of the neutrino flux assumptions mentioned in section 6.2 and in the appendix A.1.

7.4.4 Calculating the 90% confidence level upper limit

To calculate the 90% C.L. (confidence level) the average number of background events $\bar{\mu}_{90}$ has to be known. It can be calculated from the two-dimensional distribution of $P(Q|\mu_s)$ shown in Fig. 7.8 as for each μ_s the Q value up to which 10% of the PEs are included can be calculated via

$$P(Q \leq Q_{meas}|\mu_s) = \int_0^{Q_{meas}} P(Q|\mu_s)\mathrm{d}Q = 0.1, \tag{7.26}$$

which is equivalent to measure for this μ_s in 90% of the PEs a Q value above Q_{meas}

$$P(Q \geq Q_{meas}|\mu_s) = \int_{Q_{meas}}^{\infty} P(Q|\mu_s)\mathrm{d}Q = 0.9. \tag{7.27}$$

The curve that represents this is called 10% curve and is shown as red curve in Fig. 7.8 left. To determine the average 90% C.L. upper limit on the number of events $\bar{\mu}_{90}$ the crossing point of the Q value for median background observation Q_0^{median}, which is the median of the distribution $h_0(Q)$, with the 10% line is determined.

So the average 90% confidence level upper limit on the neutrino flux is calculated by multiplying the neutrino flux assumption by the factor of median background observation divided by the number of expected events.

$$90\%C.L. = \frac{dN_\nu}{dE_\nu} \cdot \frac{\bar{\mu}_{90}}{n_{exp}} = \bar{\Phi}_{90} \cdot \left(\frac{E_\nu}{\mathrm{TeV}}\right)^{-\alpha_\nu} \exp\left(-\sqrt{\frac{E_\nu}{E_{cut,\nu}}}\right). \tag{7.28}$$

Chapter 8

Results and conclusions

In this chapter the results of the pseudo-experiments for each Milagro source candidate are discussed. The cut values are optimized to maximize the model discovery potential (MDP). The cut combination with the maximum MDP is taken for a 3σ (and a 5σ) discovery. As there are more then one flux assumption per source available for the sources MGRO J1908+06, MGRO J2019+37 and MGRO J2031+41, the plots of the final results are shown for the most probable result with the highest MDP of all flux assumptions per source. A systematic study of the effect of different detector conditions on the sensitivity is shown in section 8.2 for the kappes flux assumption as this is available for all five Milagro sources.

8.1 Sensitivity and average upper limits for the five Milagro sources

The outcome of the pseudo-experiments are statistically interpreted and the following variables are extracted and summarized for the cuts optimized for a 3σ discovery with the MDP method per source and neutrino flux assumption.

- N_{tot}: total number of events in the data sample,
- n_{exp}: number of expected events derived from the MC simulation per source weighted according to the neutrino flux assumption (see section 6.3),
- β_{med}: median of the angle β between source coordinates and reconstructed muons,
- $n_{3\sigma}$: mean number of events necessary for a 3σ discovery,
- MDP: model discovery potential calculated via eqn. (7.23) $MDP(n_{exp})$,
- $\Phi_{3\sigma}$: normalization of the 3σ sensitivity see eqn. (7.25),
- $\bar{\mu}_{90}$: 90% upper limit on μ_s for median background observation,
- MRF: model rejection factor calculated via eqn. (7.24),
- $\bar{\Phi}_{90}$: normalization of the average 90% C.L. upper limit see eqn. (7.28).

In this section the mentioned values correspond to a 3σ discovery, unless it is mentioned otherwise, as the number of expected events is very low - for the studied data sample with a live-time of 744.68 days (≈ 2.04 years) it is in the order of 10^{-3}. The optimization for a 5σ discovery is done for completeness and the respective tables can be found in the appendix C.1.

flux	rlogL	rdf	N_{tot}	n_{exp} $\cdot 10^{-3}$	β_{med} (deg)	$n_{3\sigma}$	MDP $\cdot 10^{-3}$	$\Phi_{3\sigma}$ $\cdot 10^{-10}$	$\bar{\mu}_{90}$	MRF $\cdot 10^{3}$	$\bar{\Phi}_{90}$ $\cdot 10^{-9}$
kappes	5.1	-	629	2.46	0.936	1.84	3.41	9.72	3.77	1.54	2.00
	5.1	1	606	2.40	0.941	1.86	3.38	10.1	3.82	1.59	2.07
halzen1	5.1	-	629	3.69	0.933	1.82	3.78	16.3	3.63	0.983	3.24
	5.1	1	606	3.59	0.936	1.87	3.71	17.2	3.73	1.04	3.43
halzen2	5.1	-	629	6.48	0.935	1.89	4.56	9.63	3.73	0.576	1.90
	5.1	1	606	6.33	0.939	1.86	4.61	9.70	3.79	0.599	1.98
halzen3	5.1	-	629	9.29	0.936	1.81	5.67	6.43	3.79	0.408	1.35
	5.1	1	606	9.09	0.941	1.80	5.61	6.53	3.82	0.420	1.39
halzen4	5.1	-	629	2.70	0.937	1.88	3.39	27.9	3.57	1.32	5.28
	5.1	1	606	2.62	0.939	1.88	3.41	28.7	3.65	1.40	5.60
halzen5	5.1	-	629	4.35	0.937	1.79	4.02	16.5	3.67	0.843	3.37
	5.1	1	606	4.23	0.940	1.81	4.03	17.1	3.75	0.886	3.54
halzen6	5.1	-	629	5.87	0.937	1.80	4.53	12.3	3.72	0.634	2.54
	5.1	1	606	5.72	0.941	1.85	4.41	12.9	3.75	0.655	2.62
halzen7	5.1	-	629	2.10	0.945	1.89	3.23	43.2	3.49	1.66	7.97
	5.1	1	606	2.02	0.946	1.91	3.21	45.4	3.57	1.76	8.45
halzen8	5.1	-	629	3.13	0.942	1.86	3.57	28.5	3.58	1.14	5.47
	5.1	1	606	3.03	0.944	1.85	3.58	29.3	3.63	1.20	5.76
halzen9	5.1	-	629	4.01	0.941	1.90	3.79	22.7	3.62	0.903	4.33
	5.1	1	606	3.89	0.943	1.89	3.76	23.3	3.71	0.954	4.58

Table 8.1: Summary of the optimized rlogL values (with and without rdf cut) for different neutrino flux assumption for MGRO J1908+06 for a 3σ discovery, $\Phi_{3\sigma}$ and $\bar{\Phi}_{90}$ are given in $\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ units.

8.1.1 MGRO J1908+06

The results for MGRO J1908+06 are summarized in Tab. 8.1. For each neutrino flux assumption the rlogL cut was optimized with and without rdf=1 cut. For MGRO J1908+06 the best rlogL cut value is 5.1 independently of the neutrino flux assumption and rdf cut. Adding the rdf cut additionally to the rlogL cut has no significant effect. There are only tiny differences in the number of expected events n_{exp}, $n_{3\sigma}$ and MDP, so its effect is negligible for this source candidate. The best MDP can be found for the flux halzen3 without rdf cut with an MDP of $5.67 \cdot 10^{-3}$ and $n_{3\sigma} = 1.81$ necessary events for a 3σ discovery.

With the flux parametrization of eqn. (6.2) discussed in section 6.2 with the relation of gamma ray flux parameters to neutrino flux parameters $k_\nu = (0.694 - 0.16\alpha_\gamma)k_\gamma$, $\alpha_\nu = \alpha_\gamma$, $E_{cut,\nu} = 0.59 E_{cut,\gamma}$ and $k_\gamma = \frac{K}{E_{norm}^{-\alpha_\gamma}}$ (for the halzen flux), the neutrino flux halzen3 with $K =$

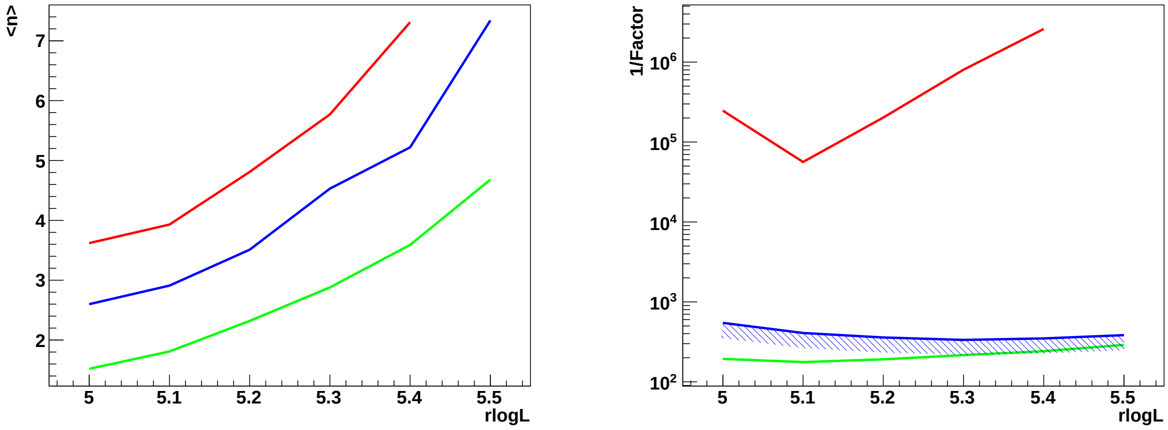


Figure 8.1: Left: Mean number of events $\langle n\rangle$ needed for a 3σ (green), 4σ(blue) and 5σ(red) discovery; Right: 1/MDP for a 3σ (green) and 5σ (red) discovery and MRF (blue); depending on the rlogL cut value for MGRO J1908+06 for neutrino flux assumption halzen3.

$6.1\cdot 10^{-13}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}$, $E_{\mathrm{norm}} = 4\,\mathrm{TeV}$, $\alpha_\gamma = 1.9$, $E_{\mathrm{cut},\gamma} = 300\,\mathrm{TeV}$ gets

$$\frac{dN}{dE_\nu} = 3.3\cdot 10^{-12}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}\cdot (E_\nu/\mathrm{TeV})^{-1.9}\exp(-\sqrt{E_\nu/177\,\mathrm{TeV}}),$$

which corresponds to the highest cut-off energy and lowest spectral index of the halzen neutrino flux parametrizations for this source (see Tab. A.1 in appendix A.1), which gives the highest neutrino flux and so the highest number of expected events. For this flux the trend of the necessary mean number of events $\langle n\rangle$ with rlogL value for 3σ (green), 4σ (blue) and 5σ (red) is shown in Fig. 8.1 on the left. $\langle n\rangle$ rises for all $n\sigma$. If some values of rlogL have no visible point for 5σ optimization, then that value could not be determined for this significance. On the right plot of Fig. 8.1 the model discovery potential is shown as 1/MDP for 3σ (green) and 5σ (red), the MRF is also shown in blue. From this curve one can derive that the best rlogL value is that value where 1/MDP is minimal. For 3σ optimization the best rlogL cut is 5.1 as already mentioned above. For this cut and the halzen3 flux the discovery potential is shown in Fig. 8.2 left. $n_{3\sigma} = 1.81$ is determined from the red curve, where a 50% chance for discovery is given. For the same cut the number of necessary events for a 5σ discovery is 3.93. On the right plot of Fig. 8.2 the sensitivity is shown for the halzen3 flux for the mentioned cut, which is a factor $1.9\cdot 10^2$ higher then the assumed flux normalization. The average 90% C.L. upper limit is for this case $\bar{\Phi}_{90} = 1.35\cdot 10^{-9}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is shown as blue curve in Fig. 8.2. This leads to the normalization of the sensitivity of $\Phi_{3\sigma} = 6.3\cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} = 6.3\cdot 10^{-7}\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

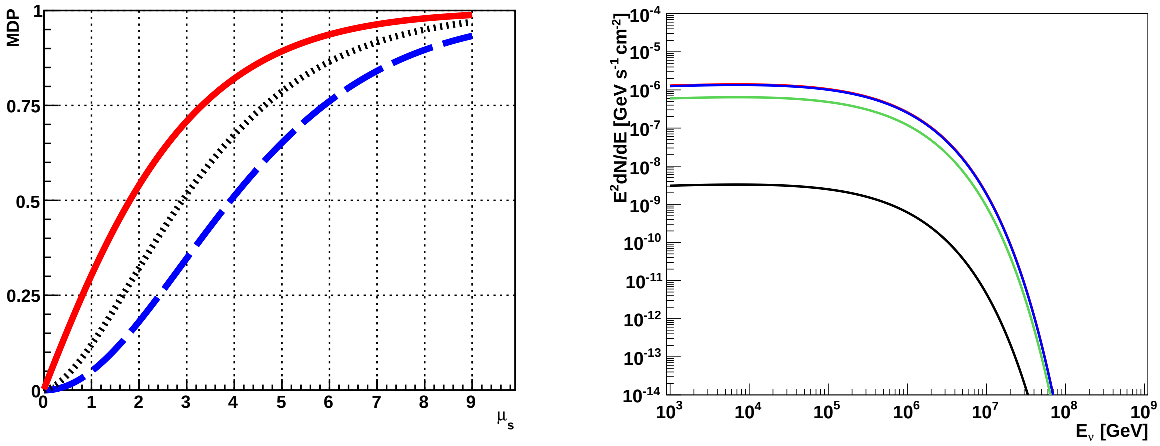


Figure 8.2: Left: MDP for a 3σ (green), 4σ (black), 5σ (blue) discovery; Right: sensitivity for a 3σ (green) and 5σ (red) discovery, limit (blue) and halzen3 neutrino flux assumption (black); for MGRO J1908+06 with rlogL cut 5.1 for neutrino flux assumption halzen3.

8.1.2 MGRO J2019+37

For MGRO J2019+37 the summary of the different values determined for 3σ optimization is shown in Tab. 8.2. Here the maximal MDP is found for the neutrino flux assumption halzen3 with rdf=1 cut for rlogL value 5.2. The tendency of cutting additionally on rdf=1 leads here mostly to a looser rlogL cut with a better MDP. The neutrino flux halzen3 with the values extracted from Tab. A.2 in appendix A.1 gets

$$\frac{dN}{dE_\nu} = 1.8 \cdot 10^{-12}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} \cdot (E_\nu/\mathrm{TeV})^{-1.8} \exp(-\sqrt{E_\nu/177\,\mathrm{TeV}}),$$

which corresponds to the highest cut-off energy and lowest spectral index of the halzen neutrino flux parametrizations for this source.

In Fig. 8.3 left the trend of the necessary number of events $\langle n \rangle$ for 3σ (green), 4σ (blue) and 5σ (red) is shown in dependence of the rlogL cut applying the additional rdf=1 cut. $\langle n \rangle$ rises for 3σ and 4σ, but is more or less constant for 5σ. On the right plot of Fig. 8.3 the model discovery potential is shown as 1/MDP for 3σ (green) and 5σ (red) and the MRF is shown in blue. For 3σ optimization the best rlogL cut is 5.2 with rdf=1 cut as already mentioned above, as the 1/MDP curve is minimal at this point. For this cut and the halzen3 flux the discovery potential is shown in Fig. 8.4 left. $n_{3\sigma} = 0.93$ are necessary for a 50% chance of a 3σ discovery. On the right plot in Fig. 8.3, the sensitivity is shown for the halzen3 flux for the mentioned cuts. which is a factor $2.0 \cdot 10^2$ higher then the assumed flux normalization. The normalization of the sensitivity is $\Phi_{3\sigma} = 3.5 \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} = 3.5 \cdot 10^{-7}\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. An average upper limit could be set to $\bar{\Phi}_{90} = 1.32 \cdot 10^{-9}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

For the optimization for a 5σ discovery, different rlogL cuts are determined. For the halzen3 flux the best cut values are rlogL value of 5.6 with rdf=1 (see Tab. C.2 in appendix C.1), which is the top border of the optimization cuts. For this $n_{5\sigma} = 4.73$, with a sensitivity of $\Phi_{5\sigma} = 8.3 \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} = 8.3 \cdot 10^{-7}\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ are determined.

flux	rlogL	rdf	N_{tot}	n_{exp} $\cdot 10^{-3}$	β_{med} (deg)	$n_{3\sigma}$	MDP $\cdot 10^{-3}$	$\Phi_{3\sigma}$ $\cdot 10^{-10}$	$\bar{\mu}_{90}$	MRF $\cdot 10^{3}$	$\bar{\Phi}_{90}$ $\cdot 10^{-9}$
kappes	5.1	-	629	1.07	0.358	0.790	3.62	9.60	2.43	2.27	2.95
	5.2	1	850	1.55	0.378	0.970	3.76	8.14	2.58	1.67	2.17
halzen1	5.1	-	629	0.933	0.383	0.830	3.45	16.0	2.46	2.64	4.75
	5.3	1	1207	1.87	0.447	1.31	3.52	12.6	2.86	1.53	2.75
halzen2	5.3	-	1282	4.08	0.422	1.33	4.51	5.87	2.95	0.723	1.30
	5.3	1	1207	3.81	0.417	1.19	4.70	5.62	2.85	0.747	1.34
halzen3	5.1	-	629	3.29	0.338	0.760	5.68	4.16	2.41	0.733	1.32
	5.2	1	850	4.72	0.362	0.930	6.13	3.55	2.54	0.539	0.97
halzen4	5.1	-	629	0.737	0.404	0.850	3.29	30.0	2.46	3.34	8.68
	5.3	1	1207	1.51	0.473	1.38	3.29	23.8	2.92	1.93	5.02
halzen5	5.1	-	629	1.40	0.376	0.810	3.84	15.0	2.44	1.74	4.52
	5.3	1	1207	2.77	0.438	1.29	4.01	12.1	2.86	1.03	2.68
halzen6	5.1	-	629	2.14	0.358	0.790	4.53	9.60	2.43	1.13	2.94
	5.2	1	850	3.09	0.378	0.980	4.80	8.25	2.59	0.837	2.18
halzen7	5.1	-	629	0.618	0.430	0.880	3.16	54.1	2.50	4.04	15.4
	5.3	1	1207	1.30	0.507	1.43	3.20	41.8	2.98	2.30	8.74
halzen8	5.1	-	629	1.06	0.400	0.830	3.57	29.8	2.48	2.33	8.85
	5.3	1	1207	2.16	0.469	1.32	3.66	23.2	2.90	1.34	5.09
halzen9	5.3	-	1282	3.23	0.450	1.37	4.04	16.1	3.00	0.929	3.53
	5.3	1	1207	3.01	0.446	1.26	4.16	15.9	2.89	0.961	3.65

Table 8.2: Summary of the optimized rlogL values (with and without rdf cut) for different neutrino flux assumption for MGRO J2019+37 for a 3σ discovery, $\Phi_{3\sigma}$ and $\bar{\Phi}_{90}$ are given in $\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ units.

