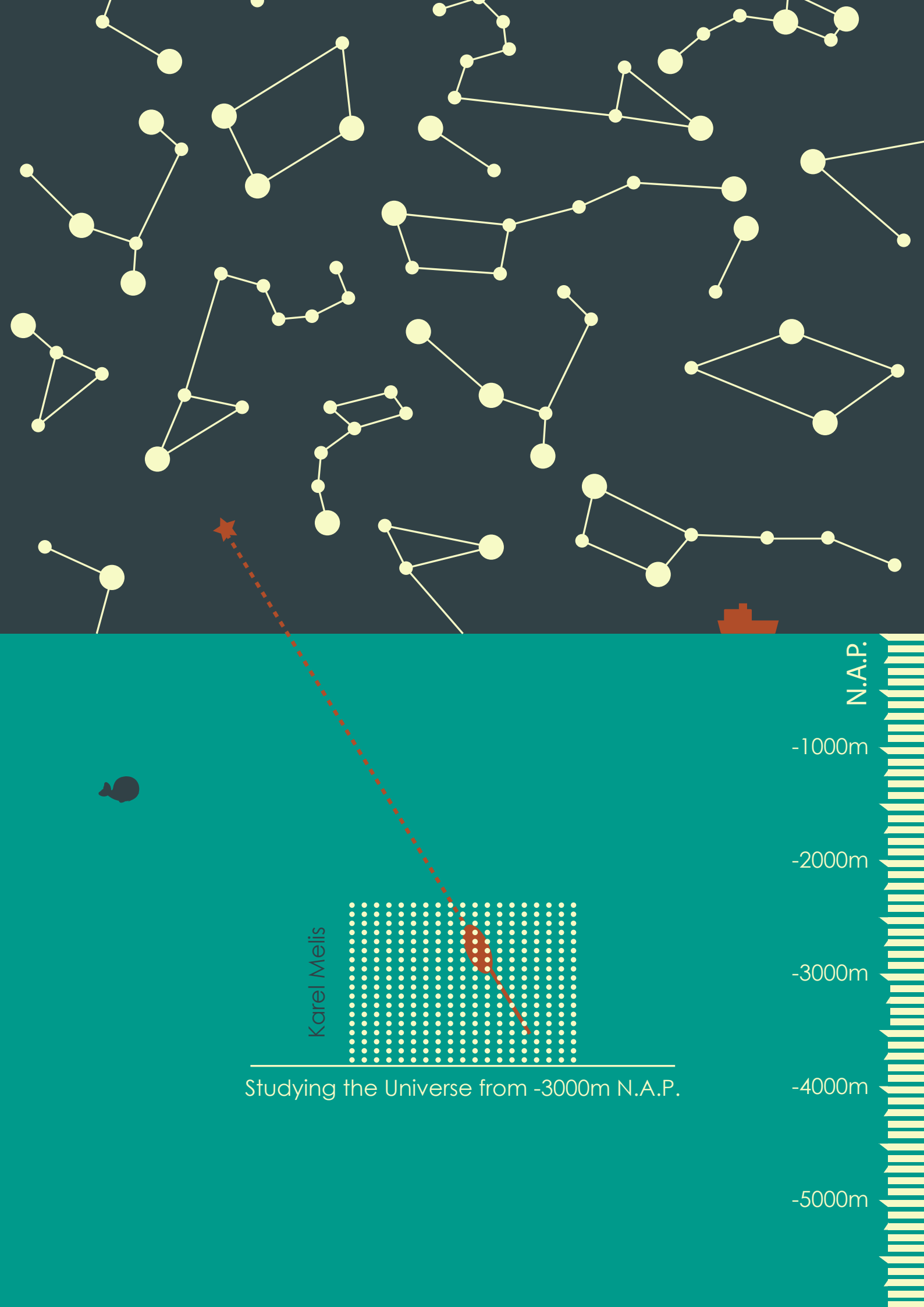
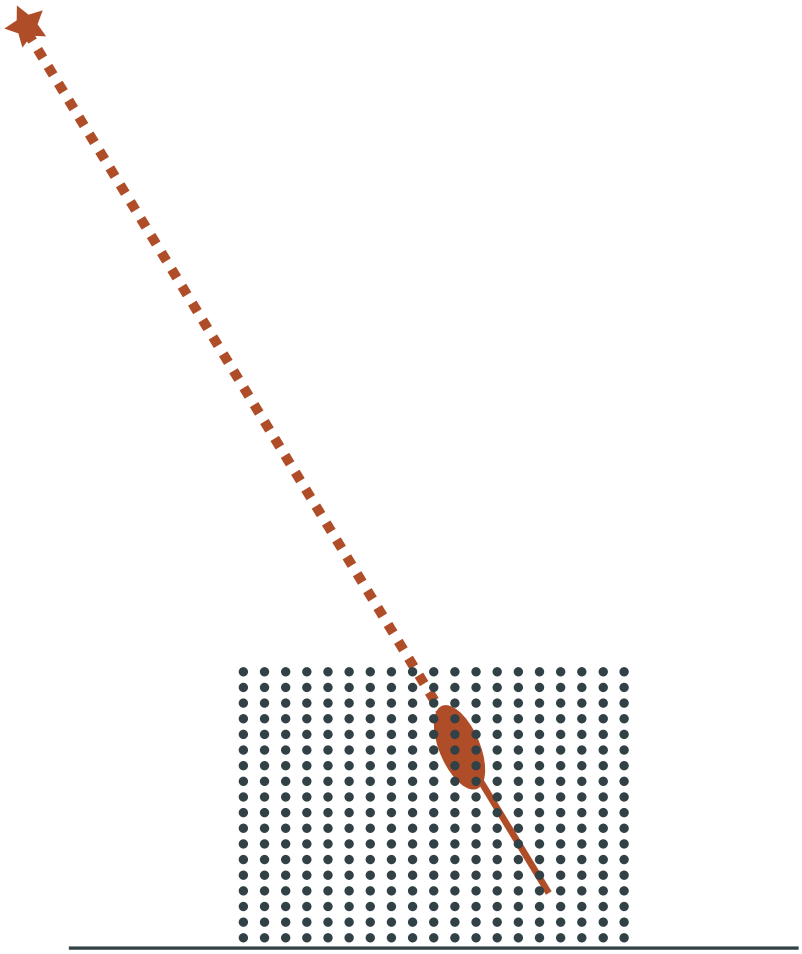

N.A.P.
-1000m
-2000m
-3000m
-4000m
-5000m
Karel Melis
Studying the Universe from -3000m N.A.P.



Studying the Universe from -3000m N.A.P.

The cover portrays the observation of a neutrino from a distant source in the Universe with an underwater neutrino telescope. Such an instrument is currently deployed by the KM3NeT collaboration in the Mediterranean Sea, and provides the data used for the studies presented in this thesis. The title refers to its approximate depth below the sea level; N.A.P. stands for 'Normaal Amsterdams Peil', a height datum formerly used as a reference point for measured depths. N.A.P. originates from a marked stone located in the docks of Amsterdam, The Netherlands[1, 2].

This work is part of the research programme of the Netherlands Organisation for Scientific Research (NWO). This research was carried out at the *Nationaal Instituut voor Subatomaire Fysica* (Nikhef) in Amsterdam, The Netherlands.

This thesis is available online at www.karel-melis.nl/PhDthesis.

mail@karel-melis.nl

ISBN 978-94-6421-506-9

Studying the Universe from -3000m N.A.P.

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 10 november 2021, te 11.00 uur

door Karel Willem Melis
geboren te Winterswijk

Promotiecommissie

Promotor:	prof. dr. P.M. Kooijman	Universiteit van Amsterdam
Copromotor:	dr. R. Bruijn	Universiteit van Amsterdam
Overige leden:	prof. dr. M. de Jong	Universiteit Leiden
	prof. dr. F.L. Linde	Universiteit van Amsterdam
	prof. dr. ing. B. van Eijk	Universiteit Twente
	prof. dr. A.P. Colijn	Universiteit van Amsterdam
	dr. H.L. Snoek	Universiteit van Amsterdam
	prof. dr. ir. P.J. de Jong	Universiteit van Amsterdam
	dr. C. Kopper	Michigan State University

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Als je iets doet, moet je het goed doen.

Contents

Samenvatting

Een pasgeborene ontdekt de wereld voornamelijk door middel van geur, smaak, tast en geluid. Dit omdat zijn/haar ogen nog geen kleuren kunnen onderscheiden en alles nog wazig zien. In de eerste maanden ontwikkelen de ogen zich volwaardig. Het duurt echter niet al te lang voordat hij/zij zich begint te interesseren in wat juist níet zichtbaar is; de inhoud van tassen en kastjes, de plaatjes verstopt achter flapjes in een flapjesboek en bovenal de gezichten die onder het uitroepen van 'kiekeboe' vanachter een paar handen tevoorschijn komen.

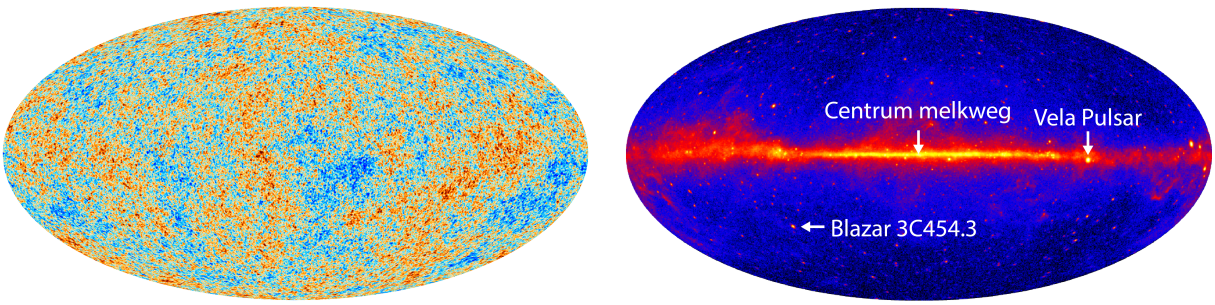
De nieuwsgierigheid van een kind neemt in de jaren daarna enkel toe, en leidt tot een eindeloze stroom van vragen. Naast essentiële vragen als "waarom moet ik mijn bord leeg eten?" zijn enkele veelgehoorde voorbeelden: "hoe groot is het heelal?", "wat is het allerkleinste dat er bestaat?" en "zijn er dingen waarvan we het bestaan niet weten?". Aan de hand van een aantal van dit soort vragen zal het onderwerp van deze dissertatie geïntroduceerd worden.

Hoe groot is het heelal? Om deze vraag te kunnen beantwoorden moet eerst de vraag "hoe oud is het heelal?" beantwoord worden. Metingen laten zien dat het heelal is ontstaan uit één enkel punt. Dit moment wordt de oerknal genoemd, en vond plaats ongeveer 13.8 miljard jaar geleden[3]. Sindsdien dijt het heelal alsmaar meer uit, waardoor alle objecten in het heelal steeds verder uit elkaar komen te liggen. Berekeningen laten zien dat het heelal tegenwoordig een bolvormig volume heeft met een straal van ongeveer 46.5 miljard lichtjaar, oftewel ongeveer $44 \cdot 10^{22}$ (44 met 22 nullen) kilometer. Dit wordt de grens van het *waarneembare heelal* genoemd.

Hoe ziet de rand van het heelal eruit? Hoewel de naam anders doet vermoeden is de rand van het waarneembare heelal niet waarneembaar. Vlak na de oerknal was de temperatuur van het heelal namelijk zo hoog dat er niets uit kon ontsnappen. Pas ongeveer 370 duizend jaar[4] na de oerknal was het heelal genoeg afgekoeld om elektronen en protonen samen waterstofkernen te laten vormen, waardoor lichtdeeltjes (fotonen) konden ontsnappen. Deze fotonen zijn nog steeds aanwezig, als laagenergetische *kosmische achtergrondstraling*, en markeren de grens van het *zichtbare heelal*.

Fotonen met zeer hoge energieën (genaamd gammastralen) bieden ons inzicht in hoogenergetische kosmische processen in het heelal. Gammastralen worden onder andere geproduceerd rond de centra van sterrenstelsels. Doordat gammastralen een wisselwerking aangaan ('interageren') met andere materie en fotonen kunnen de gammastralen van verre sterrenstelsels de aarde niet bereiken. Hierdoor is de fractie van het heelal die we met

gammastralen kunnen bestuderen slechts beperkt[5] (zie figuur 1).



Figuur 1: *De grenzen van het heelal wat we met licht van verschillende energieën kunnen zien.* ***Links****: De laagenergetische kosmische achtergrondstraling[6].* ***Rechts****: De flux van hoogenergetische gammastralen (rechten: NASA/DOE/Fermi/LAT Collaboration).*

Wat is het allerkleinste dat bestaat? Elk alledaags object kan opgesplitst worden tot atomen, welke weer uit elektronen en een atoomkern opgebouwd zijn. De atoomkern kan opgesplitst worden in protonen en neutronen, welke opgebouwd zijn uit quarks. Hier houdt het opsplitsen van deeltjes echter op.

Niet-opsplitsbare deeltjes worden *elementaire deeltjes* genoemd, en vormen de bouwstenen van het *standaardmodel van de deeltjesfysica*. Naast quarks bevat het standaardmodel vijf krachtdragende deeltjes (waaronder het foton) en zes *leptonen*. Elk quark en elk lepton heeft daarnaast zijn eigen antideeltje.

Er zijn drie *generaties* leptonen; het elektron (e^-) vormt samen met het elektron-neutrino (ν_e) een generatie, net als het muon (μ^-) met het muon-neutrino (ν_μ) en het tau (τ^-) met het tau-neutrino (ν_τ). De generatie van een neutrino wordt ook wel de *smaak* genoemd. De antideeltjes van de leptonen worden aangeduid met e^+, $\bar{\nu}_e$, μ^+, $\bar{\nu}_\mu$, τ^+ en $\bar{\nu}_\tau$.

Hoe weten we dat deze deeltjes bestaan? Tot 1932 was alleen het bestaan van elektronen[7], fotonen[8], protonen[9, 10] en neutronen[11] aangetoond. In 1933 publiceerde Anderson de ontdekking van een deeltje met precies dezelfde massa als het elektron, maar met tegenovergestelde lading; het positron (e^+)[12]. Hij ontdekte dit deeltje tijdens onderzoek naar hoogenergetische deeltjes uit het heelal, genaamd *kosmische stralen*.

Als een kosmische straal de aarde nadert, interageert het al snel met een atoom in de atmosfeer. Hierin ontstaan andere deeltjes, welke op hun beurt vervallen of interageren met andere deeltjes. Aldus ontstaat een zogeheten *kosmische deeltjeslawine*. Naast het positron is ook het muon ontdekt in kosmische deeltjeslawines, eveneens door Anderson[13].

Het merendeel van de deeltjes in een kosmische deeltjeslawine vervallen of verliezen hun energie voordat ze het aardoppervlak bereiken. Muonen en neutrinos zijn echter wel stabiel genoeg om het aardoppervlak te bereiken. Deze *atmosferische* muonen en neutrinos spelen een belangrijke rol in dit werk.

Tegenwoordig weten we dat kosmische stralen bestaan uit elektronen, protonen, heliumkernen en zwaardere atoomkernen. Een handvol kosmische stralen, (met bijnamen als

het 'oh-my-god'-deeltje) vormen de meest hoogenergetische deeltjes ooit gedetecteerd[14]. Deze deeltjes hebben een geschatte energie grofweg 100 miljoen keer zo hoog als de hoogstenergetische deeltjes geproduceerd door de mens in de LHC op het CERN in Genève, Zwitserland.

Waar komen extreem hoogenergetische kosmische stralen vandaan? Twee supernovae (de explosie waarmee een ster 'sterft') zijn geïndentificeerd als bronnen van hoogenergetische kosmische stralen[15]. Deze supernovae verklaren echter maar een zeer klein gedeelte van de totale gedetecteerde flux. Het is tot op heden onbekend wat de overige bronnen van extreem hoogenergetische kosmische straling zijn.

Aangenomen wordt dat extreem hoogenergetische kosmische stralen geproduceerd en versneld worden in de schokgolven die ontstaan bij het samensmelten van zogenaamde neutronensterren of in de *accretieschijf* rond zeer massieve zwarte gaten in het centrum van *actieve sterrenstelsels*[16]. Deze vermoedens zijn lastig te bewijzen, omdat extreem hoogenergetische kosmische stralen slechts incidenteel gedetecteerd worden, en omdat interstellaire magnetische velden ervoor zorgen dat de (elektrisch geladen) kosmische stralen geen rechte lijn volgen vanaf hun bron.

Motivatie

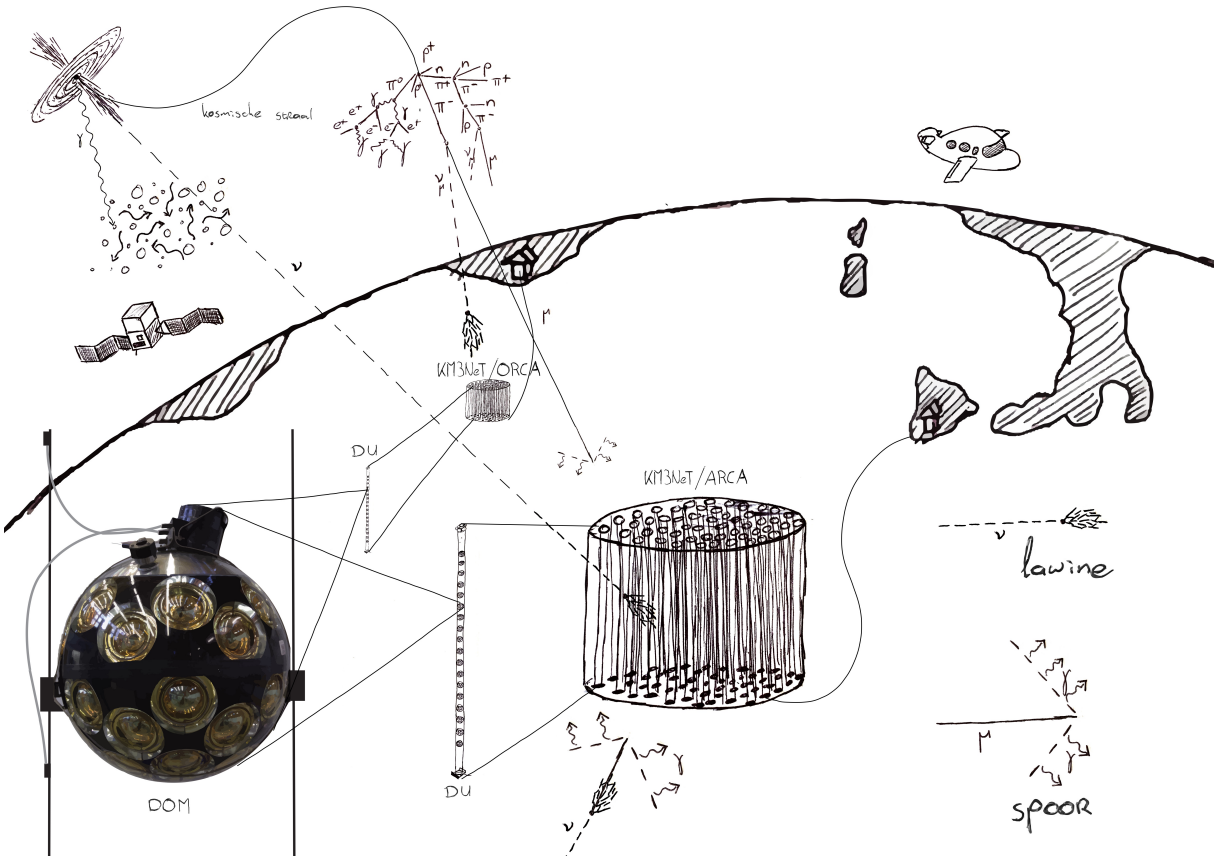
Naast fotonen en kosmische stralen is er een nog een derde deeltje dat onze aarde vanuit het heelal bereikt; het neutrino. Neutrinos zijn niet zichtbaar. Sterker nog, de meeste neutrinos vliegen dwars door onze ogen, ons lichaam en zelfs de hele aarde heen. Dit komt omdat neutrinos een bijzonder lage interactiewaarschijnlijkheid hebben. Om een idee te geven; van de 60 miljoen neutrinos met een energie van 1 TeV die door een kilometer zeewater vliegen interageert er gemiddeld slechts één.

Neutrinos zijn ideale 'boodschappers' om het heelal te bestuderen, en enkele van de hiervoor gestelde vragen te beantwoorden. Zij worden namelijk in vrijwel alle kosmische objecten geproduceerd, en door hun lage interactiewaarschijnlijkheid kunnen ze uit de meest compacte kosmische objecten ontsnappen. Daarnaast zorgt de lage interactiewaarschijnlijkheid er ook voor dat neutrinos vrijwel ongehinderd door het heelal heen kunnen vliegen. Tot slot zijn neutrinos elektrisch ongeladen, waardoor ze in een rechte lijn vliegen. Dit zorgt ervoor dat neutrinos direct terugwijzen naar hun bron.

Dit werk richt zich in het bijzonder op het bestuderen van hoogenergetische *kosmische neutrinos*, welke geproduceerd worden in interacties van kosmische stralen in de buurt van hun bron. Het energiespectrum van kosmische neutrinos is sterk gecorreleerd met het energiespectrum van de kosmische straling van waaruit ze geproduceerd worden. Hierdoor kunnen kosmische neutrinos gebruikt worden om de bronnen van extreem hoogenergetische kosmische stralen te identificeren en te bestuderen.

Neutrinos

De extreem lage interactiewaarschijnlijkheid van neutrinos maakt ze niet alleen de ideale deeltjes om het heelal mee te bestuderen, maar zorgt er ook voor dat ze notoir lastig te



Figuur 2: *De productie, propagatie en detectie van elementaire deeltjes uit het heelal.*

detecteren zijn. Hoe detecteer je immers een deeltje dat, in de meeste gevallen, zonder enige interactie dwars door de aarde heenvliegt?

Neutrinotelescopen In 1960 kwam Markov met het idee om een *neutrinotelescoop* te bouwen in een diep meer, de zee of het ijs op de noord- of zuidpool[17]. Van alle neutrinos die door het water/ijs heenvliegen zal er af en toe één interageren met een wateratoom. Mits de energie van het neutrino hoog genoeg is worden er in deze interactie een hele verzameling deeltjes geproduceerd, waarvan sommigen zorgen voor de emissie van licht. Dit licht heet Tsjerenkov straling, en is beter bekend als de blauwe gloed die in het koelwater rond kernreactoren zichtbaar is. Dit licht kan met een netwerk van gevoelige optische sensoren waargenomen worden. Uit het patroon van het gedetecteerde licht kunnen de richting, energie, smaak, interactiepositie en -tijd van het neutrino geschat, of 'gereconstrueerd' worden.

Om het licht afkomstig van neutrino interacties te kunnen detecteren moet een neutrinotelescoop op een donkere, transparante plaats gebouwd worden. Daarnaast is de telescoop bij voorkeur gelegen op grote diepte, ver van achtergronddeeltjes zoals atmosferische muonen. Enkele voorbeelden van neutrinotelescopen die dit detectieprincipe gebruiken zijn IceCube (diep in het Antarctische ijs), ANTARES en KM3NeT (beiden op grote diepte in de Middellandse Zee).