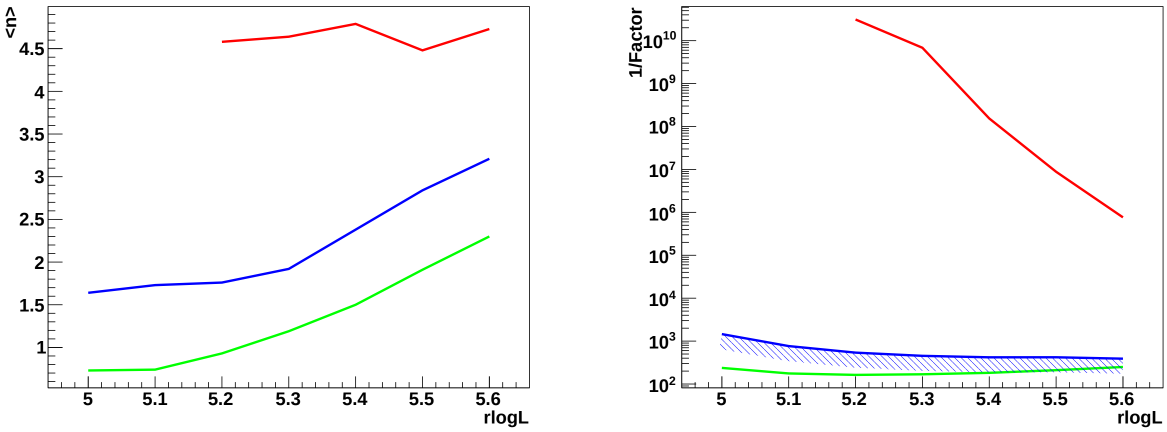


Figure 8.3: Left: Mean number of events $\langle n \rangle$ needed for a 3σ (green), 4σ (blue), 5σ (red) discovery; Right: 1/MDP for a 3σ (green) and 5σ (red) discovery and MRF (blue); depending on the rlogL cut value for MGRO J2019+37 for neutrino flux assumption halzen3 with rdf=1 cut.

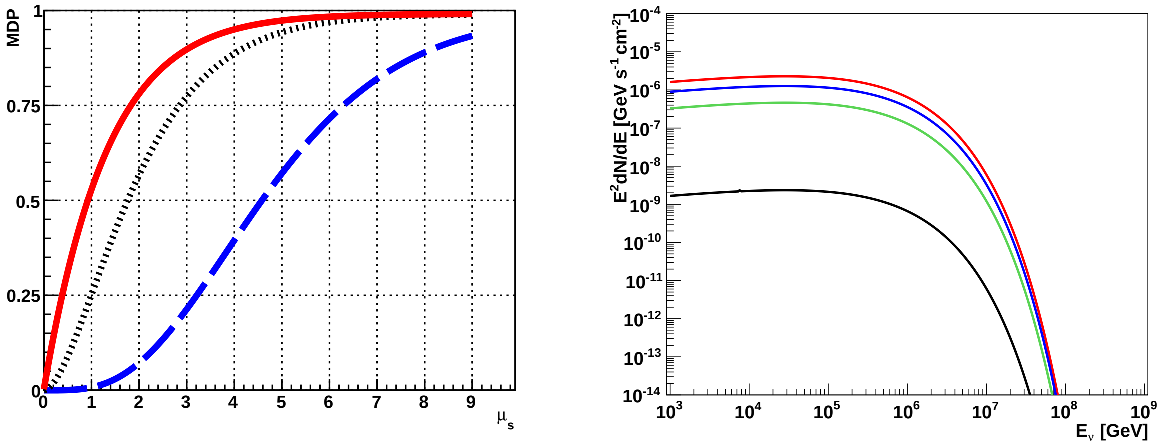


Figure 8.4: Left: MDP for a 3σ (green), 4σ (black), 5σ (blue) discovery; Right: sensitivity for a 3σ (green) and 5σ (red) discovery, limit (blue) and halzen3 neutrino flux assumption (black); for MGRO J2019+37 with rlogL cut 5.2 for neutrino flux assumption halzen3 with rdf=1 cut.

8.1.3 MGRO J2031+41

For MGRO J2031+41 the summary of the different values for 3σ optimization of the cut parameters is shown in Tab. 8.3. The best MDP is determined for the kappes neutrino flux parametrization without rdf cut for rlogL value 5.3. The kappes neutrino flux prediction is

$$\frac{dN}{dE_\nu} = 9.0 \cdot 10^{-13}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} \cdot (E_\nu/\mathrm{TeV})^{-2.0} \exp(-\sqrt{E_\nu/177\,\mathrm{TeV}}),$$

with a spectral index of 2.0, which is lower than the spectral index of 2.6 from the halzen parametrization with the same cut-off energy of $E_{cut,\gamma} = 300\,\mathrm{TeV}$.

For this flux the trend of the necessary number of events $\langle n \rangle$ with rlogL value for a 3σ (green), a 4σ (blue) and a 5σ (red) discovery is shown in Fig. 8.5 on the left. $\langle n \rangle$ rises for 3σ and 4σ, for 5σ a slight decrease at rlogL values 5.3 and 5.4 is visible. On the right plot of Fig. 8.5 the model discovery potential is shown as 1/MDP for 3σ (green) and 5σ (red) and the MRF is shown in blue. The minimal point of the green curve is at an rlogL value of 5.3. For this cut and the kappes flux the discovery potential is shown in Fig. 8.6 left. $n_{3\sigma} = 1.07$ is determined from the red curve, where a 50% chance for a discovery is given. On the right the sensitivity is shown for the kappes flux for the mentioned cut, which is a factor $8.9 \cdot 10^2$ higher then the assumed neutrino flux normalization. The normalization of the sensitivity is $\Phi_{3\sigma} = 8.03 \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} = 8.03 \cdot 10^{-7}\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is shown in green in Fig. 8.6 right.

For the kappes flux with an optimization for a 5σ discovery gives $n_{5\sigma} = 4.88$ for an rlogL cut of 5.5. The sensitivity gets then $2.2 \cdot 10^{-9}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} = 2.2 \cdot 10^{-6}\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

flux	rlogL	rdf	N_{tot}	n_{exp} $\cdot 10^{-4}$	β_{med} (deg)	$n_{3\sigma}$	MDP $\cdot 10^{-3}$	$\Phi_{3\sigma}$ $\cdot 10^{-9}$	$\bar{\mu}_{90}$	MRF $\cdot 10^{3}$	$\bar{\Phi}_{90}$ $\cdot 10^{-8}$
kappes	5.3	-	1282	12.0	0.348	1.07	3.43	0.803	2.92	2.44	0.22
	5.3	1	1207	10.6	0.342	1.00	3.36	0.849	2.83	2.67	0.24
halzen1	5.2	-	889	1.67	0.465	1.16	2.78	16.0	2.87	17.2	3.96
	5.3	1	1207	2.20	0.513	1.27	2.80	13.3	3.04	13.8	3.17
halzen2	5.3	-	1282	3.78	0.500	1.36	2.84	8.28	3.12	8.24	1.90
	5.3	1	1207	3.23	0.476	1.21	2.84	8.62	2.99	9.26	2.13
halzen3	5.3	-	1282	4.76	0.472	1.34	2.91	6.47	3.10	6.51	1.50
	5.3	1	1207	4.08	0.450	1.16	2.92	6.54	2.99	7.32	1.68
halzen4	5.3	-	1282	2.60	0.591	1.60	2.78	20.3	3.16	12.1	3.99
	5.2	1	850	1.36	0.4.67	1.06	2.77	25.7	2.74	20.2	6.67
halzen5	5.3	-	1282	3.55	0.551	1.47	2.83	13.7	3.17	8.92	2.94
	5.2	1	850	1.92	0.438	0.990	2.81	17.0	2.74	14.3	4.72
halzen6	5.3	-	1282	4.28	0.529	1.40	2.87	10.8	3.15	7.36	2.43
	5.3	1	1207	3.62	0.501	1.28	2.86	11.7	3.05	8.44	2.79
halzen7	5.1	-	629	0.873	0.489	0.940	2.75	48.5	2.58	29.5	13.3
	5.3	1	1207	2.24	0.606	1.53	2.75	30.7	3.13	14.0	6.30
halzen8	5.3	-	1282	3.56	0.618	1.57	2.79	19.8	3.23	9.08	4.09
	5.3	1	1207	2.93	0.569	1.44	2.80	22.1	3.15	10.8	4.86
halzen9	5.3	-	1282	4.13	0.590	1.57	2.82	17.1	3.19	7.72	3.47
	5.3	1	1207	3.42	0.551	1.39	2.84	18.3	3.11	9.09	4.09

Table 8.3: Summary of the optimized rlogL values (with and without rdf cut) for different neutrino flux assumption for MGRO J2031+41 for a 3σ discovery, $\Phi_{3\sigma}$ and $\bar{\Phi}_{90}$ are given in $\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ units.

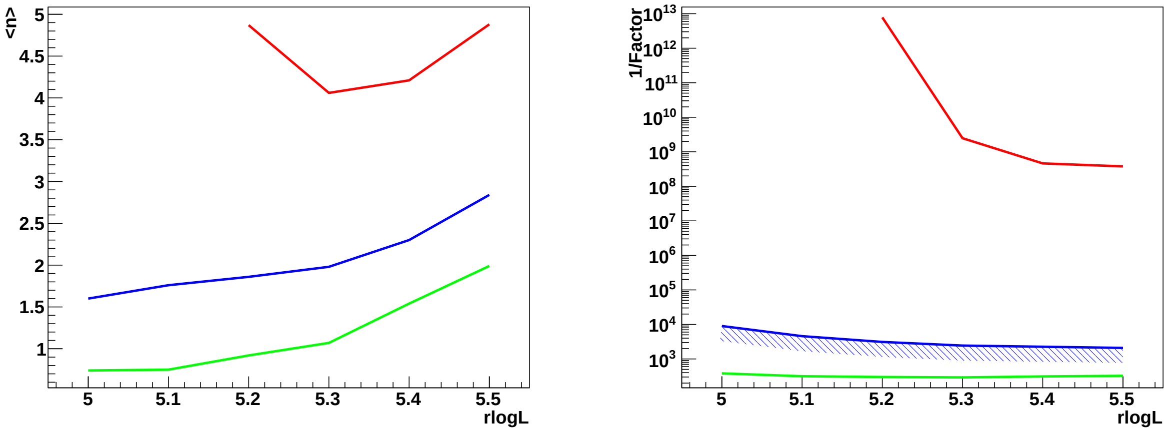


Figure 8.5: Left: Mean number of events $\langle n \rangle$ needed for a 3σ (green), 4σ (blue), 5σ (red) discovery; Right: 1/MDP for a 3σ (green) and 5σ (red) discovery and MRF (blue); depending on the rlogL cut value for MGRO J2031+41 for neutrino flux assumption kappes.

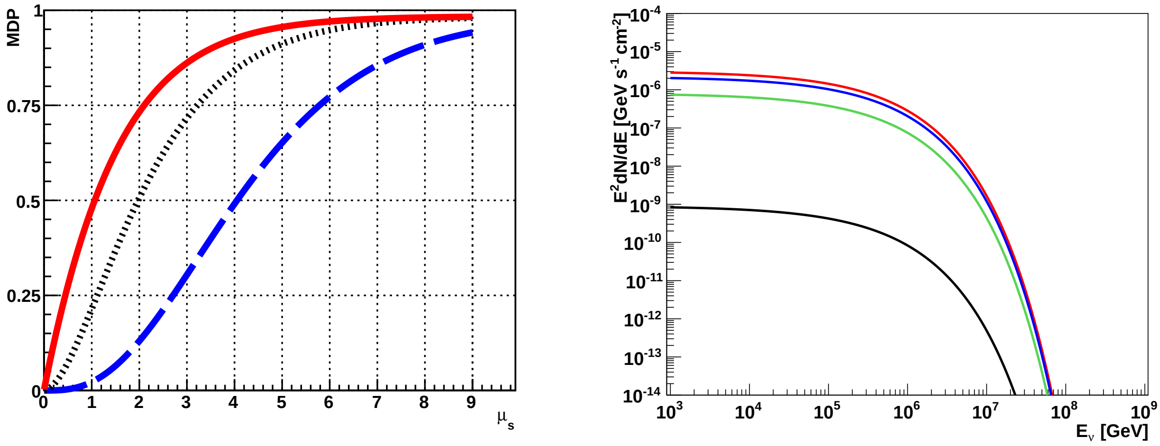


Figure 8.6: Left: MDP for a 3σ (green), 4σ (black), 5σ (blue) discovery; Right: sensitivity for a 3σ (green) and 5σ (red) discovery, limit (blue) and kappes neutrino flux assumption (black); for MGRO J2031+41 with rlogL cut 5.3 for neutrino flux assumption kappes.

8.1.4 C1

For C1 only one flux parametrization is available, the kappes neutrino flux parametrization, which is

$$\frac{dN}{dE_\nu} = 6.4 \cdot 10^{-13}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} \cdot (E_\nu/\mathrm{TeV})^{-2.0} \exp(-\sqrt{E_\nu/177\,\mathrm{TeV}}).$$

The optimal cut value for a 3σ discovery is rlogL value 5.2, where the MDP is a little bit better with additional rdf=1 cut then without, which can be seen in Tab. 8.4.

For this flux with an additional rdf=1 cut, the trend of the necessary number of events $\langle n \rangle$ depending on the rlogL value for 3σ (green), 4σ (blue) and 5σ (red) is shown in Fig. 8.7 left. $\langle n \rangle$ shows a rising tendency for 3σ and 4σ, whereas the 5σ curve (red) decreases towards an rlogL value of 5.3 and rises slowly for higher rlogL values. On the right plot of Fig. 8.7 the model discovery potential is shown as 1/MDP for 3σ(green) and 5σ(red), the MRF is shown in blue. From this the best rlogL cut of 5.2 is determined using the minimum of the 3σ 1/MDP curve. For this cut the discovery potential is shown in Fig. 8.8 left. From

source	rlogL	rdf	N_{tot}	n_{exp} $\cdot 10^{-4}$	β_{med} (deg)	$n_{3\sigma}$	MDP $\cdot 10^{-3}$	$\Phi_{3\sigma}$ $\cdot 10^{-10}$	$\bar{\mu}_{90}$	MRF $\cdot 10^{3}$	$\bar{\Phi}_{90}$ $\cdot 10^{-9}$
C1	5.2	-	889	7.83	0.383	1.12	3.13	9.15	2.76	3.52	2.25
	5.2	1	850	7.45	0.380	1.01	3.16	8.68	2.65	3.56	2.28
C2	5.2	-	889	7.70	0.393	1.05	3.18	6.68	2.72	3.53	1.73
	5.2	1	850	7.28	0.387	0.940	3.21	63.3	2.62	3.60	1.76

Table 8.4: Summary of the optimized rlogL values (with and without rdf cut) for different neutrino flux assumption for C1 and C2 for a 3σ discovery for the kappes neutrino flux assumption, $\Phi_{3\sigma}$ and $\bar{\Phi}_{90}$ are given in $\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ units.

the red curve, where a 50% chance for discovery is given, $n_{3\sigma} = 1.01$ is determined. On the right the sensitivity is shown for the mentioned cuts. The sensitivity for a 3σ discovery is $\Phi_{3\sigma} = 8.7 \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} = 8.7 \cdot 10^{-7}\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for the kappes flux, which is a factor $1.36 \cdot 10^3$ higher then the assumed neutrino flux normalization, and is shown in green in Fig. 8.8 right. An average upper limit can be set to $\bar{\Phi}_{90} = 2.28 \cdot 10^{-9}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

For the optimization for a 5σ discovery a rlogL cut of 5.6 with rdf=1 is determined (see appendix C.1 Tab. C.4). The normalization of the sensitivity is here $\Phi_{5\sigma} = 1.8{\cdot}10^{-9}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ with $n_{5\sigma} = 4.91$.

8.1.5 C2

For C2 the kappes neutrino flux parametrization is

$$\frac{dN}{dE_\nu} = 4.8 \cdot 10^{-13}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} \cdot (E_\nu/\mathrm{TeV})^{-2.0} \exp(-\sqrt{E_\nu/177\,\mathrm{TeV}}).$$

The optimal cut values for a 3σ discovery are a rlogL of 5.2 with additional rdf=1 cut (see Tab. 8.4). The trend of the necessary number of events $\langle n \rangle$ with rlogL value for 3σ (green), 4σ (blue) and 5σ (red) is shown in Fig. 8.9 on the left. On the right plot of Fig. 8.9 the model discovery potential is shown as 1/MDP for 3σ (green) and 5σ (red) and the MRF is shown in blue. From this curve one can derive that the best rlogL value is 5.2. For this cut the discovery potential is shown in Fig. 8.10 left. $n_{3\sigma} = 0.94$ is determined from the red curve, where a 50% chance for discovery is given. On the right the sensitivity is shown for the mentioned cuts, which is a factor $1.29 \cdot 10^3$ higher then the assumed neutrino flux normalization , which is $\Phi_{3\sigma} = 6.3 \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1} = 6.3 \cdot 10^{-7}\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and is shown in green in Fig. 8.10 right.