Afhankelijk van de smaak van het neutrino en het type interactie zijn er verschillende lichtpatronen, of *signaturen*, detecteerbaar. De meest voorkomende signaturen heten *lawines* en *sporen*. De Tsjerenkov straling van een lawine vormt ongeveer een uitdijende bol van licht. Een spoor zorgt voor een uitdijende kegel van licht, uitgezonden vanaf een lijn. Een spoor kan, afhankelijk van het muon dat het spoor veroorzaakt, kilometers lang zijn, en heeft één of meer lawines langs zijn pad.

Neutrino astronomie. Het bestuderen van kosmische objecten met behulp van neutrinos wordt *neutrino astronomie* genoemd. Vaak worden waarnemingen van neutrinos, fotonen, kosmische stralen en zwaartekrachtsgolven gecombineerd om tot nieuwe ontdekkingen te komen. Zo is supernova SN1987A middels een gecorreleerde waarneming van licht en neutrinos geïdentficeerd als neutrinobron[18–21], en is TXS 0506+056 (een actief sterrenstelsel met een zeer hoogenergetische 'jet' van materie gericht op de aarde) de aannemelijke bron van de gecorreleerde observatie van gammastralen en een hoogenergetisch neutrino[22–24]. Er bestaan echter nog twijfels of het gedetecteerde neutrino daadwerkelijk van TXS 0506+056 afkomstig is[25].

Icecube vergaart al meer dan tien jaar data, en heeft in die tijd een groot aantal hoogenergetische neutrinos gedetecteerd. Met statistische analyses op basis van de gereconstrueerde energieën is aangetoond dat deze neutrinos uit de kosmos afkomstig zijn [26–28]. Deze kosmische neutrino flux is tevens met ANTARES gemeten[29].

Omdat IceCube in het Antarctische ijs gebouwd is, waarin licht relatief veel verstrooit, kunnen de richtingen van deze neutrinos onvoldoende nauwkeurig bepaald worden om de bron(nen) van deze neutrinos te identificeren.

KM3NeT. Een neutrinotelescoop waarvan verwacht wordt dat de richtingen van kosmische neutrinos wel goed kan reconstrueren heet KM3NeT/ARCA. KM3NeT staat voor 'kubieke kilometer neutrinotelescoop' en ARCA voor 'Astroparticle Research with Cosmics in the Abyss'. KM3NeT/ARCA wordt gebouwd voor de kust van Sicilië (Italië), op een diepte van ongeveer drie kilometer.

Naast KM3NeT/ARCA bestaat KM3NeT uit KM3NeT/ORCA ('Oscillation Research with Cosmics in the Abyss'), welke gebouwd wordt voor de kust van Toulon (Frankrijk). KM3NeT/ORCA zal worden gebruikt om, middels een meting van de atmosferische neutrino flux, de eigenschappen van het neutrino te bestuderen. Beide detectoren worden tevens gebruikt voor oceanografie en tomografie van de aarde.

In figuur 2 wordt de opbouw van KM3NeT/ARCA en KM3NeT/ORCA schematisch weergegeven. Beide detectoren zijn opgebouwd uit zogeheten digitaal-optische modules (DOMs). De bolvormige glazen behuizing van elke DOM beschermt 31 optische sensoren genaamd fotomultiplicatoren ('photomultiplier tubes', of PMTs) en de bijbehorende elektronica tegen de druk van het water. 18 DOMs zijn bevestigd aan verticale lijnen ('detection units', of DUs), welke op de zeebodem verankerd zijn en door boeien overeind gehouden worden. 115 DUs worden over een cirkelvormig oppervlak verspreid om zo één 'bouwblok' te vormen.

KM3NeT/ARCA zal een volume van ongeveer één kubieke kilometer omvatten met twee bouwblokken. KM3NeT/ORCA zal bestaan uit één bouwblok en is, met een volume van ongeveer zes kubieke hectometer, veel compacter dan KM3NeT/ARCA. De detectoren

zijn via een netwerk van glasvezels en stroomkabels met de kust verbonden om elke DOM van stroom te voorzien en de data te verzamelen. Beide detectoren worden in fases gebouwd.

Naast licht afkomstig van neutrinos detecteren KM3NeT/ARCA en KM3NeT/ORCA ook licht afkomstig van bacteriën, algen en vissen (bioluminescentie), radioactieve vervallen van ^{40}K in zeezout en atmosferische muonen. Deze achtergronden worden benut in de kalibratie en validatie van KM3NeT.

Eerste data

Het eerste deel van de analyses gepresenteerd in deze dissertatie richt zich op de kalibratie van de detectoren met behulp van achtergrondlicht (de 'in-situ' kalibratie) en de analyse van de eerste data vergaard met twee KM3NeT/ARCA DUs en één KM3NeT/ORCA DU. Hieronder worden enkele resultaten samengevat.

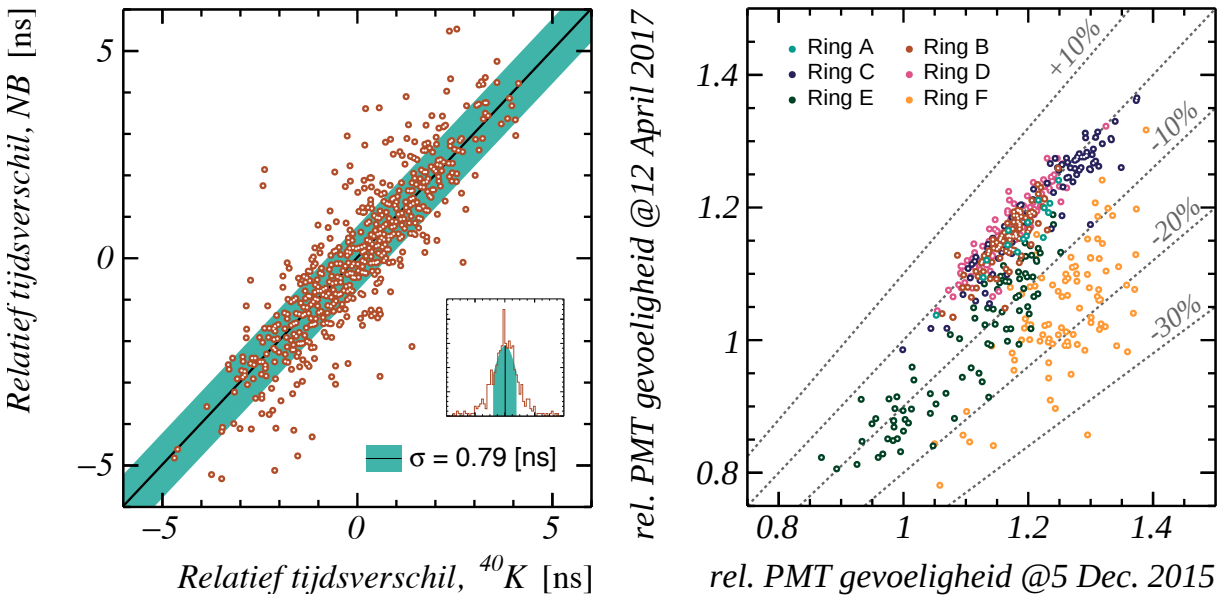
Kalibratie van KM3NeT. Een door een PMT gedetecteerd foton veroorzaakt een elektrisch signaal. Als dit signaal boven een bepaalde drempelwaarde komt wordt het moment waarop dit gebeurt en de tijd dat het signaal boven de drempel blijft (de time-over-threshold, of ToT) geregistreerd. Dit wordt een 'level-0 hit' (L0) genoemd.

De tijd dat het elektrisch signaal gedetecteerd wordt en de ToT worden geregistreerd met een nauwkeurigheid van een miljardste van een seconde (een nanoseconde). Toch is het niet evident om hieruit de aankomsttijd van het foton te herleiden. Elke PMT heeft namelijk een vertraging tussen het moment dat een foton de PMT raakt en het moment waarop het elektrisch signaal gedetecteerd wordt. Deze vertraging is ongeveer 30 nanoseconden, en verschilt per PMT. Daarnaast zijn de klokken (één per DOM) die gebruikt worden om de L0'en te registreren in fase, maar niet per sé gesynchroniseerd.

Het bepalen van de totale tijdsvertraging voor elke PMT wordt gedaan met verschillende kalibratietechnieken. In dit werk worden drie methodes beschreven om zowel KM3NeT/ORCA als KM3NeT/ARCA te kalibreren. Deze methodes maken gebruik van het licht afkomstig van ^{40}K vervallen en van atmosferische muonen. Naast het bepalen van de tijdskalibratie worden de gevoeligheid (de kans dat een foton gedetecteerd wordt) en de statistische spreiding van het tijdsverschil tussen de aankomsttijd van het foton en het gedetecteerde elektrische signaal van elke PMT bepaald.

De resultaten van de ontwikkelde kalibratietechnieken zijn vergeleken met andere kalibratie methodes, simulaties en overige waarnemingen. Een voorbeeld is gegeven in figuur 3 (links). Uit deze vergelijking blijkt dat de nauwkeurigheid van de tijdskalibratie beter is dan één nanoseconde, wat nodig is om de eigenschappen van neutrinos met goede nauwkeurigheid te kunnen bepalen.

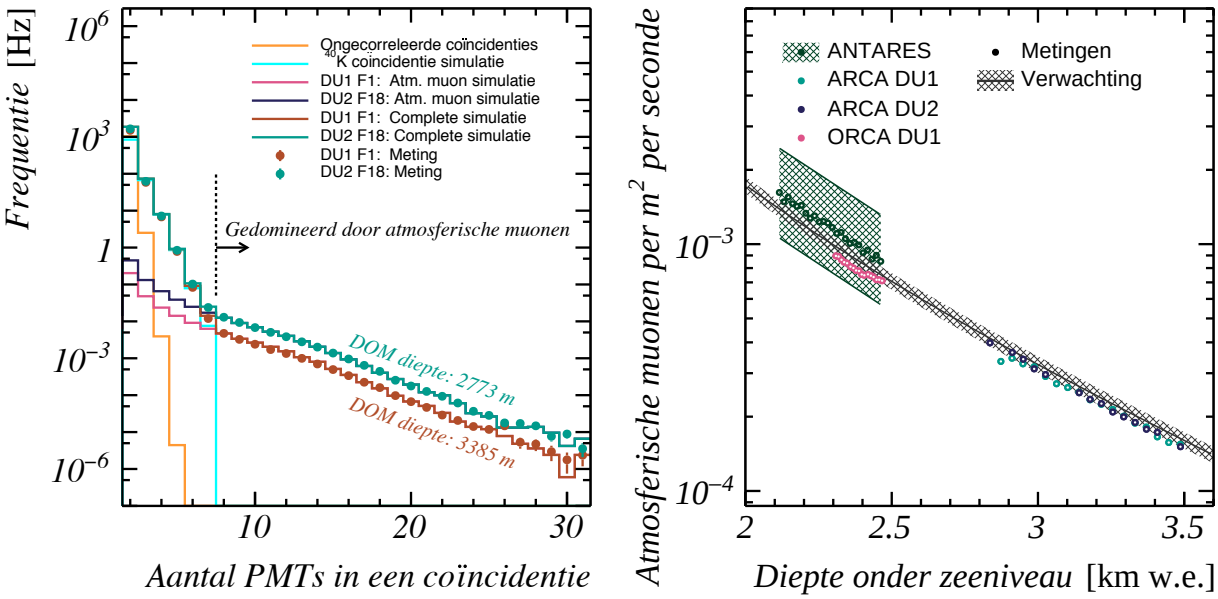
Door het gebruik van licht afkomstig van achtergronden is een grote hoeveelheid data beschikbaar, en zijn de eigenschappen van de detector continue te volgen. De nauwkeurigheid van de kalibratie van de PMT gevoeligheden is een paar procent, wat nauwkeurig genoeg is om de bevestiging van de DOMs aan de DUs en de effecten van sedimentatie op de bovenkant van de DOMs te onderscheiden (zie figuur 3, rechts). De eigenschappen van de PMTs zelf zijn zeer stabiel.



Figuur 3: *Vergelijkingen tussen de eigenschappen van de PMTs van KM3NeT/ARCA (2 DUs) bepaald met de in-situ kalibratie.* ***Links:*** *Een vergelijking van de relatieve tijdsverschillen tussen de 31 PMTs in een DOM, zoals bepaald uit de aankomsttijden van licht uitgezonden door LEDs ('nanobeacons', of NB) geplaatst in de DOMs[30, 31] en de kalibratie met behulp van ^{40}K vervallen. Beide methodes komen overeen met een standaardafwijking van* 0.79 *ns.* ***Rechts:*** *De ratio tussen de gevoeligheid van elke PMTs bepaald op 5 december 2015 en op 12 april 2017. Het effect van sedimentatie is zichtbaar op de PMTs die naar boven gericht zijn (ringen E en F). De gevoeligheid van de PMTs die omlaag kijken (ringen A-D) is zeer stabiel.*

Diepte-afhankelijkheid van de atmosferische muon flux. De kalibratie van de PMT gevoeligheden wordt gebruikt in gedetailleerde, realistische computersimulaties. Deze simulaties worden onder andere gebruikt om de kalibratie van de PMT gevoeligheden te verifiëren. Een voorbeeld van de dergelijke verificatie is gegeven in figuur 4 (links). In deze figuur wordt aangegeven hoe vaak meerdere PMTs ongeveer tegelijkertijd licht detecteren. Dit wordt een 'coïncidentie' genoemd. Een atmosferisch muon zendt veel meer licht uit dan een ^{40}K verval. Dit is te zien in de distributie van de coïncidenties; het licht afkomstig van een ^{40}K verval wordt vrijwel nooit door meer dan zeven PMTs gedetecteerd, terwijl coïncidenties van zeven of meer PMTs vrijwel uitsluitend veroorzaakt worden door atmosferische muonen.

Een muon verliest energie terwijl het door water beweegt. Hierdoor neemt de flux van atmosferisch muonen (het aantal atmosferische muonen per m^2 per seconde) af met de diepte onder het zeeoppervlak. De verwachte atmosferische muon flux is gegeven in figuur 4 (rechts). De atmosferische muon flux is gemeten met behulp van de ≥7-voudige coïncidentie frequentie op elke DOM van KM3NeT/ARCA en KM3NeT/ORCA. Twee verschillende simulaties van atmosferische muonen, één met nominale PMT gevoeligheden en één met realistische gevoeligheden (bepaald met de kalibratie besproken in de vorige paragraaf), zijn gebruikt om de ≥7-voudige coïncidentie frequentie om te rekenen naar de atmosferische muon flux. De metingen zijn weergegeven in figuur 4 (rechts), en komen



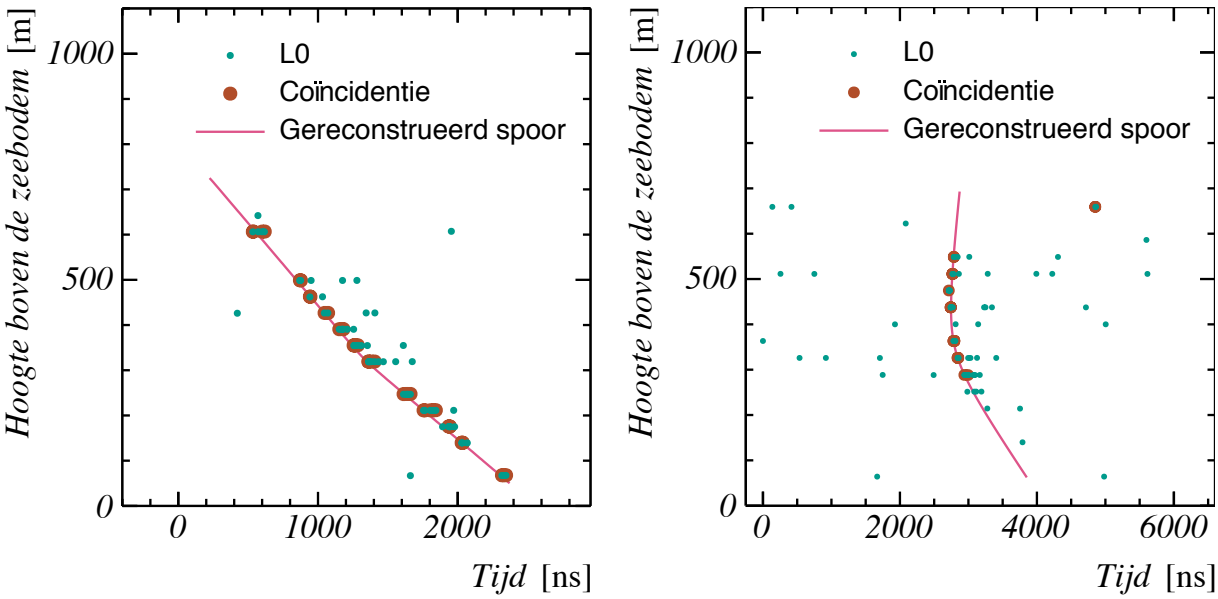
Figuur 4: ***Links:*** *De distributie van het aantal PMTs dat ongeveer tegelijkertijd een L0 detecteert. Metingen van twee DOMs op verschillende dieptes zijn vergeleken met de simulaties van ^{40}K vervallen en atmosferische muonen.* ***Rechts:*** *De richtingsgeïntegreerde flux van atmosferische muonen als functie van de diepte onder zeeniveau. De dichtheid van zout water (1.038 g/cm^3) is gebruikt om de equivalente diepte in zoet water uit te rekenen. De DOMs van KM3NeT/ORCA en KM3NeT/ARCA zijn gebruikt als individuele meetpunten. De metingen zijn vergeleken met de verwachte flux[32] en een onafhankelijke meting van de ANTARES detector[33].*

overeen met de verwachte flux[32] en met een onafhankelijke meting van ANTARES[33].

Reconstructie van sporen Muonen kunnen niet alleen geïdentificeerd worden aan de hand van de hoeveelheid licht gedetecteerd op een enkele DOM, maar ook uit het patroon van de aankomsttijden van licht gedetecteerd op verschillende DOMs. Als dit patroon overeenkomt met een spoor, lawine of andere signatuur wordt dit een *gebeurtenis* genoemd. Een voorbeeld van een gebeurtenis, gedetecteerd op 4 december 2015, is gegeven in figuur 5 (links). Dit is één van de allereerste gebeurtenissen welke met de eerste DU van KM3NeT/ARCA gedetecteerd is. In het figuur zijn de tijden van de gedetecteerde L0'en en ≥2-voudige coïncidenties weergegeven als functie van de hoogte van de DOM waarop de deze gedetecteerd is.

De DOM op een hoogte van 606 m detecteert het licht afkomstig van het muon ongeveer 1783 ns eerder dan de DOM op een hoogte van 68 m. Dit patroon komt overeen met een muon dat met bijna de lichtsnelheid (0.3 m/ns) ongeveer verticaal langs de DU naar beneden vliegt. Daarnaast zijn enkele L0'en veroorzaakt door licht afkomstig van ^{40}K vervallen.

Met behulp van een algoritme zijn de meest waarschijnlijke eigenschappen van het spoor gereconstrueerd. Dit gebeurde binnen 15 minuten na de vergaring van de eerste data. Het meest waarschijnlijke spoor beweegt met een hoek van 84 ° ten opzichte van de horizon naar beneden. Deze gebeurtenis is zeer waarschijnlijk veroorzaakt door een



Figuur 5: *De tijden waarop L0'en zijn gedetecteerd op verschillende DOMs van de eerste DU van KM3NeT/ARCA, met op de verticale as de hoogte van de DOM vanaf de zeebodem. Coïncidenties van twee of meer L0'en op één DOM zijn dikgedrukt. Het linker figuur geeft één van de eerst gedetecteerde gebeurtenissen weer, het rechter figuur een gebeurtenis welke mogelijk afkomstig is van een neutrino. De meest waarschijnlijke eigenschappen van het spoor zijn gereconstrueerd. De lijn geeft de verwachte aankomsttijd van het eerste detecteerbare licht afkomstig van deze sporen.*

atmosferisch muon.

Een voorbeeld van een gebeurtenis waarvan de meest waarschijnlijke richting omhoog is, is gegeven in figuur 5 (rechts). Atmosferische muonen kunnen weliswaar tot enkele kilometers doordringen in de aarde, maar er niet helemaal doorheen vliegen. De rechter gebeurtenis is dan ook mogelijk afkomstig van een neutrino (een muon-neutrino om precies te zijn), welke wél door de aarde heen kunnen vliegen. 100% zeker is dit echter niet. Statistische fluctuaties in de gedetecteerde L0 tijden kunnen er namelijk voor zorgen dat een atmosferisch muon spoor 'verkeerd gereconstrueerd' wordt, waarbij de gereconstrueerde richting sterk afwijkt van de werkelijke richting. Ook is het mogelijk dat dit licht niet afkomstig is van een spoor, maar van een hoogenergetische lawine langs een atmosferisch muon.

Kosmische neutrinobronnen.

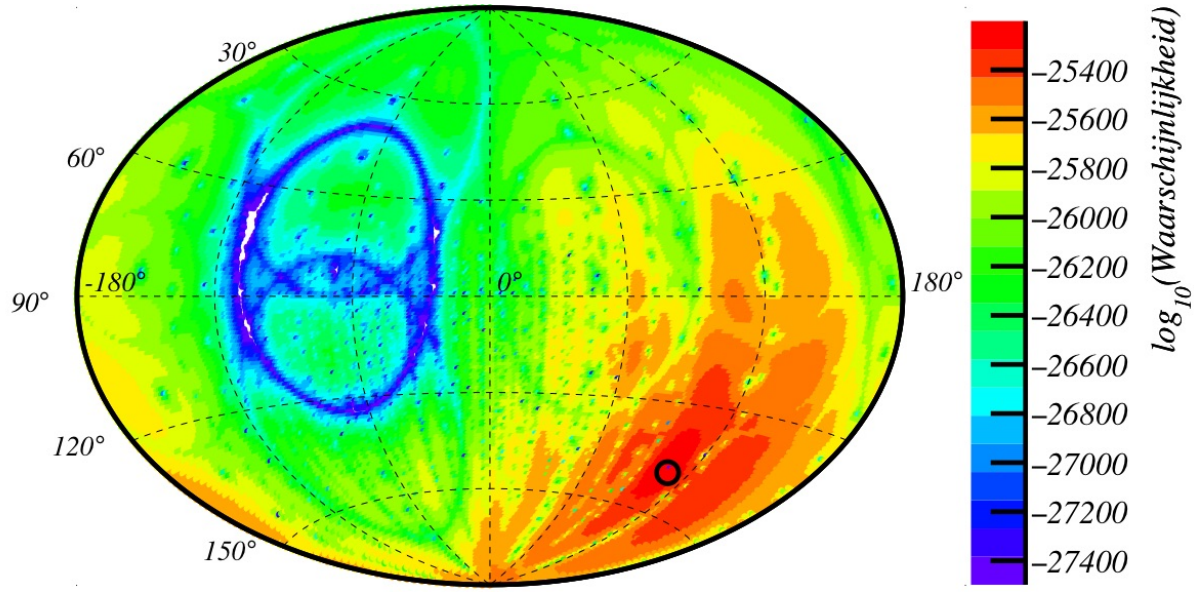
KM3NeT/ARCA wordt gebouwd om kosmische neutrinobronnen te identificeren en bestuderen. Het overgrote merendeel van de gedetecteerde gebeurtenissen zijn echter afkomstig van atmosferische neutrinos en atmosferische muonen. Een nieuwe statistische methode is ontwikkeld om het zwakke signaal van één of meerdere kosmische neutrinobron(nen) te onderscheiden van deze achtergronden.