For the optimization for a 5σ discovery a rlogL cut of 5.6 with rdf=1 is determined (see appendix C.1 Tab. C.4). The normalization of the sensitivity is here $\Phi_{5\sigma} = 1.3{\cdot}10^{-9}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

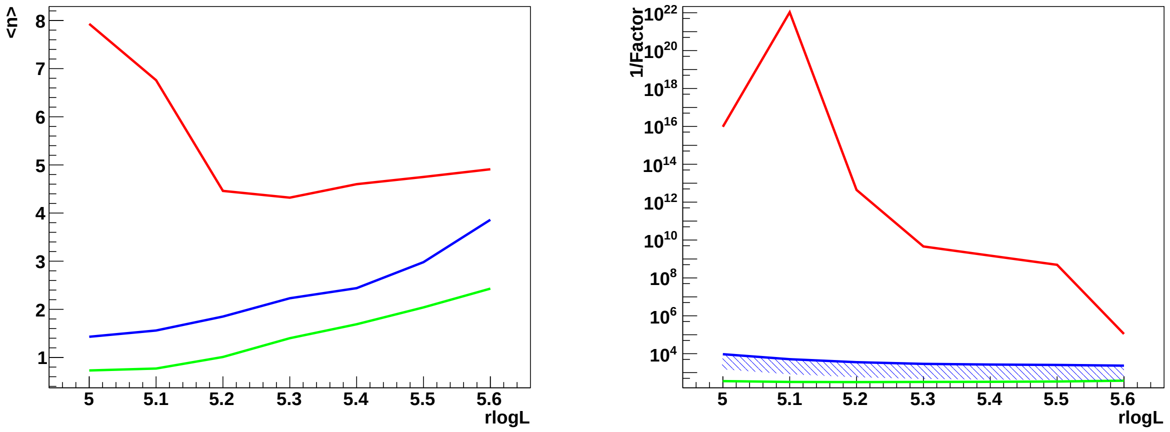


Figure 8.7: Left: Mean number of events $\langle n \rangle$ needed for a 3σ (green), 4σ (blue), 5σ (red) discovery; Right: 1/MDP for a 3σ (green) and 5σ (red) discovery and MRF (blue); depending on the rlogL cut value for C1 for neutrino flux assumption kappes with rdf=1 cut.

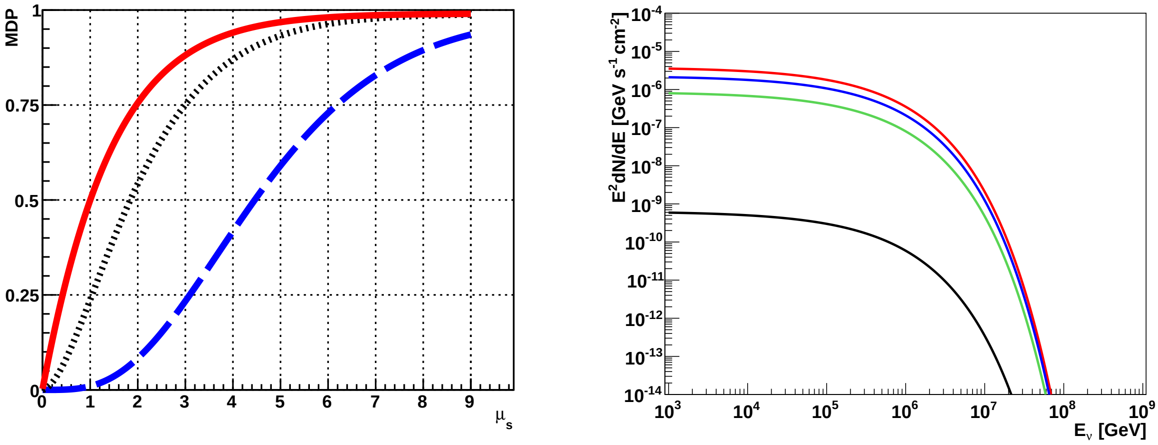


Figure 8.8: Left: MDP for a 3σ (green), 4σ (black), 5σ (blue) discovery; Right: sensitivity for a 3σ (green) and 5σ (red) discovery, limit (blue) and kappes neutrino flux assumption (black); for C1 with rlogL cut 5.2 for neutrino flux assumption kappes with rdf=1 cut.

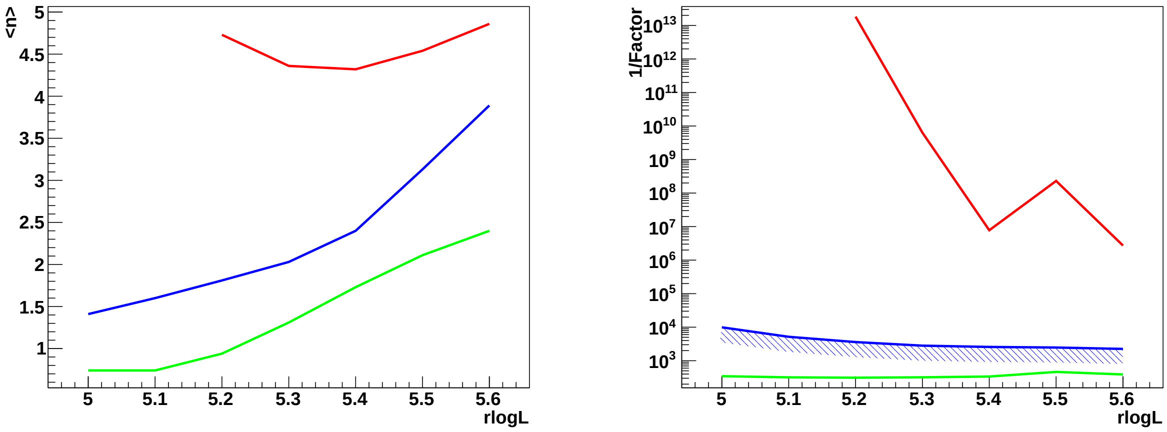


Figure 8.9: Left: Mean number of events $\langle n \rangle$ needed for a 3σ (green), 4σ (blue), 5σ (red) discovery; Right: 1/MDP for a 3σ (green) and 5σ (red) discovery and MRF (blue); depending on the rlogL cut value for C2 for neutrino flux assumption kappes with rdf=1 cut.

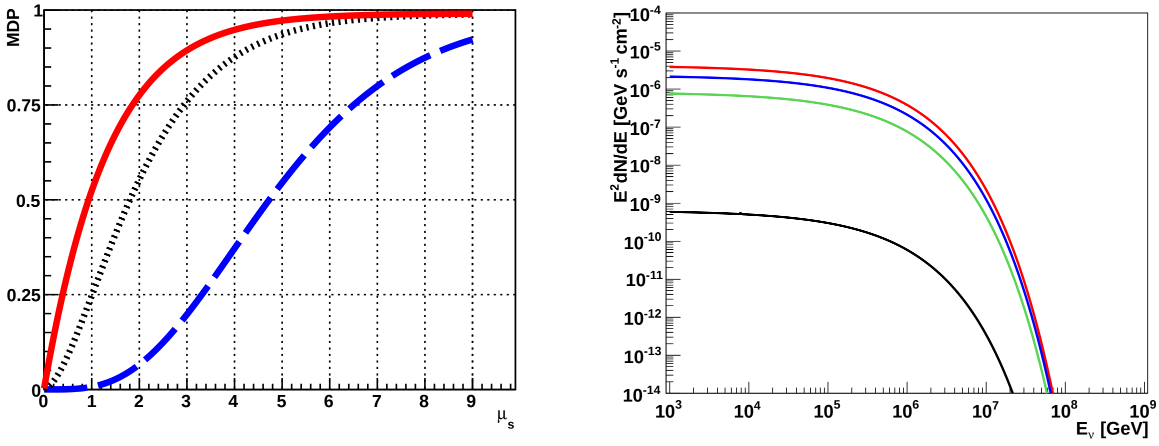


Figure 8.10: Left: MDP for a 3σ (green), 4σ (black), 5σ (blue) discovery; Right: sensitivity for a 3σ (green) and 5σ (red) discovery, limit (blue) and kappes neutrino flux assumption (black); for C2 with rlogL cut 5.2 for neutrino flux assumption kappes with rdf=1 cut.

8.2 Systematics of detector conditions

As already mentioned in section 6.3 the neutrino events are simulated for each Milagro source separately assuming a standardised, representative working condition of the ANTARES detector, which means primarily optical background rates and active working OMs. Different detector conditions are studied by systematically changing one of these parameters named cond1 to cond3 (changing the baseline rate to 100 kHz, 150 kHz and 200 kHz) and cond4 and cond5 (changing the number of active OMs to 600 and 500). The definition of the various conditions can be found in section 6.3. To systematically study the effect of changing the detector conditions on the sensitivity only the kappes neutrino flux expectation is considered as this is available for each of the five Milagro sources.

Pseudo-experiments are performed using the signal PDF determined from the simulation of the point-like sources for the specified detector conditions. The results of the evaluation of the pseudo-experiments for a 3σ optimization are shown in Tab. 8.6 to Tab. 8.10. For getting an estimate of the error on the sensitivity due to different detector conditions an approximation of the trend of the sensitivity is made for each parameter separately, see section 8.2.2, and an interval, which consists 68% of runs around the median of the distributions, is taken as basis for the error calculation.

8.2.1 Differences in signal PDF

The only part which is influenced by different detector conditions in the process of generating the pseudo-experiments is the signal PDF and therefore also the signal generation as the overall detector conditions are already included in the background PDF (as this is determined from the data events of all considered runs). As different detector conditions affect the performance of the track reconstruction and therefore the distribution of the space angle between source and reconstructed event, this is studied here separately. To compare the effect on the shape of the PSF the signal PDF is plotted for different detector conditions applying the same quality cut, see Fig. 8.11, where the signal PDF is shown for all five Milagro sources for events reconstructed with an quality cut rlogL < 5.5.

From the curves for different numbers of active OMs (cond4 and cond5) one can derive that the number of expected events decreases with decreasing numbers of active OMs, as the light blue and pink curves lie clearly below the black one of the standard condition. This can also be seen in the tables Tab. 8.6 to Tab. 8.10, where the actual determined values are given for optimized cut parameter per detector condition.

The situation for rising baseline rates is a bit different. Changing the baseline rate from 77.6 kHz (standard) to 100 kHz (cond1) the curves are comparable in the shape as well as in the number of expected events. So it won't have a huge effect changing the baseline rate to this value. The majority of runs selected for this study (85%) have baseline rates up to 100 kHz. So the error on the sensitivity will be low regarding different baseline rates. For higher baseline rates, as for example 200 kHz, the number of expected events is significantly decreased and for the case of MGRO J1908+06 the point of turnover from a constant to a shaped behaviour is higher (see Fig. 8.11 top left).

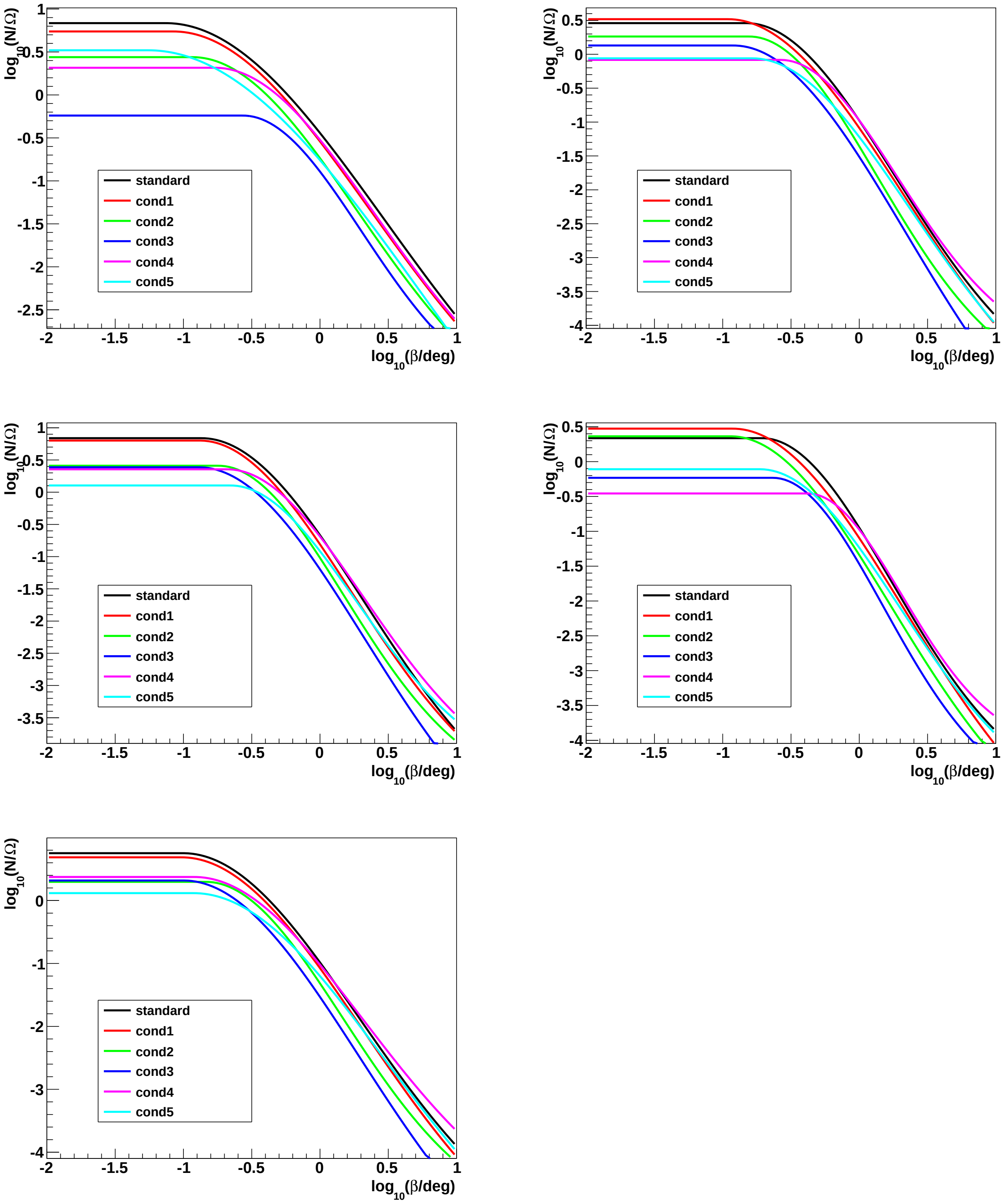


Figure 8.11: Signal PDFs determined for the different detector conditions for the Milagro sources MGRO J1908+06, MGRO J2019+37 and MGRO J2031+41 (left: top to bottom) and C1 and C2 (right: top to bottom) for rlogL < 5.5 and their kappes neutrino flux assumption.

8.2.2 Calculation of error intervals

For the selection of runs for this systematic study two parameters are used, this is the baseline rate and the number of active OMs. While one of these parameters is constant the other parameter is changed. Assuming that these parameters are linear independent the error corresponding to one of these parameters can be calculated separately and they can be added using a quadratic error addition.

For the run selection described in section 5.1 a two-dimensional histogram is created, where the number of active OMs dependent on the baseline rate is shown (see Fig. 8.12). No correlation between these two variables is assumed and found, the correlation factor is $-8.3 \cdot 10^{-2}$. So they can be treated as independent parameters in the error calculation.

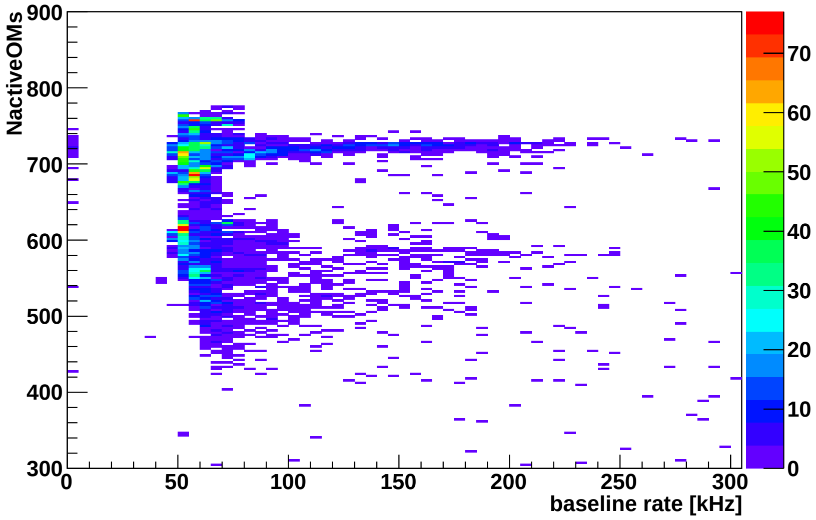


Figure 8.12: Distibution of the number of active OMs dependent on the baseline rate for the run selection used in this analysis (see section 5.1).

For each variable the one-dimensional distribution shown in section 6.3 (for the baseline rate see Fig. 6.4 and for the number of active OMs see Fig. 6.5) is used to determine the region around the median value of the distribution where 68% of runs are included. It is defined symmetrically, so that 34% of runs are below the median value and 34% are above the median value. The median value and the boundaries a_{min} and a_{max} for the active OMs and b_{min} and b_{max} for the baseline rate are given in Tab. 8.5.

	baseline rate (kHz)	nactiveOMs
median	62.5	685.5
lower boundary	52.5	559.5
upper boundary	82.5	727.5

Table 8.5: Median, lower and higher boundaries for the distributions of the baseline rate and the number of active OMs, where 68% of the runs are included.

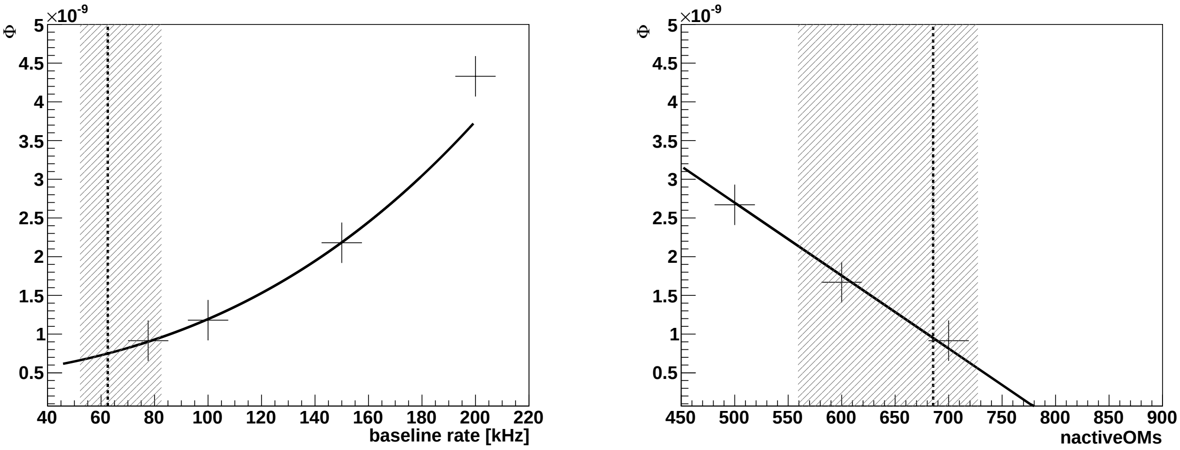


Figure 8.13: Determined sensitivity in ($\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$) for the different detector conditions varying the baseline rate (left) and the number of active OMs (right) with fitted functions f(x), respectively g(x), for source C1. The dotted line represents the median of the baseline (number of active OMs) distribution and the grey part defines the interval around the median that contains 68% of the runs.