De meest belangrijke eigenschap van de nieuwe methode is het *onderscheidend ver-*

mogen. Hoe beter het onderscheidend vermogen, des te minder data nodig is om een kosmische neutrinobron te identificeren, en des te nauwkeuriger de eigenschappen van deze bron(nen) te bepalen zijn. De nieuwe methode is ontworpen om, voor elke mogelijke bron, het theoretisch best mogelijke onderscheidend vermogen dichter te benaderen dan met bestaande methodes. Dit wordt bereikt door zoveel mogelijk informatie over de bron, de detector en de data te gebruiken.

Waarschijnlijkheidslandschappen. In de nieuwe methode worden zogeheten *waarschijnlijkheidslandschappen* gebruikt om zoveel mogelijk informatie uit elke gedetecteerde gebeurtenis te halen. Een waarschijnlijkheidslandschap beschrijft de waarschijnlijkheid $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)$ dat een gebeurtenis $\mathcal{E}$ afkomstig is van een neutrino uit een bepaalde richting, welke interageert op een bepaalde positie en tijd, en in deze interactie een signatuur veroorzaakt bestaande uit een enkel spoor, een enkele lawine of een lawine waaruit een spoor komt. Met deze parametrisatie is het waarschijnlijkheidslandschap van een gebeurtenis afhankelijk van acht variabelen $\boldsymbol{x}_s$ (twee voor de richting, drie voor de positie, één voor de tijd en twee energiën). Deze parametrisatie kan ook gebruikt worden om atmosferische muonen te beschrijven, namelijk een spoor zonder lawine afkomstig van een positie ver boven de detector.

De 'hoogte' van een waarschijnlijkheidslandschap $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)$ in een bepaald punt $\boldsymbol{x}_s$ hangt af van de kansen dat elke PMT wel of geen L0(en) detecteert in een bepaald tijdsvenster. Van de geraakte PMTs wordt daarnaast de kans gebruikt dat de L0 juist op dat moment (en dus niet eerder of later) gedetecteerd wordt. Een dwarsdoorsnede van het waarschijnlijkheidslandschap van een gebeurtenis is gegeven in figuur 6.



Figuur 6: *Een dwarsdoorsnede van het waarschijnlijkheidslandschap van een gesimuleerde gebeurtenis. De richting van de signatuur is gevarieerd, de overige grootheden zijn vastgezet op de waardes van de gesimuleerde gebeurtenis. De zwarte cirkel geeft de gesimuleerde richting aan.*

Verwachte signatuur flux De nieuwe methode kan gebruikt worden om het bestaan van elk mogelijke (verzameling van) kosmische bron(nen) te toetsen (een $\mathcal{H}_1$ hypothese). Dit gebeurt door uit te rekenen wat de flux $\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}_1)$ van het aantal verwachte signaturen per eenheid richting, volume, tijd en energieën van $\mathcal{H}_1$ is. Naast signaturen veroorzaakt door kosmische neutrinos bestaat de $\mathcal{H}_1$ flux ook uit de verwachte flux van achtergronden (atmosferische neutrinos en atmosferische muonen). De hypothese met enkel achtergronden wordt de nulhypothese ($\mathcal{H}_0$) genoemd.

Om de verwachte signatuur flux van de $\mathcal{H}_0$ of een $\mathcal{H}_1$ hypothese uit te rekenen zijn parametrisaties gebruikt van de verwachte neutrinoflux van kosmische neutrinos en achtergronden, de propagatie van neutrinos door de aarde, de interactiewaarschijnlijkheid van neutrinos in en rond de detector en de kansverdeling dat een interagerend neutrino met een bepaalde smaak en energie leidt tot een signatuur met parameters $\boldsymbol{x}_s$.

Detectiegevoeligheid Een derde ingrediënt van de nieuwe methode is een functie die de gevoeligheid van de detector beschrijft; de kans $P^{det}(\boldsymbol{x}_s)$ dat een signatuur met parameters $\boldsymbol{x}_s$ gedetecteerd wordt. Simulaties zijn gebruikt om een parametrisatie van de detectiegevoeligheid te verkrijgen.

Toetsingsgrootheid De drie ingrediënten zijn gecombineerd in een (acht-dimensionale) integraal over alle mogelijke signaturen $\boldsymbol{x}_s$:

$$\mathcal{L}(\mathcal{E}|\mathcal{H}) \equiv \int \mathcal{L}(\mathcal{E}_i|\boldsymbol{x}_s) \cdot P^{det}(\boldsymbol{x}_s) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) d\boldsymbol{x}_s. \tag{1}$$

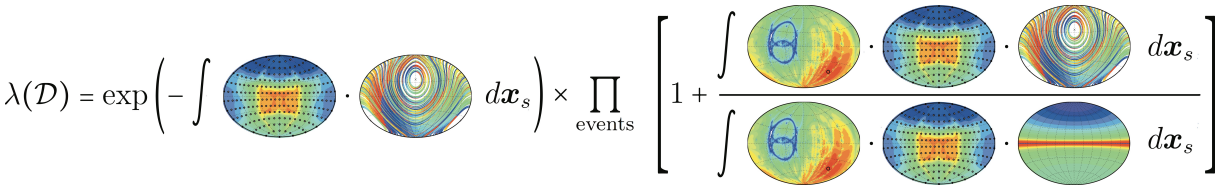
De oplossing van deze integraal geeft de kans dat een gebeurtenis $\mathcal{E}$ veroorzaakt is door een neutrino of muon afkomstig van de flux onder een hypothese $\mathcal{H}$. Als een gebeurtenis bijvoorbeeld zeer waarschijnlijk veroorzaakt is door een neutrino uit een bepaalde richting, dan is $\mathcal{L}(\mathcal{E}|\ \mathcal{H})$ zeer hoog voor hypotheses met een bron in deze richting. Omdat (bijna) alle informatie van de gebeurtenis, detector en verwachte flux in deze integraal gebruikt zijn, is de uitgerekende kans $\mathcal{L}(\mathcal{E}|\mathcal{H})$ zeer nauwkeurig.

De kansen $\mathcal{L}(\mathcal{E}|\mathcal{H})$ van alle gedetecteerde gebeurtenissen (de data $\mathcal{D}$) zijn gebruikt om uit te rekenen hoe veel waarschijnlijker het is dat de data afkomstig is van een $\mathcal{H}_1$ hypothese dan van de $\mathcal{H}_0$ hypothese:

$$\lambda(\mathcal{D}|\mathcal{H}_1) = \frac{\mathcal{L}(\mathcal{D}|\mathcal{H}_1)}{\mathcal{L}(\mathcal{D}|\mathcal{H}_0)} = \exp\left(-\mu^{det}_{\mathcal{H}_1-\mathcal{H}_0}\right) \times \prod_{\text{gebeurtenissen}} \left[\frac{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)}{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)}\right], \tag{2}$$

waarin $\mathcal{L}(\mathcal{D}|\mathcal{H})$ de kans is om de data te observeren als hypothese $\mathcal{H}$ waar is en $\mu^{det}_{\mathcal{H}_1-\mathcal{H}_0}$ het verwachte aantal gedetecteerde gebeurtenissen afkomstig van de kosmische neutrinobron is. In figuur 7 is deze formule grafisch weergegeven.

Deze kansratio $\lambda(\mathcal{D}|\mathcal{H}_1)$, genaamd de *toetsingsgrootheid*, is groot indien er één gebeurtenis vrijwel zeker afkomstig is van $\mathcal{H}_1$, ofwel als er veel gebeurtenissen zijn met een kleine kans van $\mathcal{H}_1$ afkomstig te zijn. Als de toetsingsgrootheid groter is dan een bepaalde waarde λ_{kritiek} kan de $\mathcal{H}_1$ hypothese geaccepteerd worden, en is het bestaan van de kosmische neutrinobron die correspondeert met $\mathcal{H}_1$ aangetoond. De waarde van λ_{kritiek} hang af van de $\mathcal{H}_1$ hypothese die getoetst wordt, en wordt zó gekozen dat de kans dat er een



Figuur 7: *Een grafische weergave van de toetsingsgrootheid $\lambda(\mathcal{D})$. Naast de waarschijnlijkheidslandschappen (versimpeld weergegeven met) hangt deze integraal af van de detectie gevoeligheid ( ) en de verwachte signatuur flux van achtergronden ( ) en van kosmische neutrinobronnen ( ). In werkelijkheid zijn deze functies acht-dimensionaal.*

foutieve ontdekking gedaan wordt ($\mathcal{H}_1$ wordt geaccepteerd terwijl $\mathcal{H}_0$ waar is, een fout van de eerste orde) kleiner is dan 0.000057%.

Voor- en nadelen De meest significante voordelen van de nieuwe methode om neutrinobronnen te identificeren ten opzichte van bestaande methodes zijn:

- De nieuwe methode is zeer flexibel, in de zin dat elk van de ingrediënten in de toetsingsgrootheid integraal eenvoudig inwisselbaar is met een andere/betere functie:
  - Door enkel de gebruikte flux te veranderen kan het bestaan van elke mogelijke kosmische neutrinobron getoetst worden.
  - Data vergaard met een verschillend aantal DUs kan gecombineerd worden door de detectie efficïentie te vervangen. Dit maakt het mogelijk om ook gebeurtenissen te gebruiken terwijl KM3NeT nog onder constructie is.
  - Een nieuw ontwikkelde suggestie om ook het aantal gedetecteerde L0'en en de ToT informatie van de L0'en te gebruiken in de waarschijnlijkheidslandschappen wordt bediscussieerd in appendix I. Deze suggestie kan gebruikt worden door de waarschijnlijkheidslandschappen te vervangen.
- De nieuwe methode gebruikt meer informatie dan conventionele methodes, waarvan bewezen is dat het leidt tot een beter oplossend vermogen.
  - In conventionele methodes worden alleen de gereconstrueerde, i.e. de meest waarschijnlijke, eigenschappen van het neutrino gebruikt. In de nieuwe methode worden alle mogelijke oplossingen gebruikt.
  - De onnauwkeurigheid van de bepaalde neutrino eigenschappen wordt in conventionele methodes benaderd met relatief simpele parametrisaties. In de nieuwe methode worden de onnauwkeurigheden op de meest nauwkeurig mogelijke manier beschreven met behulp van het waarschijnlijkheidslandschap.
- In conventionele methodes worden selectiecriteria gebruikt om slecht en verkeerd gereconstrueerde gebeurtenissen (en gebeurtenissen die slecht door de reconstructie beschreven zijn) buiten de analyse te houden. In de nieuwe methode bestaan er

geen verkeerd gereconstrueerde gebeurtenissen, omdat het waarschijnlijkheidslandschap elke mogelijke oplossing beschrijft. Hierdoor zijn selectiecriteria overbodig, waardoor alle gebeurtenissen kunnen bijdragen aan het identificeren van (een) neutrinobron(nen).

- In conventionele methodes worden lawines en sporen verschillend behandeld. Een 'smaakidentificatie' stap wordt gebruikt om de meest waarschijnlijke signatuur (en dus smaak van het neutrino) te bepalen. Deze stappen zijn niet meer nodig in de nieuwe methode, omdat zowel lawines, sporen en een combinatie van een spoor en een lawine beschreven worden. Door deze verbeterde parametrisatie kan er meer informatie uit elke gebeurtenis gehaald worden. Daarnaast is het relatief eenvoudig om de nieuwe methode uit te bereiden naar meer exotische signaturen.

De nieuwe methode heeft ook een nadeel; het kost veel computerkracht om elk van de functies in de integraal, alsmede de integraal zelf, uit te rekenen. Dit nadeel was reeds voorzien in een vroeg stadium van het ontwerp, en is doelbewust geaccepteerd om te kijken in hoeverre de nieuwe methode toepasbaar is. Het gevolg hiervan is echter dat de integraal $\mathcal{L}(\mathcal{E}|\mathcal{H})$ is uitgerekend voor 60 gesimuleerde gebeurtenissen en naast de $\mathcal{H}_1$hypothese voor 121 verschillende $\mathcal{H}_1$ hypotheses (elk met een diffuse kosmische neutrino flux). Met het uitreken van deze integraal is aangetoond dat de nieuwe methode in principe werkt (het uitrekenen van de toetsingsgrootheid $\lambda(\mathcal{D}|\mathcal{H}_1)$ en de minimale toetsingsgrootheid om $\mathcal{H}_1$ te accepteren is een standaard statistische handeling). Zestig gebeurtenissen zijn echter niet voldoende om ook daadwerkelijk aan te tonen dat de nieuwe methode een beter oplossend vermogen heeft dan conventionele methodes.

Conclusies.

Technieken zijn ontwikkeld om de KM3NeT detectoren te kalibreren met behulp van het licht afkomstig van ^{40}K vervallen en atmosferische muonen. De tijdskalibratie is binnen één nanoseconde nauwkeurig, wat nauwkeurig genoeg is om gebeurtenissen goed te kunnen reconstrueren. De kalibratie van de PMT gevoeligheden is precies genoeg om kleine verschillen tussen de PMTs waar te kunnen nemen. De gemeten PMT gevoeligheden worden onder anderen gebruikt in een meting van de diepte-afhankelijkheid van de atmosferische muon flux met KM3NeT/ORCA en KM3NeT/ARCA.

In het laatste deel van deze dissertatie wordt een nieuwe methode geïntroduceerd om kosmische neutrinobronnen te identificeren en bestuderen. Deze methode is ontworpen om de theoretisch best mogelijke gevoeligheid dichter dan ooit te naderen. Dit gebeurt door middel van waarschijnlijkheidslandschappen, gedetailleerde parametrisaties van de flux van verwachte gebeurtenissen en een nieuwe manier om signaturen te beschrijven. Bijkomend voordeel van deze aanpak is dat de nieuwe methode zeer flexibel is; alle mogelijke kosmische bronnen kunnen getoetst worden, hiervoor kan alle data gebruikt worden en complexere signaturen dan lawines en sporen zijn redelijk eenvoudig te implementeren.

Deze methode kost echter veel computerkracht. Om tegemoet te komen aan limitaties in de beschikbare computerkracht kan een hybride methode aanbevolen worden. In deze hybride methode wordt de conventionele methode gebruikt voor laagenergetische

'standaard' gebeurtenissen (goed beschreven met conventionele reconstructiemethodes), en wordt nieuwe methode toegepast op gebeurtenissen die 1) anders niet meegenomen zouden worden in de analyse (slecht gereconstrueerde gebeurtenissen) 2) de meest hoog-energetische gebeurtenissen (bijvoorbeeld de gebeurtenissen met zeer veel L0s) en 3) gebeurtenissen met een gereconstrueerde richting dicht bij een bekend kosmisch object.

KM3NeT probeert het onzichtbare heelal zichtbaar te maken vanaf een plek waar bijna niets te zien is. Dit proces voltrekt zich met kleine stapjes, als een pasgeborene die begint met zien. Met de constructie van de eerste detectie elementen zijn de 'ogen' van de detector geopend. Met het kalibreren ziet de detector scherp. De identificatie van de eerste neutrinos tussen de atmosferische muon achtergrond kan vergeleken worden met het onderscheiden van de eerste kleuren. De nieuwe methode om kosmische neutrinobronnen te identificeren is een eerste stapje richting volledig zicht op het heelal. Wie weet wat zich dan openbaart.

Summary

It takes a few months before the eyes of a newborn are fully developped and he/she can see all colours clearly. However, it doesn't take too long after this moment before he/she is mostly fascinated by the invisible world: The contents of bags and drawers, the hidden pictures in a lift-the-flap book and the faces emerging from behind a pair of hands saying 'peek-a-boo'.

Over the years, the curiosity of a child only increases, leading to an endless stream of questions such as "why do I have to finish my plate?". For some people, called scientists, such questions evolve to "what is the smallest/biggest thing that exists?" and "are there things we don't know exist?". The topics of this dissertation will be introduced by answering a few of those questions.

What are the smallest things? Every object can be split into atoms, which in turn consist of a nucleus and one or more surrounding electrons. The nucleus can be split into protons and neutrons, which in turn are made out of quarks. It is currently unknown if even smaller particles exist.

Indivisible particles are called *elementary particles*. These particles form the building blocks of the *Standard Model* of particle physics. The Standard Model consists of six types of quarks, six leptons and five force-carrying particles. Quarks make up protons and neutrons in the nucleus of atoms, while the electrons surrounding the nucleus are leptons. Force-carrying particles called photons scatter of these particles and may end up in our eyes, allowing us to see the world around us.

What is the biggest thing? The biggest thing that exists is (by definition) the universe. It emerged about 13.8 billion years ago[3] from a single point, in an explosion-like event called the *Big Bang*. At present, everything that we can possibly observe is contained in a spherical volume with a radius* of about $44 \cdot 10^{22}$ kilometers (46.5 billion light-years), called the *observable universe*. Light emitted from points outside this volume should have been emitted before the big bang in order to reach the Earth, which cannot be the case.

What does the edge of the observable universe look like? Photons from the edge of the observable universe were emitted right after the Big Bang. However, during the first 370 thousand years[4] after the Big Bang, the universe was so dense that all photons interacted almost immediately with another particles, making the early universe completely opaque.

*taking into account that the universe itself is expanding over time

The edge of the observable universe is therefore, in contrast to what its name suggests, actually not observable.

Which fraction of the universe can we see? The first photons that 'escaped' from the early universe can still be observed, forming the low-energy *cosmic microwave background radiation* (CMB). The CMB marks the boundary of the *visible universe* (see figure 8). The probability for a photon to interact with matter and other photons increases with the energy of the photon, resulting in a smaller fraction of the universe to be visible at higher photon energies. With high-energy photons (called *gamma-rays*), only the nearest galaxies can be observed.

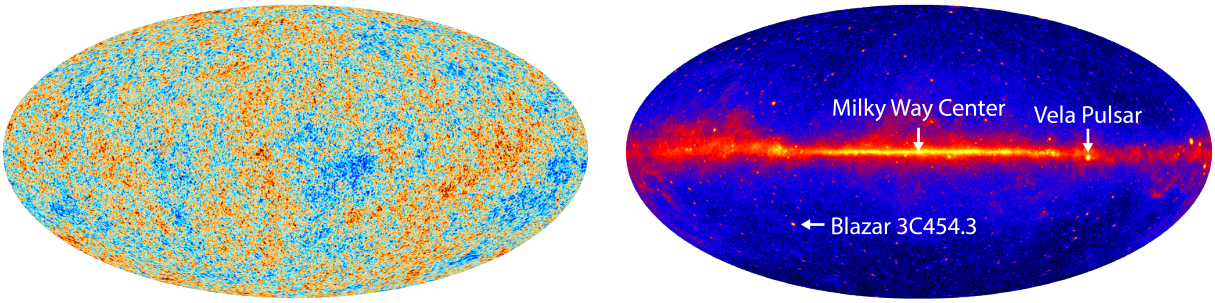


Figure 8: *The edge of the visible universe, as observed with photons of different energies.* ***Left****: The low-energy cosmic microwave background radiation[6].* ***Right****: The flux of high-energy gamma-rays (rights: NASA/DOE/Fermi/LAT Collaboration).*

Are there no other ways to study the universe? In 1922, Victor Hess showed that electrically charged particles penetrate the atmosphere from the universe. He named these particles *cosmic rays*. They consist of electrons, protons, helium nuclei and heavier atomic nuclei. Cosmic rays interact high in the atmosphere, causing the production of secondary particles. These particles also interact, producing a 'shower' of secondary particles. Most of the particles in the shower decay before reaching the surface of the Earth. Long-living particles called muons (the more massive 'brother' of the electron) can reach the surface, and may even penetrate the Earth for a few kilometres. Such muons are called *atmospheric muons*.

A handful of cosmic rays (with nicknames such as the 'oh-my-god particle') are the highest-energy particles ever detected[14], with energies roughly 100 million times higher than the highest-energy particles produced by humans (in the LHC at CERN in Switzerland). A few cosmic objects have been proposed as potential sources of ultra-high-energy cosmic rays. However, none of these sources have been identified yet, mainly because interstellar magnetic fields deflect the (charged) cosmic rays, causing them to follow curved trajectories through space (see figure 9).

Elementary particles called *neutrinos* provide a third way to study the universe. Neutrinos are produced in almost every object and have a very low probability to interact with anything. Consequently a huge number of neutrinos reach the Earth. Neutrinos are electrically uncharged, so they propagate through the Universe in a straight line. These

properties make neutrinos ideal 'messengers' to study the densest and most distant regions of the Universe, and to pinpoint the sources of ultra high energy cosmic rays.

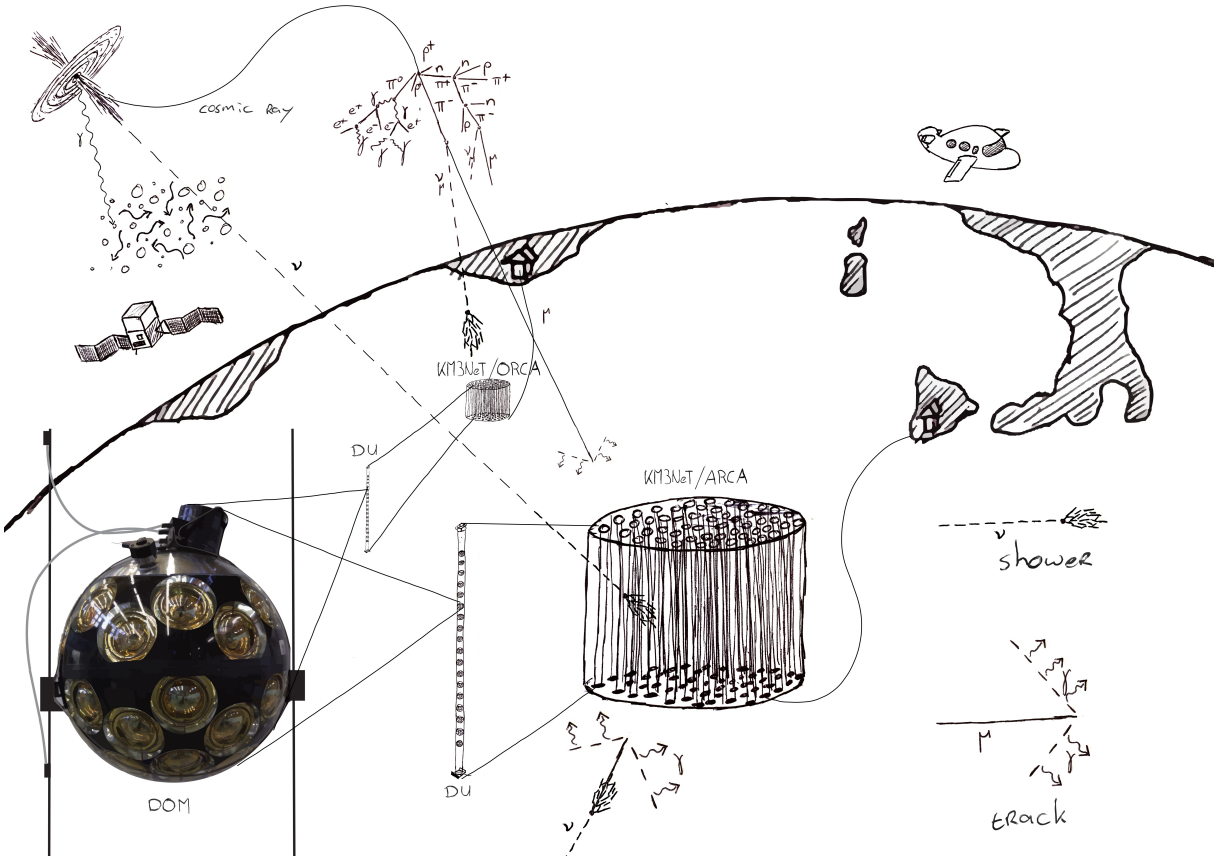


Figure 9: *The production, propagation and detection of elementary particles from the universe with the KM3NeT detectors.*