To get an estimate of the sensitivity that could be determined for these boundaries the sensitivity depending on the baseline rates is plotted for the sensitivities determined by standard and cond1 to cond3 detector condition. The function

$$f(x) = a + b \cdot (x - c)^3 \tag{8.1}$$

is fitted to the points (see Fig. 8.13 left for example). The lower and upper sensitivities are determined by this function $\Phi_{min} = f(b_{min})$ and $\Phi_{max} = f(b_{max})$ and the lower and upper error is calculated to the measured sensitivity for standard detector condition. The same procedure is done for the number of active OMs, where a linear dependency of the sensitivity is observed, so a linear function $g(x) = -a \cdot x + b$ is fitted (see Fig. 8.13 right for example). A quadratic error calculation is performed to get the total upper and lower error on the sensitivity.

For each of the sources the results shown in Tab. 8.6 to Tab. 8.10 are taken and the errors are calculated with the method mentioned above. The result and the uncertainties are given here:

- MGRO J1908+06: $\Phi_{3\sigma} = 9.72^{+10.75}_{-2.41} \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
- MGRO J2019+37: $\Phi_{3\sigma} = 9.60^{+13.06}_{-3.79} \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
- MGRO J2031+41: $\Phi_{3\sigma} = 8.03^{+12.01}_{-4.85} \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
- C1: $\Phi_{3\sigma} = 9.15^{+12.22}_{-4.38} \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$
- C2: $\Phi_{3\sigma} = 6.68^{+8.68}_{-2.39} \cdot 10^{-10}\,\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

The relative upper errors range from 110% to 150%, the relative lower errors range from 25% to 60%. The source with the lowest relative errors is MGRO J1908+06, that is the source of the five that is best visible with ANTARES and MGRO J2013+41 has the largest relative errors, which is the source that has the lowest visibility within ANTARES. To get a rough estimate of the uncertainties on the sensitivity for the halzen neutrino flux assumption the mean relative errors are calculated. The mean relative upper uncertainty gets 131% and the mean relative lower uncertainty is 41%. These are rather large uncertainties so a full rbr simulation or a more dedicated run selection would be needed before unblinding the data.

condition	rlogL	N_{tot}	n_{exp} $\cdot 10^{-3}$	β_{med} (deg)	$n_{3\sigma}$	MDP $\cdot 10^{-3}$	$\Phi_{3\sigma}$ (TeV^{-1}cm^{-2}s^{-1})
standard	5.1	629	2.46	0.936	1.84	3.41	$9.72 \cdot 10^{-10}$
cond1	5.2	889	2.46	0.930	2.26	3.22	$1.19 \cdot 10^{-9}$
cond2	5.1	629	0.79	0.734	1.79	2.95	$2.92 \cdot 10^{-9}$
cond3	5.1	629	0.40	0.656	2.20	2.75	$7.15 \cdot 10^{-9}$
cond4	5.0	410	1.15	0.868	1.60	3.11	$1.81 \cdot 10^{-9}$
cond5	5.2	889	1.36	1.14	2.36	2.96	$2.26 \cdot 10^{-9}$

Table 8.6: Summary of the optimized rlogL values (without rdf cut) for different detector conditions for the kappes neutrino flux assumption for MGRO J1908+06 for a 3σ discovery.

condition	rlogL	N_{tot}	n_{exp} $\cdot 10^{-4}$	β_{med} (deg)	$n_{3\sigma}$	MDP $\cdot 10^{-3}$	$\Phi_{3\sigma}$ (TeV^{-1}cm^{-2}s^{-1})
standard	5.1	629	10.7	0.358	0.790	3.62	$9.60 \cdot 10^{-10}$
cond1	5.2	889	11.5	0.352	0.990	3.47	$1.12 \cdot 10^{-9}$
cond2	5.3	1282	9.01	0.369	1.33	3.07	$1.92 \cdot 10^{-9}$
cond3	5.3	1282	5.42	0.361	1.17	2.98	$2.81 \cdot 10^{-9}$
cond4	5.1	629	7.10	0.443	0.990	3.17	$1.81 \cdot 10^{-9}$
cond5	5.1	629	3.14	0.412	0.880	2.94	$3.64 \cdot 10^{-9}$

Table 8.7: Summary of the optimized rlogL values (without rdf cut) for different detector conditions for the kappes neutrino flux assumption for MGRO J2019+37 for a 3σ discovery.

condition	rlogL	N_{tot}	n_{exp} $\cdot 10^{-4}$	β_{med} (deg)	$n_{3\sigma}$	MDP $\cdot 10^{-3}$	$\Phi_{3\sigma}$ ($\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
standard	5.3	1282	0.120	3.48	1.07	3.43	$8.03 \cdot 10^{-10}$
cond1	5.3	1282	9.01	3.28	1.08	3.25	$1.08 \cdot 10^{-9}$
cond2	5.3	1282	4.74	3.04	1.10	2.96	$2.09 \cdot 10^{-9}$
cond3	5.3	1282	2.91	2.87	0.97	2.87	$3.00 \cdot 10^{-9}$
cond4	5.3	1282	8.91	5.22	1.35	3.09	$1.36 \cdot 10^{-9}$
cond5	5.3	1282	4.95	5.14	1.55	2.86	$2.82 \cdot 10^{-9}$

Table 8.8: Summary of the optimized rlogL values (without rdf cut) for different detector conditions for the kappes neutrino flux assumption for MGRO J2031+41 for a 3σ discovery.

condition	rlogL	N_{tot}	n_{exp} $\cdot 10^{-4}$	β_{med} (deg)	$n_{3\sigma}$	MDP $\cdot 10^{-3}$	$\Phi_{3\sigma}$ ($\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
standard	5.2	889	7.83	3.83	1.12	3.13	$9.15 \cdot 10^{-10}$
cond1	5.2	889	5.68	3.73	1.05	3.05	$1.18 \cdot 10^{-9}$
cond2	5.2	889	2.82	3.47	0.960	2.89	$2.18 \cdot 10^{-9}$
cond3	5.2	889	1.42	3.23	0.960	2.79	$4.33 \cdot 10^{-9}$
cond4	5.1	629	3.41	4.56	0.890	2.94	$1.67 \cdot 10^{-9}$
cond5	5.2	889	3.09	5.07	1.29	2.83	$2.67 \cdot 10^{-9}$

Table 8.9: Summary of the optimized rlogL values (without rdf cut) for different detector conditions for the kappes neutrino flux assumption for C1 for a 3σ discovery.

condition	rlogL	N_{tot}	n_{exp} $\cdot 10^{-4}$	β_{med} (deg)	$n_{3\sigma}$	MDP $\cdot 10^{-3}$	$\Phi_{3\sigma}$ ($\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
standard	5.2	889	7.70	3.93	1.05	3.18	$6.68 \cdot 10^{-10}$
cond1	5.3	1282	8.16	4.05	1.35	3.04	$8.11 \cdot 10^{-10}$
cond2	5.3	1282	4.47	3.76	1.30	2.88	$1.43 \cdot 10^{-9}$
cond3	5.2	889	1.65	3.37	1.00	2.77	$2.97 \cdot 10^{-9}$
cond4	5.1	629	3.20	4.81	0.990	2.90	$1.52 \cdot 10^{-9}$
cond5	5.3	1282	4.49	5.66	1.61	2.81	$1.76 \cdot 10^{-9}$

Table 8.10: Summary of the optimized rlogL values (without rdf cut) for different detector conditions for the kappes neutrino flux assumption for C2 for a 3σ discovery.

8.3 Summary of the results and comparison to other analyses

A short summary of the neutrino flux predictions and optimized cuts for the best 3σ discovery potential per source is presented in Tab. 8.11. The signal PDFs per source for the mentioned fluxes and cuts in Tab. 8.11 are shown in the appendix C.2 Fig. C.1. The sensitivity of these Milagro sources for a 3σ discovery is shown in Fig. 8.14.

source	flux	rlogL	rdf	$n_{3\sigma}$	sensitivity ($\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
MGRO J1908+06	halzen3	5.1	-	1.81	$6.3 \cdot 10^{-10}$
MGRO J2019+37	halzen3	5.2	1	0.93	$3.5 \cdot 10^{-10}$
MGRO J2031+41	kappes	5.3	-	1.07	$8.0 \cdot 10^{-10}$
C1	kappes	5.2	1	1.01	$8.7 \cdot 10^{-10}$
C2	kappes	5.2	1	0.94	$6.3 \cdot 10^{-10}$

Table 8.11: Summary of the flux assumptions and optimized cuts for the best 3σ discovery potential per source. The number of $n_{3\sigma}$ and the normalization of the sensitivity are also given.

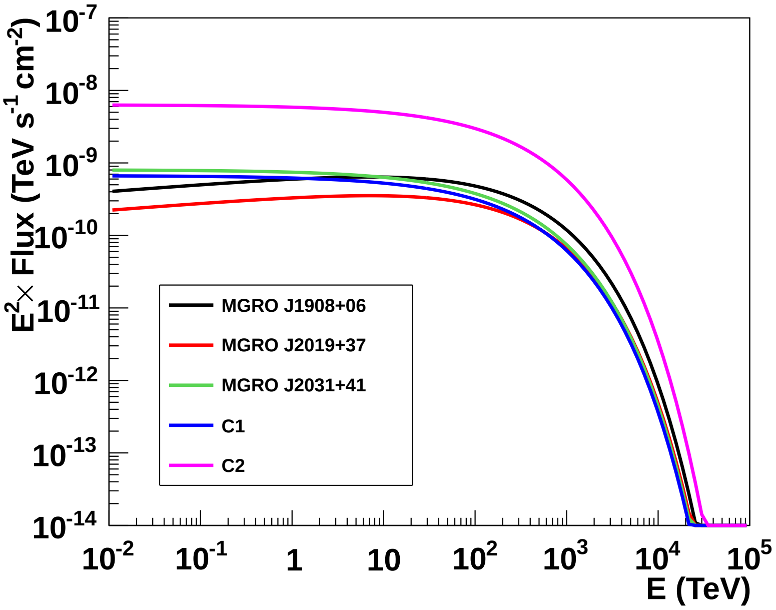


Figure 8.14: Sensitivity of the five Milagro sources for a 3σ discovery for the flux assumptions and optimized cuts of Tab. 8.11.

With the neutrino flux determined from the updated measurements of the gamma ray flux of the sources MGRO J1908+06, MGRO J2019+37 and MGRO J2031+41, it is stated in [5] that IceCube might be able to exclude two of the five Milagro sources, MGRO J1908+06 and MGRO J2019+37, being PeVatrons within several years, if no signal event is found. The best results could as well be obtained for the neutrino flux assumptions with the highest cut-off energy of

$E_{cut,\gamma} = 300\,\text{TeV}$ and the lowest spectral index which was also found in this thesis, and can be explained by the amount of expected signal events, which is highest for this flux assumption. The larger the spectral index the longer it takes to exclude these sources. For MGRO J2031+41 the lowest spectral index is 2.6 (from the halzen parametrization [5]) and IceCube would need more than 10 years to exclude this source. As the detector volume is large and four of the five sources are better visible with the IceCube detector than with ANTARES, the chance to measure a significant amount of cosmic neutrinos from these sources with IceCube is higher than for ANTARES. For example for MGRO J2019+37 with the corresponding flux halzen3 a 3σ discovery would be possible within 4 years and for MGRO J1908+06 within 3 years with the IceCube detector [5]. As the number of expected neutrino events for ANTARES is of the order of 10^{-3} for a live-time of about 2 years, enlarging the live-time by a factor 2 won't change the significance for these sources much for the ANTARES detector, but could lead to a better upper limit. In the latest point source analysis from IceCube with four years of data they studied the two sources MGRO J1908+06 and MGRO J2019+37 and could not find any signal event so a 90% C.L. upper limit is set on the neutrino flux assuming an overall E^{-2} energy spectrum instead of a neutrino flux assumption derived from gamma ray measurements per source [7]. The upper limits are for MGRO J1908+06 $\Phi_{90} = 0.71 \cdot 10^{-12}\,\text{TeV}^{-1}\text{cm}^{-2}\text{s}^{-1}$ (shown in Fig. 8.15 pink line) and for MGRO J2019+37 $\Phi_{90} = 1.36 \cdot 10^{-12}\,\text{TeV}^{-1}\text{cm}^{-2}\text{s}^{-1}$ [7]. The energy range for northern hemisphere sources is for IceCube 1 TeV to 1 PeV.

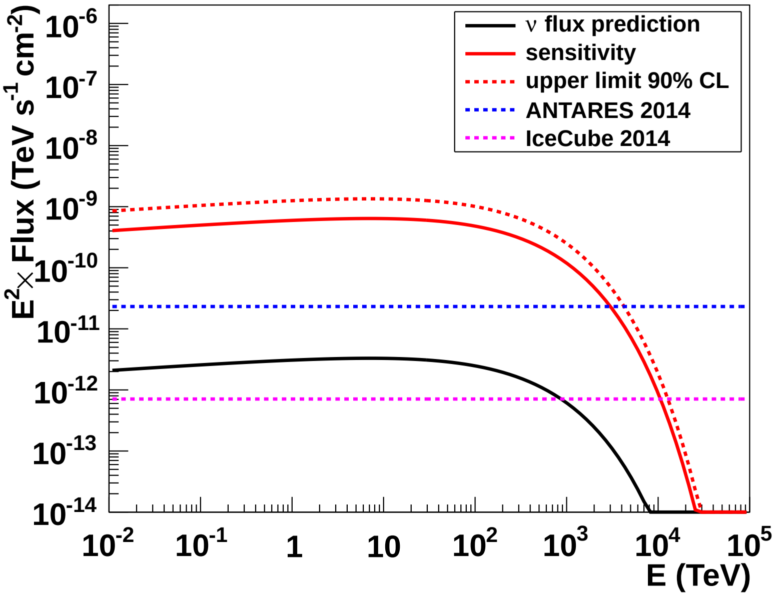


Figure 8.15: Neutrino flux assumption (black), sensitivity (red continuous line) for a 3σ discovery and average upper limit 90% C.L. (red dotted line) for MGRO J1908+06 for the halzen3 neutrino flux assumption. Two 90% C.L. upper limits are shown assuming an E^{-2} energy spectrum without cut-off energy determined for MGRO J1908+06 for ANTARES in blue (derived in [6]) and for IceCube in pink (derived in [7]).

MGRO J1908+06, with a visibility of 48% within the ANTARES detector, was also studied in the candidate list search of the official ANTARES point source analysis with 2007-2012 data using the AAFit reconstruction strategy [6]. In that study an overall E^{-2} energy spectrum without cut-off energy was assumed. No significant amount of signal events was found in data, so a 90% C.L. upper limit was set with a flux normalization of $E^2 \times \Phi_{90} = 2.32 \cdot 10^{-8}\,\mathrm{GeV\,cm^{-2}s^{-1}}$. This upper limit is compared in Fig. 8.15 (blue line) to the average upper limit (red dotted line) and sensitivity for a 3σ discovery (red continuous line) determined in this thesis for the neutrino flux assumption halzen3 (black line), derived from gamma ray spectra measurements of this source, which are better motivated for this source than an E^{-2} energy spectrum. The blue curve is nearer to the assumed neutrino flux, but it cannot exclude this model. The facts that the underlying neutrino flux spectra is without cut-off energy and that the considered live-time is a factor 1.8 higher (1338 days compared to 744.68 days) explains this difference between the upper limit derived in the analysis shown in this thesis and the official point source analysis, in which the events would be mainly high-energetic.

8.4 Conclusion

The sensitivity that could be determined performing pseudo-experiments for the five Milagro sources (see Tab. 8.11 in section 8.3) is for all sources about a factor 10^2 higher then their expected neutrino flux. The error on the sensitivity for different detector conditions is most influenced by the number of active OMs. The average relative upper uncertainties on the sensitivity is about 131% and the average relative lower one is about 41%. Further the number of expected events determined from the simulation of the assumed point-like sources for the ANTARES detector is of the order of 10^{-3} (or even below for MGRO J2031+41, which is with a declination of 41° at the borders of the visibility of ANTARES) for a live-time of 2.04 years. Therefore it is unlikely to measure a significant amount of neutrinos sufficient for a 3σ or a 5σ discovery. This is caused by the fact that the visibility of these sources within the ANTARES detector is low (around 23% to 29%) except for MGRO J1908+06 with a visibility of 48% and by the fact that the expected neutrino spectrum has an exponential cut-off and thus the neutrinos will be mostly low-energetic.

Neither the latest point source search of ANTARES with a live-time of a factor 1.8 higher [6], nor the analysis of IceCube, with a higher visibility of these sources and a much larger detector volume, could measure any event connected to MGRO J1908+06 assuming an E^{-2} energy spectrum [7]. To not constrain future analysis plans on these five Milagro sources with a trial factor, the data are not unblinded in this thesis. Additionally a complete run-by-run signal simulation, which is with several CPU-years very time consuming, would be needed before unblinding the data as the relative uncertainties on the sensitivity determined by systematically changing the detector conditions are too large.

Some possible adjustments to the analysis method that could be made to improve the upper limit on the neutrino flux predictions are discussed here. There is no cut on the neutrino or muon energy applied in this study, which could further suppress the background of atmospheric muons and neutrinos. But it would also lead to a decrease of expected cosmic neutrino events from these sources. So I suspect that only cutting on the energy will not increase the sensitivity to a level where an unblinding process would be sensible. A better solution would be using in the signal PDF additionally to the spacing distribution of the events to the sources also an energy dependent component, which was shown to lower the number of necessary events (assuming a power-law without exponential cut-off) [8]. Further it could be studied if a stacking analysis approach with these five sources could lower the upper limit, but it is unlikely to lead to a detection as the trial factor would rise to 5 and therefore the number of events necessary for 3σ significance would rise as well. A good compromise would be to maybe combine the enlarged signal PDF with energy information with a stacking analysis approach and using events that are reconstructed with KrakeFit and/or AAFit to get a higher amount of neutrinos after quality cuts.

In this thesis it is shown that ANTARES is not able to detect or constrain physically motivated neutrino flux predictions derived from gamma ray measurements of these sources, which are better motivated than an E^{-2} energy spectrum. By adjusting the method used in this thesis, the upper limit could maybe be improved, but a detection of these Milagro sources is unlikely.