How can we detect these neutrinos? Their very low interaction probability makes neutrinos notoriously hard to detect. One way to 'catch' them is with very large *neutrino telescopes*; networks of optical sensors in (sea)water or ice. Neutrinos interacting with an atom in the detector cause the production of a shower of secondary particles, similar to the way a shower is produced in cosmic ray interactions. While propagating through the water or ice, a fraction of the particles in the shower induce the emission of light. This *Cherenkov radiation* can be detected by the optical sensors of the telescope, providing an indirect way to detect neutrinos.

The properties of the interacting neutrino such as its direction and energy, can be estimated from the pattern of detected light. These patterns, or *events*, can be categorised in different 'signatures'. The most frequently detected signatures are *tracks* and *showers*. A track is caused by a high-energy muon, producing a trace-like pattern of light in the detector. A shower-like signature is characterised by an expanding sphere of light, emitted from a single point close to the interaction.

Since neutrinos are so difficult to 'catch', neutrino telescopes are generally very large. They are commonly located at a great depth, far from daylight and to suppress the background of atmospheric muons. A cost-efficient location for a neutrino telescope is the deep sea or the (Ant)Arctic ice. Two examples are ANTARES in the Mediterranean sea and IceCube in the Antarctic ice, both operational for over more than ten years. A neutrino telescope called KM3NeT ('cubic kilometer neutrino telescope') is currently under construction in the Mediterranean Sea.

KM3NeT consist of two detectors; KM3NeT/ARCA and KM3NeT/ORCA (*Astroparticle-* and *Oscillation Research with Cosmics in the Abyss*, respectively). KM3NeT/ARCA aims at the detection of high-energy neutrinos from the cosmos, KM3NeT/ORCA is build to determine the elementary properties of neutrinos. The detectors are also used for oceanographic research and tomography of the Earth. KM3NeT/ARCA instruments a cylinder with a height of 612 m and a radius of 538 m, at a depth of ≈3 km below the sea surface. The KM3NeT/ORCA detector uses the same instruments, but is (much) more compact than KM3NeT/ARCA.

The structure of KM3NeT/ARCA and KM3NeT/ORCA is schematically shown in figure 9. The principal detection elements of both detectors are very sensitive optical sensors called photomultiplier tubes (PMTs). 31 PMTs and the accompanying electronics are protected against the pressure of the deep sea by the glass sphere of a digital-optical module (DOM). Eighteen DOMs are attached to vertical lines ('detection units', or DUs), which are anchored to the seabed and kept upright by a buoy on the top. 115 DUs are placed within a circular surface on the seabed, forming one 'building block'.

Motivation and research topics

Over the years, Icecube and ANTARES have detected a large number of high-energy neutrinos, some of which are proven to be of cosmic origin[26–29]. Unfortunately, the directional accuracy of these detectors is not sufficient to pinpoint their source(s).

KM3NeT/ARCA has (once completed) a much better angular resolution than ANTARES and IceCube, and can identify the source(s) of high-energy cosmic neutrinos within a few years after completion[34]. The work presented in this thesis marks important stepping stones towards this goal:

- The calibration of the KM3NeT detectors, needed to achieve the best possible accuracy of the determined neutrino properties.
- Proving that the detectors works as expected, by analysing data obtained with the first deployed DUs.
- The development of a new statistical method to identify and study the source(s) of high-energy cosmic neutrinos. This method is unique in the sense that is proven to be capable of approaching the theoretical best possible sensitivity closer than ever before.

Some of the obtained results are highlighted in the following sections.

First data

KM3NeT/ARCA and KM3NeT/ORCA are build in several deployment phases, each time adding a few detection units to the detector(s). The data obtained with the first deployed DUs has been used to develop calibration techniques, characterise the detector and to verify the simulations.

In-situ calibration. An electrical signal is produced in case a photon is detected by one of the PMTs. The moment at which this signal exceeds a certain threshold (the *leading edge*) and the time that the signal remains above the threshold (the *time-over-threshold*, or ToT) are recorded with an accuracy of one billionth of a second (a nanosecond). Both quantities are stored to disc, forming a 'level-0 hit' (L0).

It takes on average about 30 ns between the moment a photon hits a PMT and the moment when an L0 is recorded. Determining this PMT-specific time delay is one of the objectives of the 'in-situ' calibration. The in-situ calibration exploits the light induced by ^{40}K decays and atmospheric muons, providing an inexpensive way to monitor the detector during operation, over a long period of time.

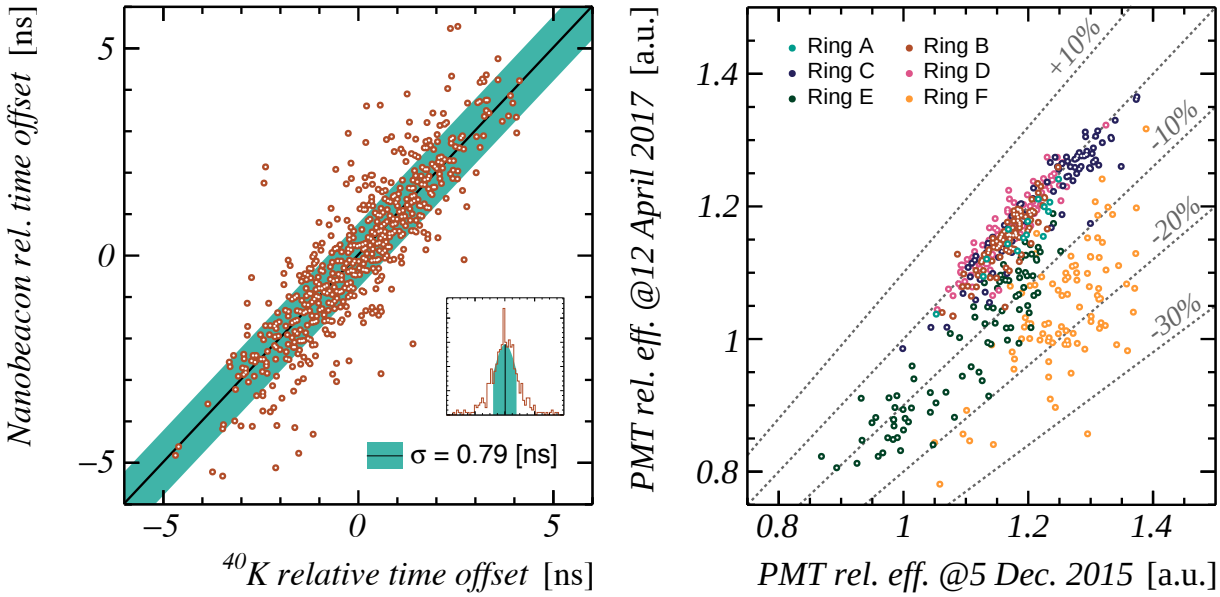


Figure 10: *Comparisons between the properties of the PMTs of KM3NeT/ARCA (2 DUs), obtained with the in-situ calibration.* ***Left:*** *A comparison of the relative time differences between the 31 PMTs in a DOM, as determined from the arrival times of light emitted by LEDs ('nanobeacons', or NB) placed in the DOMs[30, 31] and the calibration using ^{40}K decays. Both methods agree, with a standard deviation of* 0.79 *ns.* ***Right:*** *The efficiency of PMTs determined on December 5, 2015 as a function of the efficiency on April 12, 2017. The effect of sedimentation is visible on the PMTs facing upwards (rings E and F). The efficiency of the PMTs looking down (rings A-D) is very stable over time.*

The developped in-situ calibration techniques have been compared to other calibration methods. An example of such a comparison is given in figure 10 (left). The agreement

between both calibration techniques is better than one nanosecond. This is well within the specifications to determine the properties of detected neutrinos with the foreseen accuracy.

In addition to the time calibration, the in-situ calibration is also used to determine the probability that an incident photon is detected by the PMT (the PMT efficiency) and the statistical spread in the transit time of each PMT. The determined PMT efficiencies are accurate enough to distinguish the attachment of the DOMs to the DUs and to monitor the effects of sedimentation on the top of each DOMs (see figure 10, right).

Depth dependence of the atmospheric muon flux. The PMT efficiencies, determined with the in-situ calibration, are used in detailed computer simulations of events caused by neutrinos, atmospheric muons and ^{40}K decays. These simulations are cross-checked with data, after which they can be used to characterise the performance of the detector. An example of such a cross-check is given in figure 11 (left), in which the distribution of the number of PMTs per DOM recording light at approximately the same time (a *coincidence*) is shown. Coincidences of two or three PMTs are primarily caused by ^{40}K decays, while coincidences of seven or more PMTs are almost exclusively caused by atmospheric muons. The distributions of the simulations and the data agree.