Appendix A

Definitions for the Milagro sources

A.1 Definition of the neutrino flux parametrization per Milagro source

The sources MGRO J2031+41, MGRO J1908+06 and MGRO J2019+37 have 9 different neutrino flux parametrizations from [5] and one from [4]. Therefore a short naming: halzen1 to halzen9 and kappes is introduced to distinguish between the different parameters. The mapping of names to the corresponding parametrizations is given here.

The neutrino flux is then calculated using eqn. (6.2) from section 6.2, where the definition $k_\gamma = \frac{K}{E_{norm}^{-\alpha_\gamma}}$ is used. Please see section 6.2 for the definition of the parameters k_ν, α_ν and $E_{cut,\nu}$. The calculated k_ν is given in the tables as well, as this is the normalization of the flux which is used to calculate the sensitivity and 90% C.L. upper limits.

A.1.1 MGRO J1908+06

name	K [TeV^{-1}cm^{-2}s^{-1}]	E_{norm} [TeV]	α_γ	$E_{cut,\gamma}$[TeV]	k_ν [TeV^{-1}cm^{-2}s^{-1}]
halzen1	$6.1\cdot10^{-13}$	4.0	1.9	30.0	$3.3\cdot10^{-12}$
halzen2	$6.1\cdot10^{-13}$	4.0	1.9	100.0	$3.3\cdot10^{-12}$
halzen3	$6.1\cdot10^{-13}$	4.0	1.9	300.0	$3.3\cdot10^{-12}$
halzen4	$6.1\cdot10^{-13}$	4.0	2.1	30.0	$4.0\cdot10^{-12}$
halzen5	$6.1\cdot10^{-13}$	4.0	2.1	100.0	$4.0\cdot10^{-12}$
halzen6	$6.1\cdot10^{-13}$	4.0	2.1	300.0	$4.0\cdot10^{-12}$
halzen7	$6.1\cdot10^{-13}$	4.0	2.3	30.0	$4.8\cdot10^{-12}$
halzen8	$6.1\cdot10^{-13}$	4.0	2.3	100.0	$4.8\cdot10^{-12}$
halzen9	$6.1\cdot10^{-13}$	4.0	2.3	300.0	$4.8\cdot10^{-12}$

Table A.1: Definition of the names and main parameters of the Halzen neutrino flux parametrizations for MGRO J1908+06.

A.1.2 MGRO J2019+37

name	K [TeV^{-1}cm^{-2}s^{-1}]	E_{norm} [TeV]	α_γ	$E_{cut,\gamma}$[TeV]	k_ν [TeV^{-1}cm^{-2}s^{-1}]
halzen1	$7.0{\cdot}10^{-14}$	10.0	1.8	30.0	$1.8{\cdot}10^{-12}$
halzen2	$7.0{\cdot}10^{-14}$	10.0	1.8	100.0	$1.8{\cdot}10^{-12}$
halzen3	$7.0{\cdot}10^{-14}$	10.0	1.8	300.0	$1.8{\cdot}10^{-12}$
halzen4	$7.0{\cdot}10^{-14}$	10.0	2.0	30.0	$2.6{\cdot}10^{-12}$
halzen5	$7.0{\cdot}10^{-14}$	10.0	2.0	100.0	$2.6{\cdot}10^{-12}$
halzen6	$7.0{\cdot}10^{-14}$	10.0	2.0	300.0	$2.6{\cdot}10^{-12}$
halzen7	$7.0{\cdot}10^{-14}$	10.0	2.2	30.0	$3.8{\cdot}10^{-12}$
halzen8	$7.0{\cdot}10^{-14}$	10.0	2.2	100.0	$3.8{\cdot}10^{-12}$
halzen9	$7.0{\cdot}10^{-14}$	10.0	2.2	300.0	$3.8{\cdot}10^{-12}$

Table A.2: Definition of the names and main parameters of the Halzen neutrino flux parametrizations for MGRO J2019+37.

A.1.3 MGRO J2031+41

name	K [TeV^{-1}cm^{-2}s^{-1}]	E_{norm} [TeV]	α_γ	$E_{cut,\gamma}$[TeV]	k_ν [TeV^{-1}cm^{-2}s^{-1}]
halzen1	$2.1{\cdot}10^{-14}$	10.0	2.6	30.0	$2.3{\cdot}10^{-12}$
halzen2	$2.1{\cdot}10^{-14}$	10.0	2.6	100.0	$2.3{\cdot}10^{-12}$
halzen3	$2.1{\cdot}10^{-14}$	10.0	2.6	300.0	$2.3{\cdot}10^{-12}$
halzen4	$2.1{\cdot}10^{-14}$	10.0	2.8	30.0	$3.3{\cdot}10^{-12}$
halzen5	$2.1{\cdot}10^{-14}$	10.0	2.8	100.0	$3.3{\cdot}10^{-12}$
halzen6	$2.1{\cdot}10^{-14}$	10.0	2.8	300.0	$3.3{\cdot}10^{-12}$
halzen7	$2.1{\cdot}10^{-14}$	10.0	3.0	30.0	$4.5{\cdot}10^{-12}$
halzen8	$2.1{\cdot}10^{-14}$	10.0	3.0	100.0	$4.5{\cdot}10^{-12}$
halzen9	$2.1{\cdot}10^{-14}$	10.0	3.0	300.0	$4.5{\cdot}10^{-12}$

Table A.3: Definition of the names and main parameters of the Halzen neutrino flux parametrizations for MGRO J2031+41.

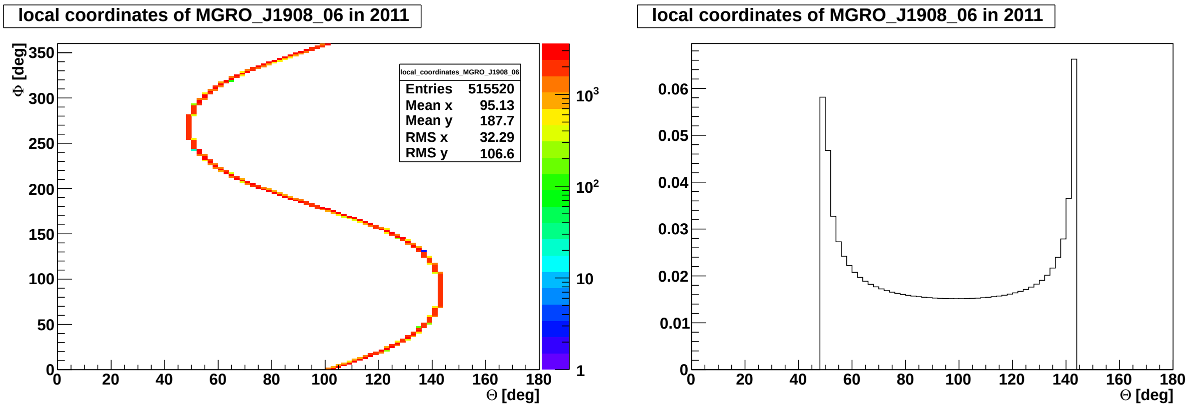
A.1.4 Kappes parametrization of the neutrino fluxes per Milagro source

source	name	k_γ [TeV^{-1}cm^{-2}s^{-1}]	α_γ	$E_{cut,\gamma}$[TeV]	k_ν [TeV^{-1}cm^{-2}s^{-1}]
MGRO J1908+06	kappes	$3.4 \cdot 10^{-12}$	2.01	300	$1.3{\cdot}10^{-12}$
MGRO J2019+37	kappes	$3.5 \cdot 10^{-12}$	2.00	300	$1.3{\cdot}10^{-12}$
MGRO J2031+41	kappes	$2.4 \cdot 10^{-12}$	2.00	300	$9.0{\cdot}10^{-13}$
C1	kappes	$1.7 \cdot 10^{-12}$	2.00	300	$6.4{\cdot}10^{-13}$
C2	kappes	$1.3 \cdot 10^{-12}$	2.00	300	$4.9{\cdot}10^{-13}$

Table A.4: Gamma ray flux parametrization of the Milagro sources determined in [4], exact values recieved from [59].

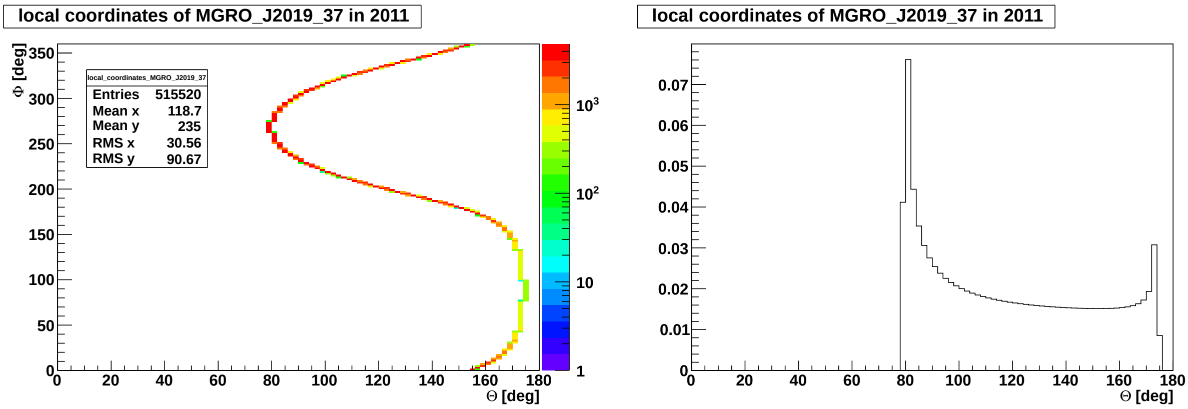
A.2 Visibility of the Milagro sources in the ANTARES detector

The coordinates of the Milagro sources given in section 6.1 are transformed into local detector coordinates Θ and Φ for one year of data taking using the astro package. The figures Fig. A.2 to Fig. A.5 show on the left the direction in local coordinates of the ANTARES detector and on the right the distribution of Θ, which is the projection on the x-axis for the year 2011 as an example. For an angle $\Theta < 90°$ the track directions are upgoing within the detector. The fraction of time in which the source is seen in upward direction gives the visibility, which is determined to 48% for MGRO J1908+06, 28% for MGRO J2019+37, 23% for MGRO J2031+41, 29% for C1 and to 28% for C2.


local coordinates of MGRO_J1908_06 in 2011

local coordinates of MGRO_J1908_06 in 2011

Figure A.1: Possible local coordinates (Θ, Φ) of MGRO J1908+06 within the ANTARES detector for one year (left), distribution of the Θ coordinate (right).


local coordinates of MGRO_J2019_37 in 2011

local coordinates of MGRO_J2019_37 in 2011

Figure A.2: Possible local coordinates (Θ, Φ) of MGRO J2019+37 within the ANTARES detector for one year (left), distribution of the Θ coordinate (right).

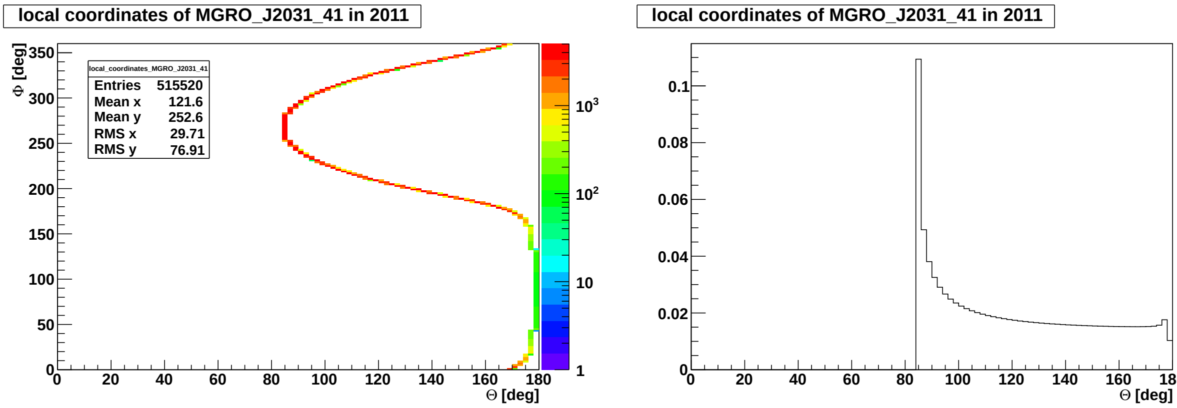


Figure A.3: Possible local coordinates (Θ, Φ) of MGRO J2031+41 within the ANTARES detector for one year (left), distribution of the Θ coordinate (right).

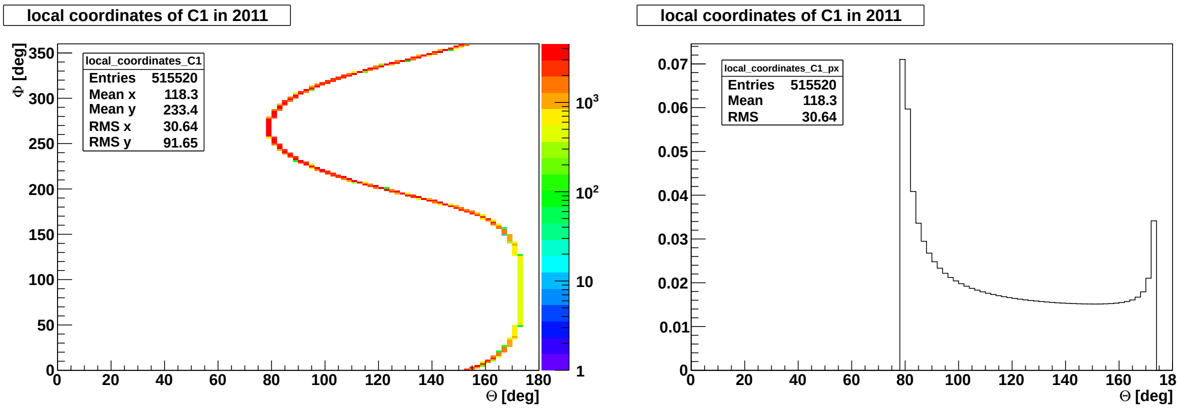


Figure A.4: Possible local coordinates (Θ, Φ) of C1 within the ANTARES detector for one year (left), distribution of the Θ coordinate (right).

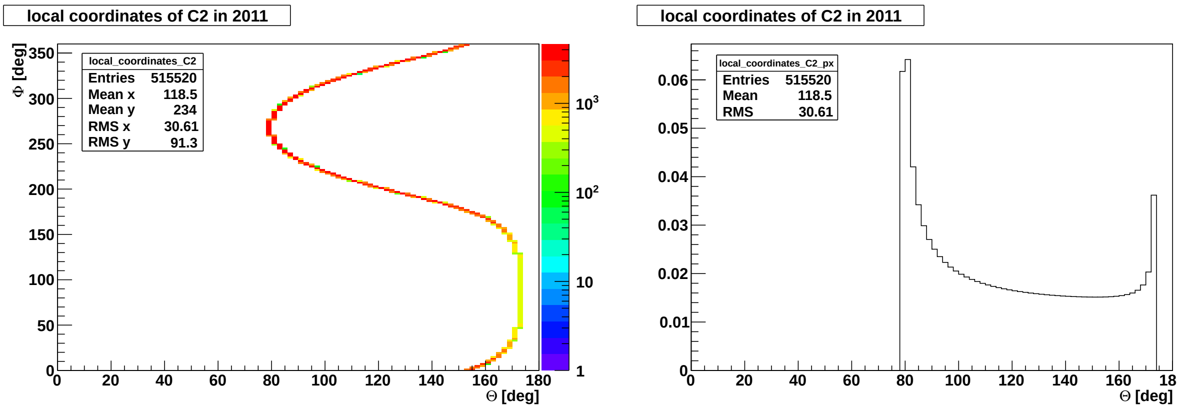


Figure A.5: Possible local coordinates (Θ, Φ) of C2 within the ANTARES detector for one year (left), distribution of the Θ coordinate (right).

A.3 Run selection for MC simulation

For each Milagro source a point-like MC simulation is performed (see section 6.3) to estimate the detector response for the sources and the number of expected events. As the MC simulation for all runs takes several CPU years for one source, a subsample of runs is selected being representatives of the majority of ANTARES runs, which is defined as standard detector condition with baseline rate $(77,6 \pm 10\%)$ kHz and number of active OMs $700 \pm 10\%$. To study the systematic effect of different detector conditions on the sensitivity per source, five more detector conditions are defined three of them changing the baseline rate by keeping the number of active OMs fixed (cond1 to cond3) and two the other way around (cond4 and cond5):

- Different baseline rates:
  - cond1: baseline rate $(100 \pm 10\%)$ kHz, nactiveOMs $700 \pm 10\%$
  - cond2: baseline rate $(150 \pm 10\%)$ kHz, nactiveOMs $700 \pm 10\%$
  - cond3: baseline rate $(200 \pm 10\%)$ kHz, nactiveOMs $700 \pm 10\%$
- Different nactiveOMs:
  - cond4: baseline rate $(77,6 \pm 10\%)$ kHz, nactiveOMs $600 \pm 10\%$
  - cond5: baseline rate $(77,6 \pm 10\%)$ kHz, nactiveOMs $500 \pm 10\%$

The runs representing these conditions used in the simulation process are summarized in Tab. A.5 to Tab. A.10.

Runs for condition standard

RunID	baseline [kHz]	burstfr [kHz]	QB	nactiveOMs	deadChannels	Runduration [min]
62112	72.6666	0.09181	4.0	721.0	141.0	136.666
64128	81.0798	0.18931	4.0	726.0	151.0	169.522
65963	70.0556	0.19007	4.0	708.0	156.0	185.643
66105	76.1213	0.12057	4.0	708.0	157.0	182.181
66185	77.3974	0.19489	4.0	712.0	157.0	165.084
66304	85.1206	0.17617	4.0	708.0	157.0	167.43
66324	79.879	0.12874	4.0	709.0	157.0	176.533
66341	80.2788	0.11173	4.0	708.0	157.0	171.042
66365	81.0201	0.14107	4.0	708.0	157.0	172.999
66378	80.7412	0.18279	4.0	708.0	157.0	174.13

Table A.5: Overview of the 10 runs selected for the MC simulation for detector condition standard.

Runs for condition 1 to 3

RunID	baseline [kHz]	burstfr [kHz]	QB	nactiveOMs	deadChannels	Runduration [min]
65249	94.5845	0.17911	4.0	721.0	155.0	155.263
65256	92.2164	0.13523	4.0	709.0	155.0	156.018
65259	93.7006	0.12962	4.0	718.0	155.0	150.785
65612	109.636	0.19226	4.0	718.0	156.0	136.708
65613	107.64	0.18842	4.0	718.0	156.0	135.7
65617	108.532	0.19703	4.0	717.0	156.0	128.666
65640	108.478	0.1885	4.0	719.0	156.0	130.023
65645	103.71	0.17682	4.0	720.0	156.0	140.133
65654	94.4125	0.18285	4.0	715.0	156.0	151.009
65656	92.5604	0.16224	4.0	713.0	156.0	156.435

Table A.6: Overview of the 10 runs selected for the MC simulation for detector condition 1.

RunID	baseline [kHz]	burstfr [kHz]	QB	nactiveOMs	deadChannels	Runduration [min]
57056	145.079	0.15234	2.0	725.0	129.0	112.565
63555	149.0	0.86268	2.0	741.0	147.0	152.589
64933	161.909	0.43491	2.0	728.0	155.0	140.778
65019	135.096	0.39196	2.0	726.0	155.0	102.423
65223	159.636	0.17331	2.0	726.0	155.0	107.269
65234	147.52	0.14598	2.0	724.0	155.0	126.194
65371	143.287	0.2658	2.0	724.0	155.0	137.451
65563	154.119	0.2841	2.0	722.0	156.0	127.325
65817	154.057	0.38093	2.0	716.0	156.0	135.639
66014	153.422	0.28963	2.0	715.0	156.0	122.809

Table A.7: Overview of the 10 runs selected for the MC simulation for detector condition 2.

RunID	baseline [kHz]	burstfr [kHz]	QB	nactiveOMs	deadChannels	Runduration [min]
63208	198.3	0.46859	2.0	737.0	147.0	123.652
64879	186.591	0.42403	2.0	727.0	155.0	134.684
64942	214.565	0.43568	2.0	728.0	155.0	101.708
64958	192.194	0.43584	2.0	714.0	155.0	136.962
64976	212.298	0.29988	2.0	728.0	155.0	137.975
65050	185.512	0.41803	2.0	728.0	155.0	127.231
65809	205.477	0.22736	2.0	716.0	156.0	108.793
65814	192.875	0.3736	2.0	719.0	156.0	102.163
65881	200.461	0.24359	2.0	724.0	156.0	116.353
65911	213.243	0.42052	2.0	714.0	156.0	105.324

Table A.8: Overview of the 10 runs selected for the MC simulation for detector condition 3.

Runs for condition 4 and 5

RunID	baseline [kHz]	burstfr [kHz]	QB	nactiveOMs	deadChannels	Runduration [min]
47949	71.1792	0.11834	4.0	601.0	114.0	181.564
49911	82.4936	0.1208	4.0	603.0	117.0	200.404
50359	71.224	0.09554	4.0	611.0	118.0	223.833
50843	80.1074	0.15363	4.0	610.0	119.0	208.614
51088	76.8463	0.17889	4.0	611.0	119.0	207.669
51133	71.8184	0.06826	4.0	621.0	119.0	172.129
51359	74.6346	0.19504	4.0	607.0	119.0	210.69
51497	70.8965	0.1419	4.0	611.0	119.0	217.513
51586	77.6147	0.17842	4.0	607.0	119.0	121.666
51731	70.147	0.17066	4.0	610.0	119.0	176.542

Table A.9: Overview of the 10 runs selected for the MC simulation for detector condition 4.

RunID	baseline [kHz]	burstfr [kHz]	QB	nactiveOMs	deadChannels	Runduration [min]
46233	69.9528	0.65421	1.0	527.0	112.0	197.206
47900	70.1451	0.55697	1.0	514.0	114.0	119.345
48372	72.6641	0.51039	1.0	540.0	115.0	117.15
49251	82.6293	0.50209	1.0	541.0	116.0	181.437
50244	77.1029	0.49717	1.0	546.0	117.0	200.173
50402	79.305	0.45885	1.0	546.0	118.0	111.647
51477	81.578	0.59931	1.0	523.0	119.0	183.791
51757	85.0753	0.62698	1.0	509.0	119.0	184.156
52092	72.4092	0.39213	1.0	525.0	120.0	165.035
56059	81.0598	0.78349	1.0	504.0	128.0	118.061

Table A.10: Overview of the 10 runs selected for the MC simulation for detector condition 5.

Appendix B

Data-MC-comparison of KrakeFit

B.1 Data-MC-comparison plots for the prefit of KrakeFit

For the official data production the output per reconstruction strategy has to be reduced to the minimal necessary information used for the different analyses as the storage space is limited. So not the whole information, e.g. hits, per fitting step of KrakeFit is available for all runs. To study the agreement between data and MC for the different prefit steps one run was chosen. A low-statistics data-MC-comparison was made for the single steps of the prefit of KrakeFit. For the run 37128 data and atmospheric muons are compared. As there are no quality cuts applied the distributions are normalized so that the shape of the curves can be compared. Here, the FilteringFit algorithm is started with $60 \cdot 10^3$ directions and an older version of the rbr MC simulation is used, (rbr v0.1 instead of v2.0.1) as the newer version was not available at the time of this comparison. The default step FilteringFit maxQ is shown in Fig. B.1, FilFitsDTrackSelection in Fig. B.2 und the prefit result FilFitsDMEstimator in Fig. B.3 where the shape of the curves of data and MC fits well. The number of atmospheric muons has to be weighted by 10 to correspond to the live-time of the selected run, which explains the statistical uncertainties for atmospheric muons.

B.2 Data-MC-comparison plots of KrakeFit

More data-MC-comparison plots for the final fit result of KrakeFit are shown here with the runselections totRuns (0Runs) with a live-time of 744.68 days (70.67 days). The distributions and ratios of data-to-MC of zenith angle, azimuth angle, number of lines and number of hits used in the reconstruction for events reconstructed as upgoing are shown for the cuts $\text{rlogL} < 6.0$ for 0Runs selection in Fig. B.4, for $\text{rlogL} < 5.5$ and rdf=1 for totRuns selection in Fig. B.5 and without cuts in Fig. B.6. The color code is in all plots the same: for data black, for atmospheric muons red, for atmospheric neutrinos green, for the sum of atmospheric neutrinos and muons blue. If the ratio of data to MC is calculated, the MC is represented by the sum of atmospheric neutrinos and muons.

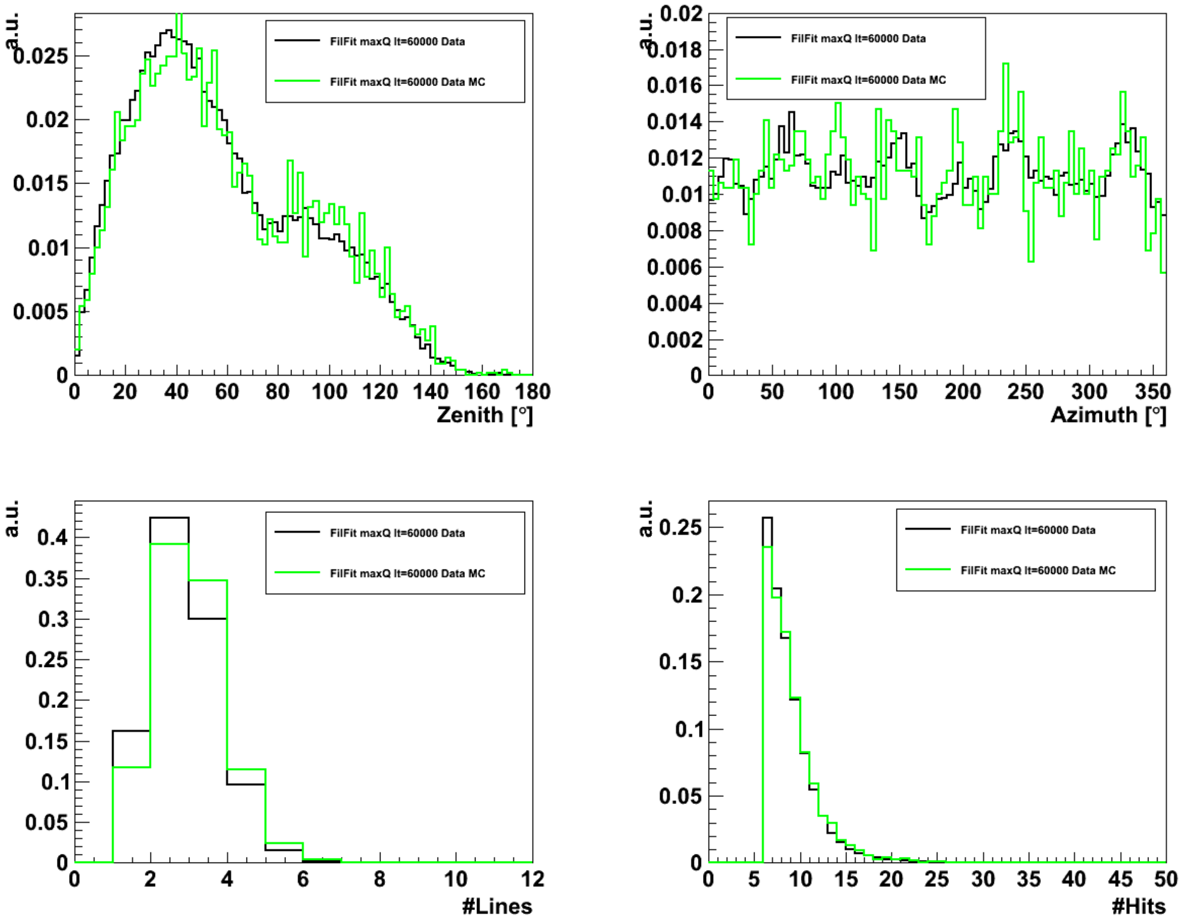


Figure B.1: Distribution of reconstructed zenith angle (top left), azimuth angle (top right), number of lines (bottom left) and hits (bottom right) used in the reconstruction of FilteringFit 'maxQ' for data (black) and atmospheric muons (green) for run 37128. All distributions are normalized.

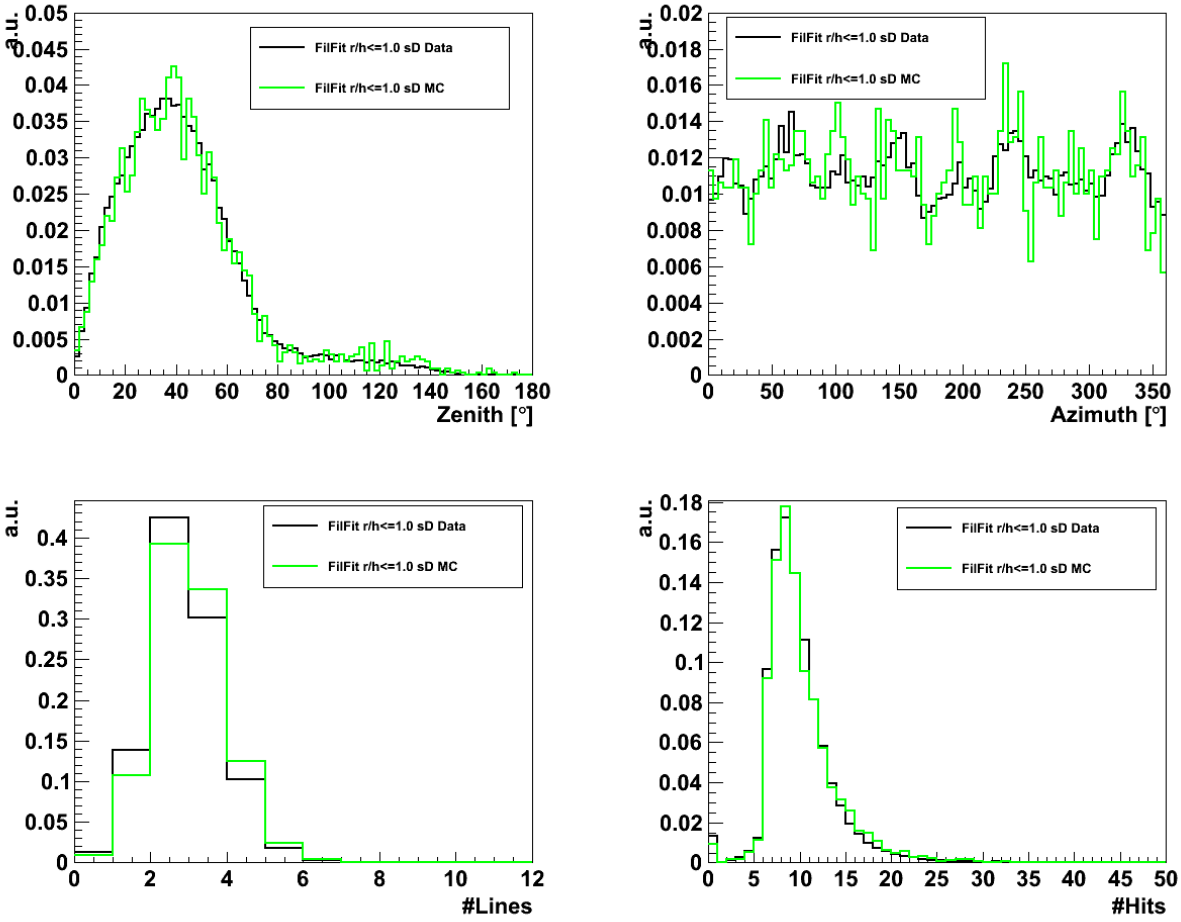


Figure B.2: Distribution of reconstructed zenith angle (top left), azimuth angle (top right), number of lines (bottom left) and hits (bottom right) used in the reconstruction of FilteringFitsD-TrackSelection for data (black) and atmospheric muons (green) for run 37128. All distributions are normalized.

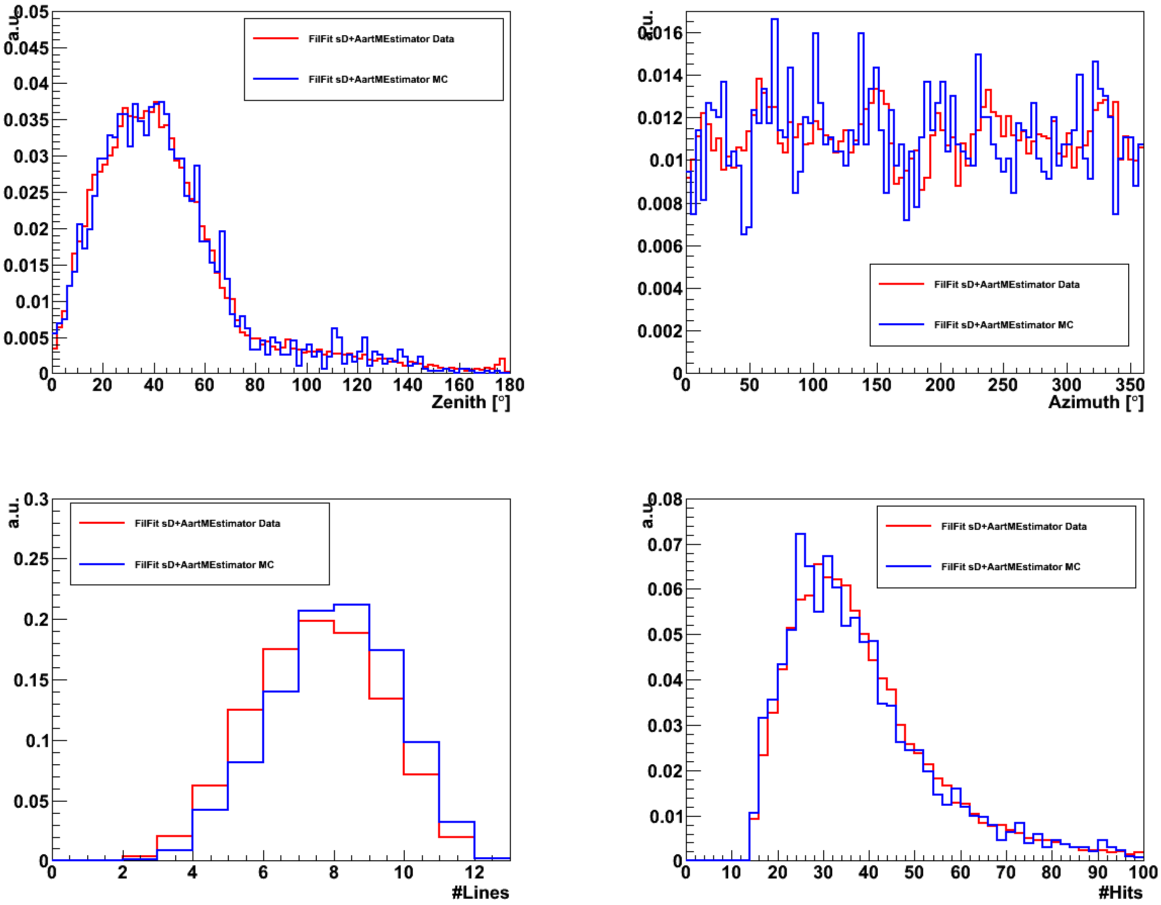


Figure B.3: Distribution of reconstructed zenith angle (top left), azimuth angle (top right), number of lines (bottom left) and hits (bottom right) used in the reconstruction of FilteringFitsDMEstimator for data (blue) and atmospheric muons (red) for run 37128. All distributions are normalized.

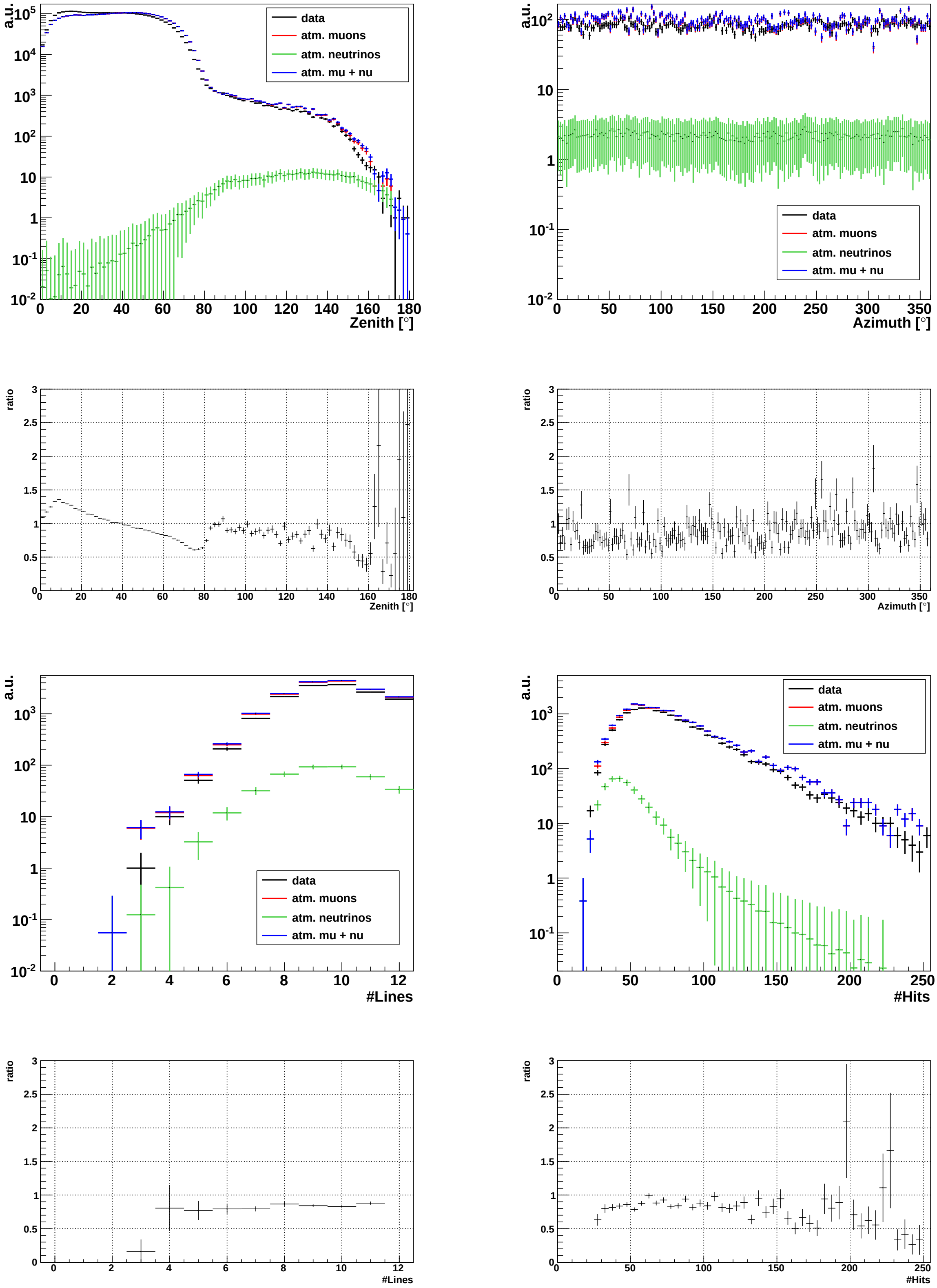


Figure B.4: 0Runs selection: Distribution of reconstructed zenith and azimuth angle, number of lines and hits used in the reconstruction with KrakeFit with rlogL < 6.0. The respective bottom plots show the ratio between data and MC for the distributions above.

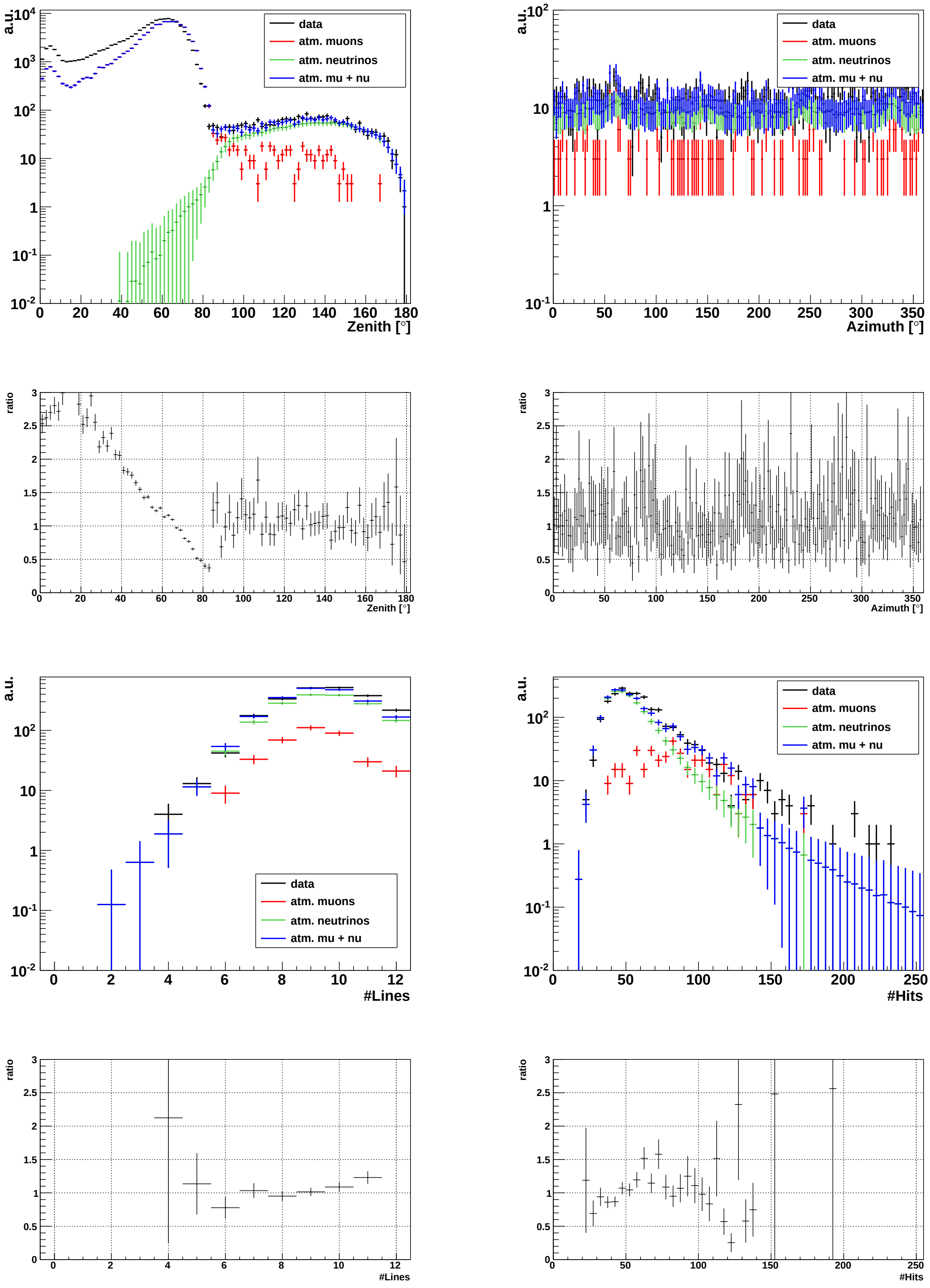


Figure B.5: Distribution of reconstructed zenith and azimuth angle, number of lines and hits used in the reconstruction with KrakeFit with rlogL < 5.5 and rdf=1. The respective bottom plots show the ratio between data and MC for the distributions above.

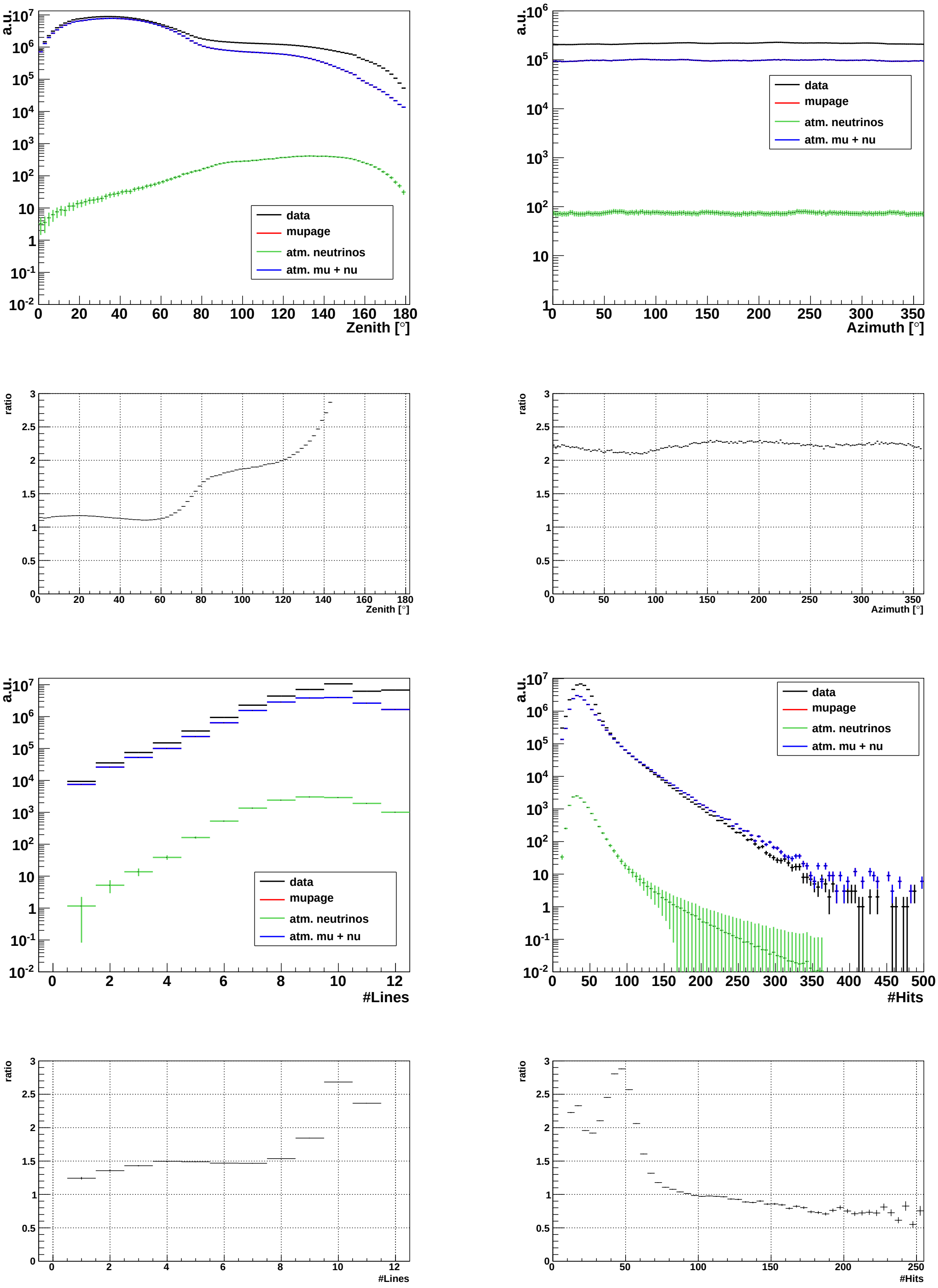


Figure B.6: Distribution of reconstructed zenith and azimuth angle, number of lines and hits used in the reconstruction with KrakeFit (no cuts applied). The respective bottom plots show the ratio between data and MC for the distributions above.

Appendix C

Point source search results

C.1 Summary of the results for 5σ optimization

The following variables are extracted and summarized for the cuts optimized for a 5σ discovery with the MDP method per source and neutrino flux assumption.

- N_{tot}: total number of events in the data sample,
- n_{exp}: number of expected events derived from the MC simulation per source weighted according to the neutrino flux assumption (see section 6.3),
- β_{med}: median of the angle β between source coordinates and reconstructed muons,
- $n_{5\sigma}$: mean number of events necessary for a 3σ discovery,
- MDP: model discovery potential calculated via eqn. (7.23) $MDP(n_{exp})$,
- $\Phi_{5\sigma}$: normalization of the 3σ sensitivity see eqn. (7.25).

flux	rlogL	rdf	N_{tot}	n_{exp} $\cdot 10^{-3}$	β_{med} (deg)	$n_{5\sigma}$	MDP $\cdot 10^{-5}$	$\Phi_{5\sigma}$ ($\mathrm{TeV^{-1}cm^{-2}s^{-1}}$)
kappes	5.5	-	2786	5.70	1.54	8.97	2.89	$2.05 \cdot 10^{-9}$
	5.2	1	850	3.19	1.08	4.31	1.23	$1.76 \cdot 10^{-9}$
halzen1	5.5	-	2786	9.35	1.56	8.97	0.915	$3.17 \cdot 10^{-9}$
	5.5	1	2193	8.87	1.55	8.53	1.00	$3.17 \cdot 10^{-9}$
halzen2	5.0	-	410	4.29	0.821	3.36	0.577	$2.58 \cdot 10^{-9}$
	5.2	1	850	8.43	1.08	4.53	1.75	$1.77 \cdot 10^{-9}$
halzen3	5.1	-	629	9.29	0.936	3.93	1.78	$1.40 \cdot 10^{-9}$
	5.0	1	403	6.08	0.824	3.44	2.82	$1.87 \cdot 10^{-9}$
halzen4	5.2	-	889	3.78	1.08	4.97	0.913	$5.26 \cdot 10^{-9}$
	5.4	1	1580	5.88	1.38	6.97	1.03	$4.74 \cdot 10^{-9}$
halzen5	5.5	-	2786	11.0	1.55	8.73	1.14	$3.17 \cdot 10^{-9}$
	5.1	1	606	4.23	0.940	3.87	1.37	$3.66 \cdot 10^{-9}$
halzen6	5.5	-	2786	14.2	1.54	8.86	2.99	$2.50 \cdot 10^{-9}$
	5.2	1	850	7.68	1.08	4.68	1.38	$2.44 \cdot 10^{-9}$
halzen7	5.5	-	2786	6.29	1.56	8.70	2.78	$6.64 \cdot 10^{-9}$
	5.5	1	2193	5.90	1.56	8.51	0.0635	$6.92 \cdot 10^{-9}$
halzen8	5.1	-	629	3.13	0.942	3.99	0.281	$6.12 \cdot 10^{-9}$
	5.5	1	2193	8.21	1.55	8.10	1.07	$4.74 \cdot 10^{-9}$
halzen9	5.0	-	410	2.58	0.834	3.51	1.18	$6.53 \cdot 10^{-9}$
	5.1	1	606	3.89	0.943	4.07	1.10	$5.02 \cdot 10^{-9}$

Table C.1: Summary of the optimized rlogL values (with rdf cut) for different neutrino flux assumption for MGRO J1908+06 for a 5σ discovery.

flux	rlogL	rdf	N_{tot}	n_{exp} $\cdot 10^{-3}$	β_{med} (deg)	$n_{5\sigma}$	MDP $\cdot 10^{-7}$	$\Phi_{5\sigma}$ ($\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
kappes	5.5	-	2786	3.41	0.530	4.53	3.73	$1.73 \cdot 10^{-9}$
	5.6	1	3518	3.65	0.576	4.84	3.96	$1.72 \cdot 10^{-9}$
halzen1	5.5	-	2786	3.28	0.570	4.53	6.94	$2.49 \cdot 10^{-9}$
	5.6	1	3518	3.54	0.632	4.84	94.4	$2.46 \cdot 10^{-9}$
halzen2	5.5	-	2786	6.25	0.525	4.52	13.5	$1.30 \cdot 10^{-9}$
	5.6	1	3518	6.68	0.570	4.72	14.3	$1.27 \cdot 10^{-9}$
halzen3	5.5	-	2786	9.72	0.494	4.73	1.81	8.76e-10
	5.6	1	3518	10.3	0.534	4.73	13.1	8.27e-10
halzen4	5.5	-	2786	2.76	0.613	4.65	5.82	$4.38 \cdot 10^{-9}$
	5.5	1	2193	2.52	0.603	4.49	2.65	$4.63 \cdot 10^{-9}$
halzen5	5.5	-	2786	4.76	0.560	4.69	105.0	$2.56 \cdot 10^{-9}$
	5.6	1	3518	5.13	0.618	4.88	0.469	$2.47 \cdot 10^{-9}$
halzen6	5.5	-	2786	6.83	0.530	4.70	0.969	$1.79 \cdot 10^{-9}$
	5.6	1	3518	7.30	0.576	4.76	91.8	$1.70 \cdot 10^{-9}$
halzen7	5.5	-	2786	2.46	0.664	4.85	5.11	$7.49 \cdot 10^{-9}$
	5.6	1	3518	2.71	0.727	4.87	96.3	$6.83 \cdot 10^{-9}$
halzen8	5.5	-	2786	3.89	0.609	4.81	4.26	$4.70 \cdot 10^{-9}$
	5.5	1	2193	3.57	0.598	4.67	3.75	$4.97 \cdot 10^{-9}$
halzen9	5.5	-	2786	5.22	0.572	4.70	0.776	$3.42 \cdot 10^{-9}$
	5.6	1	3518	5.64	0.636	4.81	6.49	$3.24 \cdot 10^{-9}$

Table C.2: Summary of the optimized rlogL values (with rdf cut) for different neutrino flux assumption for MGRO J2019+37 for a 5σ discovery.

flux	rlogL	rdf	N_{tot}	n_{exp} $\cdot 10^{-4}$	β_{med} (deg)	$n_{5\sigma}$	MDP $\cdot 10^{-8}$	$\Phi_{5\sigma}$ ($\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
kappes	5.5	-	2786	19.1	0.471	4.88	0.264	$2.30 \cdot 10^{-9}$
	5.4	1	1580	13.7	0.382	3.97	0.207	$2.61 \cdot 10^{-9}$
halzen1	5.5	-	2786	5.11	0.721	4.76	10.3	$2.14 \cdot 10^{-8}$
	5.6	1	3518	5.49	0.777	5.05	0.0379	$2.12 \cdot 10^{-8}$
halzen2	5.5	-	2786	7.17	0.662	4.85	0.0854	$1.56 \cdot 10^{-8}$
	5.6	1	3518	7.62	0.729	5.17	0.0701	$1.56 \cdot 10^{-8}$
halzen3	5.5	-	2786	8.79	0.631	5.05	0.0743	$1.32 \cdot 10^{-8}$
	5.6	1	3518	9.28	0.693	5.12	0.109	$1.27 \cdot 10^{-8}$
halzen4	5.4	-	1761	3.90	0.674	4.78	3.92	$4.04 \cdot 10^{-8}$
	5.6	1	3518	5.77	0.837	5.10	909.0	$2.92 \cdot 10^{-8}$
halzen5	5.5	-	2786	7.05	0.746	4.90	21.3	$2.29 \cdot 10^{-8}$
	5.6	1	3518	7.57	0.799	5.00	7.66	$2.18 \cdot 10^{-8}$
halzen6	5.5	-	2786	8.31	0.715	5.04	8.45	$2.00 \cdot 10^{-8}$
	5.0	1	403	0.728	0.361	2.60	2.26	$1.18 \cdot 10^{-7}$
halzen7	5.4	-	1761	4.23	0.753	5.08	8.46	$5.40 \cdot 10^{-8}$
	5.4	1	1580	3.45	0.731	4.72	916.0	$6.16 \cdot 10^{-8}$
halzen8	5.5	-	2786	7.41	0.822	5.00	15.0	$3.04 \cdot 10^{-8}$
	5.5	1	2193	5.99	0.789	4.93	6.02	$3.70 \cdot 10^{-8}$
halzen9	5.5	-	2786	8.46	0.799	5.16	8.56	$2.74 \cdot 10^{-8}$
	5.6	1	3518	9.13	0.849	5.03	9.30	$2.48 \cdot 10^{-8}$

Table C.3: Summary of the optimized rlogL values (with rdf cut) for different neutrino flux assumption for MGRO J2031+41 for a 5σ discovery.

source	rlogL	rdf	N_{tot}	n_{exp} $\cdot 10^{-3}$	β_{med} (deg)	$n_{5\sigma}$	MDP $\cdot 10^{-6}$	$\Phi_{5\sigma}$ ($\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
C1	5.5	-	2786	1.64	0.548	4.67	9.25	$1.82 \cdot 10^{-9}$
	5.6	1	3518	1.78	0.597	4.91	9.08	$1.77 \cdot 10^{-9}$
C2	5.5	-	2786	1.67	0.547	4.52	9.42	$1.33 \cdot 10^{-9}$
	5.6	1	3518	1.80	0.605	4.86	0.367	$1.32 \cdot 10^{-9}$

Table C.4: Summary of the optimized rlogL values (with rdf cut) for different neutrino flux assumption for C1 and C2 for a 5σ discovery for the kappes neutrino flux assumption.

C.2 Summary of the signal PDFs

The signal PDFs for the neutrino flux expectations and optimized cuts with the best MDP per source for optimization for 3σ is shown in Fig. C.1. A summary of the different parameters already shown in chapter 8 is given here again in Tab. C.5.

source	flux	rlogL	rdf	$n_{3\sigma}$	sensitivity ($\mathrm{TeV}^{-1}\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
MGRO J1908+06	halzen3	5.1	-	1.81	$6.3 \cdot 10^{-10}$
MGRO J2019+37	halzen3	5.2	1	0.93	$3.5 \cdot 10^{-10}$
MGRO J2031+41	kappes	5.3	-	1.07	$8.0 \cdot 10^{-10}$
C1	kappes	5.2	1	1.01	$8.7 \cdot 10^{-10}$
C2	kappes	5.2	1	0.94	$6.2 \cdot 10^{-10}$

Table C.5: Summary of the flux expectations and optimized cuts for the best 3σ discovery potential per source. The number of $n_{3\sigma}$ and the normalization of the sensitivity are also given.

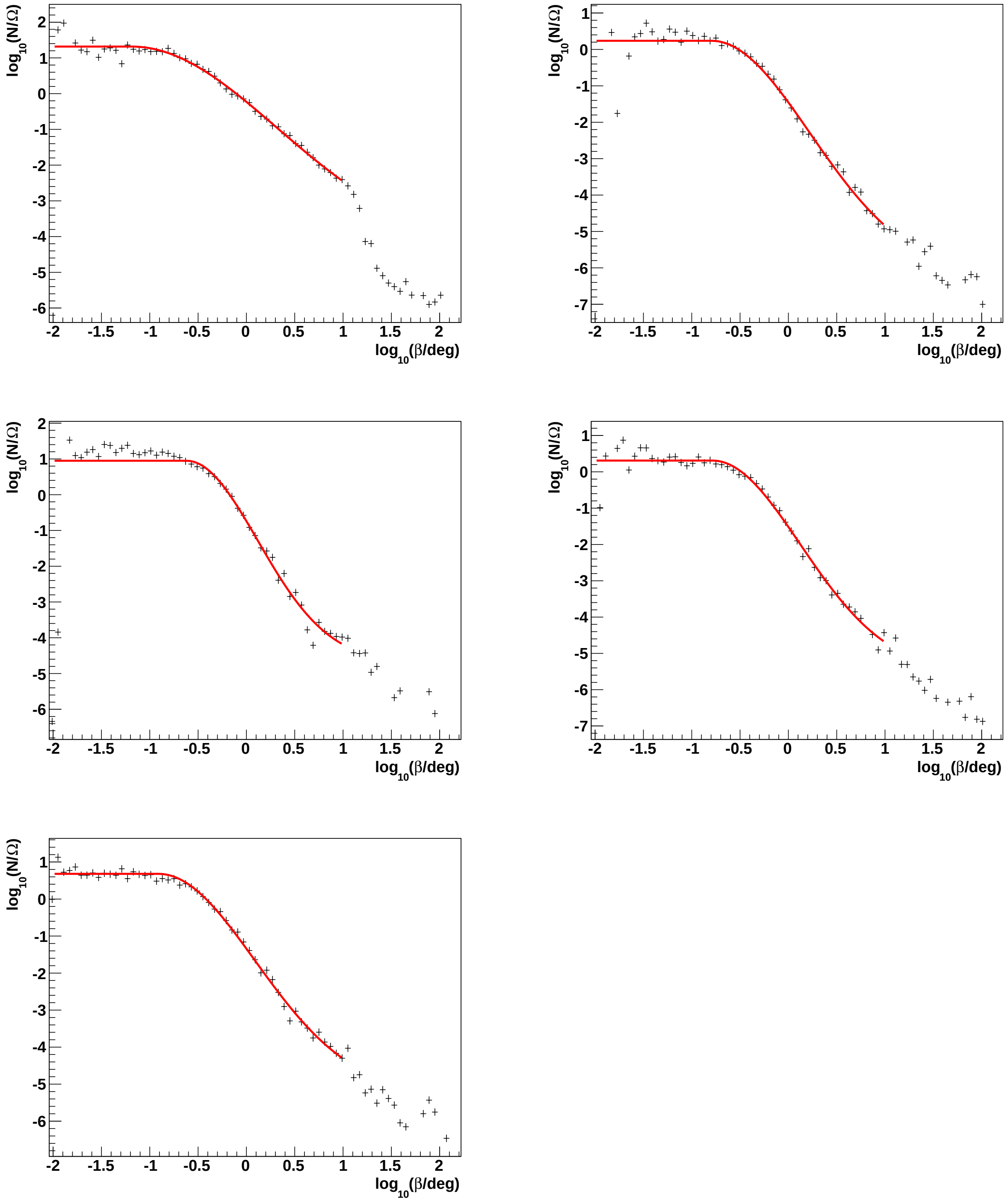


Figure C.1: Signal PDFs determined for the Milagro sources MGRO J1908+06, MGRO J2019+37 and MGRO J2031+41 (left: top to bottom) and C1 and C2 (right: top to bottom) for the neutrino flux expectations and cuts with the best MDP (see Tab. C.5 for the various conditions).

List of Figures

List of Tables

References

[1] M. G. Aartsen et al. Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector. *Science 342, 1242856*, 2013.

[2] A. A. Abdo et al. Discovery of TeV gamma-ray emission from the Cygnus region of the galaxy. *The Astrophysical Journal*, 658:L33–L36, 2007.

[3] A. A. Abdo et al. TeV gamma-ray sources from a survey of the Galactic Plane with Milagro. *The Astrophysical Journal*, 664:L91–94, 2007.

[4] F. Halzen, A. Kappes, A. Ó. Murchadha. Prospects for identifying the sources of the Galactic cosmic rays with IceCube. *Physical Review*, D78:063004, 2008.

[5] M.C. Gonzales-Garcia, F.Halzen, V. Niro. Reevaluation of the Prospect of Observing Neutrinos from Galactic Sources in the Light of Recent Results in Gamma Ray and Neutrinos Astronomy. *arXiv.1310.7194v1*, 2013.

[6] S. Adrián-Martínez et al. Searches for point-like and extended neutrino sources close to the Galactic Center using the ANTARES neutrino telescope. *The Astrophysical Journal Letters*, 786:L5 (5pp), 2014.

[7] M. G. Aartsen et al. Searches for Extended and Point-like Neutrino Sources with Four Years of IceCube Data. *arXiv:1406.6757*, 2014.

[8] J. Braun, J. Dumm, F. De Palma, C. Finley, A. Karle, T. Montaruli. Methods for point source analysis in high energy neutrino telescopes. *arXiv:0801.1604v1*, 2008.

[9] U. F. Katz, Ch. Spiering. High-Energy Neutrino Astrophysics: Status and Perspectives. *arXiv:1111.0507v1*, 2011.

[10] C. Grupen. *Astroparticle Physics.* Springer-Verlag Berlin Heidelberg, 2005.

[11] J. Blümer, R. Engel and J. R. Hörandel. Cosmic Rays from the Knee to the Highest Energies. *arXiv:0904.0725v1*, 2009.

[12] http://antares.in2p3.fr.

[13] http://icecube.wisc.edu.

[14] M.C. Gonzalez-Garcia, F. Halzen, S. Mohapatra. Identifying Galactic PeVatrins with neutrinos. *Astroparticle Physics*, 31:437–444, 2009.

[15] F.L. Villante, F. Vissani. How precisely neutrino emission from supernova remnants can be constrained by gamma ray observations? *Physical Review*, D78:103007, 2008.

[16] A. Kappes, J. Hinton, C. Stegmann, F.A. Aharonian. Potential Neutrino Signals from Galactic γ-Ray Sources. *The Astrophysical Journal*, 656:870–878, 2007.

[17] S. Adrián-Martínez et al. Search for muon neutrinos from gamma-ray bursts with the ANTARES neutrino telescope using 2008 to 2011 data. *A & A 559, A9*, 2013.

[18] R. Abbasi et al. An Absence of Neutrinos Associated with Cosmic-Ray Acceleration in Gamma-Ray Bursts. *Nature*, 484:351–354, 2012.

[19] Fritsch U. *Cluster search for neutrinos from active galactic nuclei.* PhD thesis, Friedrich-Alexander Universität Erlangen-Nürnberg, 2014.

[20] http://www-sk.icrr.u-tokyo.ac.jp/sk/index-e.html.

[21] http://www.sno.phy.queensu.ca.

[22] R. Bruijn. *The Antares Neutrino Telescope: Performance Studies and Analysis of First Data.* PhD thesis, University Amsterdam, 2008.

[23] Folger F. *Search for a diffuse cosmic neutrino flux using shower events in the ANTARES neutrino telescope.* PhD thesis, Friedrich-Alexander Universität Erlangen-Nürnberg, 2014.

[24] The ANTARES Collaboration. A Deep Sea Telescope for High Energy Neutrinos. Internal Note ANTARES-PHYS-1999-006, 1999.

[25] C. Grupen. *Teilchendetektoren.* BI-Wiss.-Verl. Mannheim, 1993.

[26] B. Hartmann. *Reconstruction of Neutrino-Induced Hadronic and Elektromagnetic Showers with the ANTARES Experiment.* PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2006.

[27] J.A. Aguilar et al. AMADEUS - The Acoustic Neutrino Detection Test System of the ANTARES Deep-Sea Neutrino Telescope. *Nuclear Inst. and Methods in Physics Research A*, 626-627:128–143, 2011. arXiv:1009.4179.

[28] J.A. Aguilar et al. ANTARES: the first undersea neutrino telescope. *Nuclear Inst. and Methods in Physics Research A*, 656:11–38, 2011. arXiv:1104.1607v1.

[29] J.A. Aguilar et al. Performance of the front-end electronics of the ANTARES Neutrino Telescope. *Nuclear Inst. and Methods in Physics Research A*, 622:59–73, 2010. arXiv:1007.2549.

[30] S. Adrián-Martínez et al. The positioning system of the ANTARES Neutrino Telescope. *JINST 7 T08002*, 2012.

[31] Garmt de Vries-Uiterweerd. *Signal and Background in the Underwater Neutrino Telescope ANTARES.* PhD thesis, University Utrecht, 2007.

[32] S. Wagner. Entwicklung und Analyse eines Hitselektionsalgorithmus für Niedrigenergieereignisse im ANTARES Neutrinoteleskop basierend auf einer Hough-Transformation. Diplomarbeit, Friedrich-Alexander Universität Erlangen-Nürnberg, 2010.

[33] J.A. Aguilar et al. A Fast Algorithm for Muon Track Reconstruction and its Application to the ANTARES Neutrino Telescope. *Astroparticle Physics*, 34:652–662, 2011.

[34] A. Heijboer. *Track reconstruction and point source searches with Antares.* PhD thesis, University Amsterdam, 2004.

[35] S. Adrián-Martínez et al. Search for Cosmic Neutrino Point Sources with Four Years of Data from the ANTARES Telescope. *JCAP 1303, 006*, 2013.

[36] C. Kopper. private communication.

[37] M. De Jong. Partial linearisation of the track fit problem. Internal Note ANTARES-SOFT-2007-001, 2007.

[38] M. C. Bouwhuis. *Detection of Neutrinos with Gamma-Ray Bursts.* PhD thesis, University Amsterdam, 2005.

[39] C. Kopper, D. Samtleben. Track reconstruction on a grid - Filteringfit. Internal Note ANTARES-SOFT-2012-006, 2012.

[40] T. Eberl, C. Kopper. The SeaTray software framework. Internal Note ANTARES-SOFT-2009-013, 2009.

[41] T. DeYoung et al. IceTray: A Software Framework for IceCube, International Conference on Computing in High-Energy Physics and Nuclear Physics (CHEP2004). available from http://www.chep2004.org/, 2004.

[42] M. De Jong. The Antares Trigger Parameters. Internal Note ANTARES-SOFT-2008-010, 2008.

[43] M. De Jong. The ANTARES Trigger Software. Internal Note ANTARES-SOFT-2005-005, 2005.

[44] S. Escoffier. Performance of the T3 Triggers on MC Data. Internal Note ANTARES-SOFT-2008-009, 2008.

[45] J. Hofestädt. Studies for low energy muon track reconstruction with the ANTARES detector. Erasmus project, Internal Note ANTARES-SOFT-2013-001, 2011.

[46] Y. Becherini, A. Margiotta, M. Sioli, M. Spurio. A parameterisation of single and multiple muons in the deep water or ice. *Astroparticle Physics*, 25:1–13, 2006.

[47] S. Geißelsöder. Classification of events for the ANTARES neutrino detector. Master's thesis, Friedrich-Alexander Universität Erlangen-Nürnberg, 2011.

[48] S. Geißelsöder. RDFClassify Tutorial. Internal Note ANTARES-SOFT-2012-002, 2012.

[49] S. Geißelsöder. private communication.

[50] S. Geißelsöder. PhD thesis in preparation. Friedrich-Alexander Universität Erlangen-Nürnberg.

[51] C. Bogazzi, JP. Gomez Gonzalez, A. Heijboer, J. Barrios Marti, T. Michael, S. Schulte, J. Zornoza. Point Source Search with 2007-2012 data. Internal Note ANTARES-PHYS-2013-006, 2013.

[52] R. Atkins et al. Status of the Milagro Gamma Ray Observatory. *arXiv:astro-ph/0110513v1*, 2001.

[53] R. Atkins et al. Observation of TeV gamma rays from the crab nebula with Milagro using a new background rejection technique. *The Astrophysical Journal*, 595:803–811, 2003.

[54] A. A. Abdo et al. Spectrum and Morphology of the two Brightest Milagro Sources in the Cygnus region: MGRO J2019+37 and MGRO J2031+41. *The Astrophysical Journal*, 753:159(8pp), 2012.

[55] A. A. Abdo et al. Milagro observation of multi-TeV emission from galactic sources in the Fermi bright source list. *The Astrophysical Journal*, 700:L127–L131(8pp), 2009.

[56] F. Aharonian et al. Detection of very high energy radiation from HESS J1908+063 confirms the Milagro unidentified source MGRo J1908+06. *A & A*, 499:723–728, 2009.

[57] http://tevcat.uchicago.edu/ at 2014-10-08.

[58] B. Bartoli et al. Observation of the TeV gamma-ray source MGRO J1908+06 with ARGO-YBJ. *arXiv1207.6280v2*, 2012.

[59] A. Kappes. private communication.

[60] R.J. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences.* John Wiley & Sons Ltd., West Sussex, England, 1999.

[61] C. Bogazzi, J. Hartmann, A. Heijboer. Point Source Search with 2007 and 2008 data. Internal Note ANTARES-PHYS-2010-008, 2010.

[62] http://mathworld.wolfram.com/GreatCircle.html.

[63] G.C. Hill, K. Rawlins. Unbiased cut selection for optimal upper limits in neutrino detectors: the model rejection potential technique. *arXiv:astro-ph/0209350v1*, 2002.

Acknowledgement

I would like to thank Uli Katz and Gisela Anton for the opportunity to do research in such an interesting field and for their support during the years.

Special thanks to Kay Graf for his supervision, his support and his encouragement in the final period of my thesis. Thanks a lot for your comments and suggestions and for proofreading this work. Further I want to thank Thomas Eberl for his support, especially in the development process of the track reconstruction strategy and for his help with computing and programming issues.

I want to thank Alexander Kappes for his interest in my studies and the discussions about point source analyses. His experience in this field helped a lot to develop this analysis procedure. In this context, thanks also to Julia Schmid and Kathrin Roensch especially for the discussions about the technical implementation of analyses methods.

I would like to thank the ANTARES collaboration and particularly the track reconstruction working group, with the leader Dorothea Samtleben, whose comments and suggestions helped to adjust KrakeFit.

Further I would like to thank the ANTARES and KM3NeT group of Erlangen, it was inspiring to work in such a collegial atmosphere. My special thanks go to my phd colleagues. Having you as companions makes this period of my life unforgettable. Guys, you know who this refers to. I thank my office mate Clancy James and my former office mate Friederike Schoeck for the nice atmosphere and valuable discussions (sometimes beyond physics). Particularly thanks to Clancy who let me win the discussion about closed and opened windows most of the time.

Finally I want to thank my family. Dankwart, thanks for your support and your encouragement through all the years and that you took the load off me in the last months. I would particularly like to thank my parents, without your support and help I would not have come so far.

Thank you! Danke!