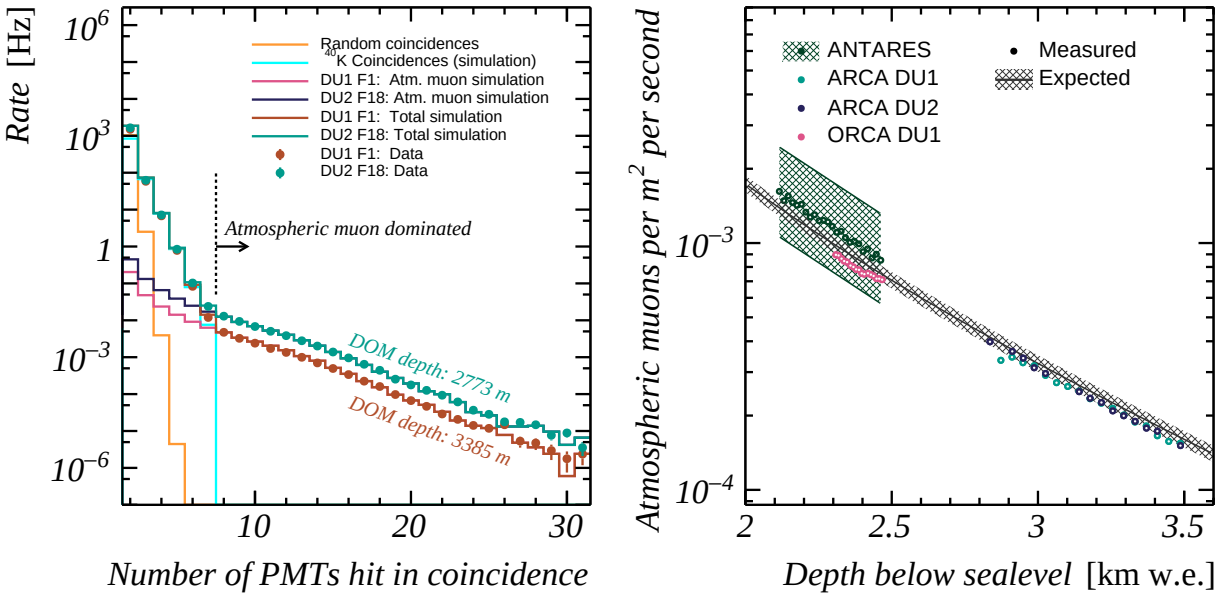


Figure 11: ***Left:*** *The distribution of the number of PMTs that detect one (or more) L0(s) coincidentally. The measured distributions of two DOMs at different depths are shown, as well as simulations.* ***Right:*** *The direction-integrated flux of atmospheric muons, as a function of depth below sea level. Each DOM of KM3NeT/ORCA and KM3NeT/ARCA provides an individual measurement of the flux. The density of seawater (*1.038 *g/cm*^{3} *) was used to calculate the equivalent depth in freshwater. The measurements are compared with the expected flux[32] and an independent measurement of the ANTARES detector[33].*

Atmospheric muons loose energy while propagating through (sea)water, leading to a decreasing *flux* (the number of atmospheric muons per m^2 per second) with increasing

depth. This effect can be observed in the ≥7-fold coincidence rate, which decreases with the depth of the DOM (see figure 11). Two different simulations of atmospheric muons, one using nominal PMT sensitivities and one with realistic PMT efficiencies (determined with the in-situ calibration technique discussed before), are used to calculate the atmospheric muon flux from the detected ≥7-fold coincidence rate. The results are given in figure 11 (right). The measured atmospheric muon flux agrees with the expected flux and with independent measurements of the ANTARES detector[33].

Reconstruction of track-like event signatures The most likely properties of events detected in the first year of operation have been estimated using a reconstruction algorithm. Two of these events are given in figure 12, showing the relation between the detection time of each L0 in the event as a function of the height of the corresponding DOM. The event in the left panel is best fit by a muon propagating almost straight down along the DU, at an angle of 84 ° relative to the horizon. Since the vast majority of all downward moving muons are produced in the atmosphere, this event is most likely caused by an atmospheric muon. The event in the right panel, on the other hand, is most likely caused by a muon moving in an upward direction. This event is therefore potentially caused by a neutrino propagating through the Earth before interacting in the vicinity of the detector.

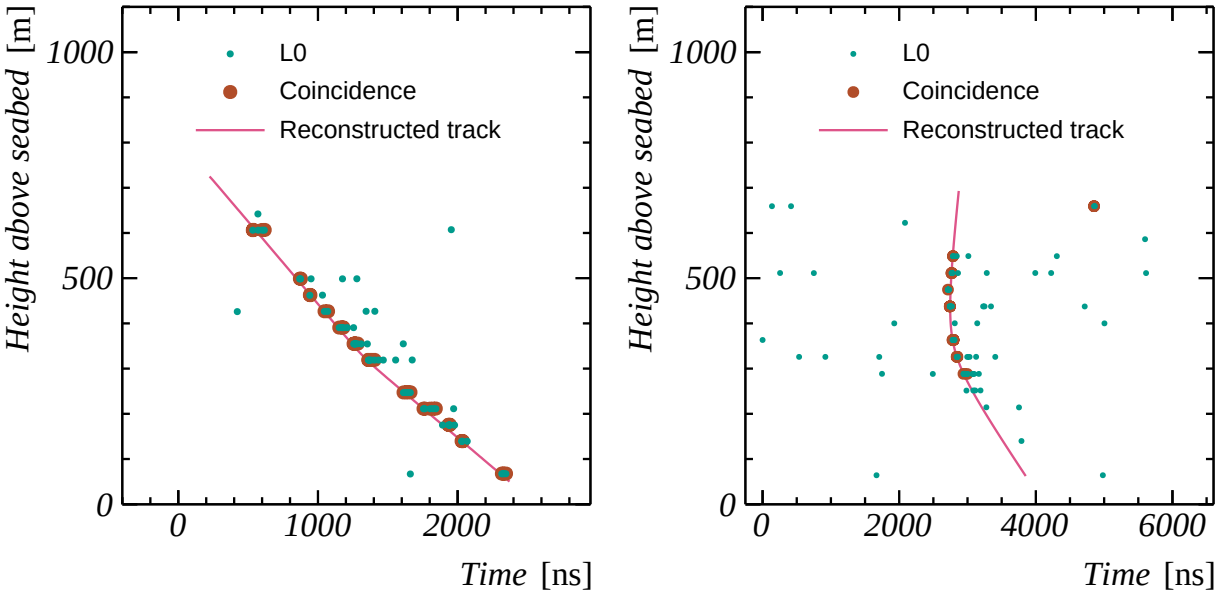


Figure 12: *The relation between the detection times (horizontal axis) of L0s on DOMs at different heights (vertical axis), of the first DU of KM3NeT/ARCA. Coincidences of two or more L0s on a single DOM are shown as well. The left figure shows one of the first detected events, on December 4, 2015. The right figure shows an event possibly originating from a neutrino. The most probable properties of a muon track causing each event have been reconstructed. The line shows the expected arrival time of the first detectable light from these reconstructed tracks.*

Cosmic neutrino sources.

A new statistical method to identify and study cosmic neutrino source(s) with KM3NeT/ARCA has been developped. It follows the principles of a statistical *hypothesis test*; the existence of a potential collection of cosmic neutrino sources (a $\mathcal{H}_1$ hypothesis) is tested against the null-hypothesis ($\mathcal{H}_0$) consisting of only backgrounds (mainly non-cosmic neutrinos and atmospheric muons).

The novel method is designed to be applicable to any potential source, providing the best possible *sensitivity* for each tested hypothesis. The better the sensitivity, the fewer detected events are needed to discover a potential source, and the more precise its properties can be determined.

Test statistic integral The backbone of the new approach is the likelihood $\mathcal{L}(\mathcal{E}|\mathcal{H})$ that a detected event $\mathcal{E}$ is caused by a hypothesis $\mathcal{H}$. The more information about the source, detector and event is used this likelihood, the more accurate it can be calculated and the better the sensitivity of the hypothesis test. The new method aims at calculating this likelihood in the theoretical best possible way, namely by evaluating the following integral

$$\mathcal{L}(\mathcal{E}|\mathcal{H}) \equiv \int \mathcal{L}(\mathcal{E}_i|\boldsymbol{x}_s) \cdot P^{det}(\boldsymbol{x}_s) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) d\boldsymbol{x}_s, \tag{3}$$

which is called the *test statistic integral*. The ingredients of this integral will be discussed in the following paragraphs.

Event likelihood landscapes. The main difference between the new method and conventional methods is the use of so-called *event likelihood landscapes*. Each of these landscapes describes the probability $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)$ that an event $\mathcal{E}$ is caused by a certain event *signature*. The event signature is parameterised by eight parameters $\boldsymbol{x}_s$, such as the direction of the interacting neutrino and two energy parameters describing both shower- and track-like events.

The event likelihood $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)$ is defined as the product of the probabilities that each of the PMTs in KM3NeT/ARCA did (or did not) record one or more L0s (in a certain timewindow) at a specific time. In this way, it contains almost all available information of the event. An impression of an event likelihood landscape is given in figure 13.

Expected event signature flux Any potential (collection of) cosmic neutrino source(s) is/are characterised by its *flux*; the expected number of neutrinos per unit time, direction and energy. This flux is multiplied by the probability that a neutrino reaches and interacts within the detector, giving an event signature with parameters $\boldsymbol{x}_s$ to calculate the *event signature flux* $\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}_1)$.

Detection probability A third ingredient of the new method is a function describing the probability $P^{det}(\boldsymbol{x}_s)$ that an event signature with parameters $\boldsymbol{x}_s$ is detected. Simulations are used to obtain a detailed parameterisation of the detection probability.

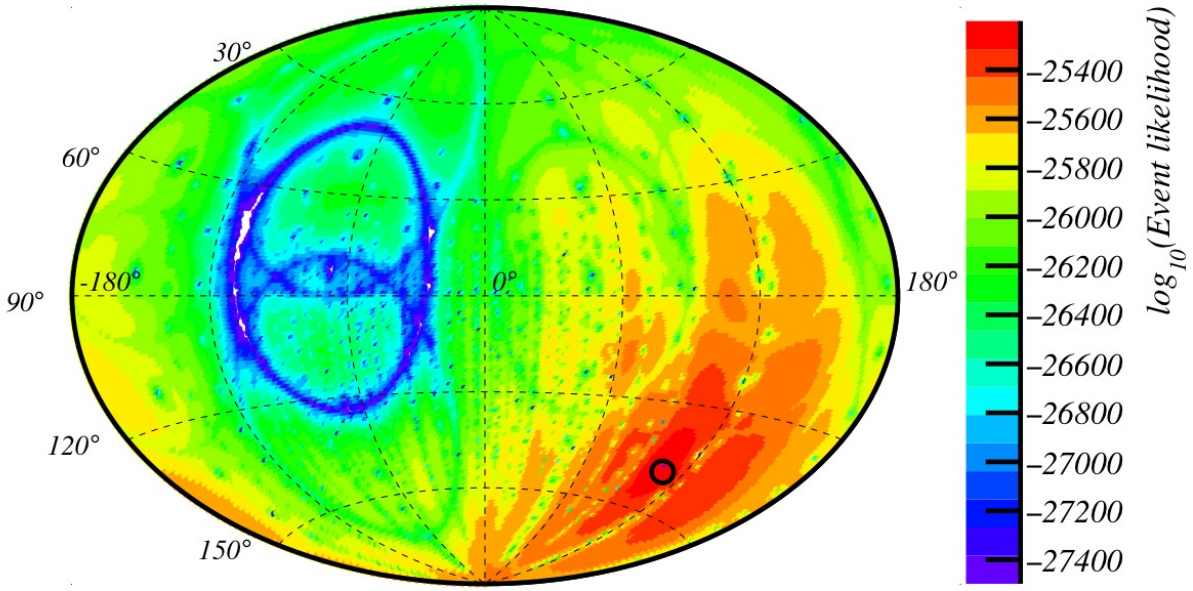


Figure 13: *A cross-section of an event likelihood landscape, obtained from simulations. The direction of the signature is varied, the other quantities are fixed to the values of the simulated event. The black circle indicates the simulated direction.*

Test statistic The test statistic integrals of the detected events (the data $\mathcal{D}$) are used to calculate how much more likely $\mathcal{H}_1$ is than $\mathcal{H}_0$:

$$\lambda(\mathcal{D}|\mathcal{H}_1) = \frac{\mathcal{L}(\mathcal{D}|\mathcal{H}_1)}{\mathcal{L}(\mathcal{D}|\mathcal{H}_0)} = \exp\left(-\mu^{det}_{\mathcal{H}_1-\mathcal{H}_0}\right) \times \prod_{\text{events}} \left[\frac{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)}{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)}\right], \tag{4}$$

where $\mathcal{L}(\mathcal{D}|\mathcal{H})$ denotes the likelihood to observe the data if hypothesis $\mathcal{H}$ is true and $\mu^{det}_{\mathcal{H}_1-\mathcal{H}_0}$ is the expected number of detected events from the assumed cosmic neutrino source(s). The likelihood ratio $\lambda(\mathcal{D}|\mathcal{H}_1)$ is called the *test statistic*. A graphical representation of equation 4 is given in figure 14.

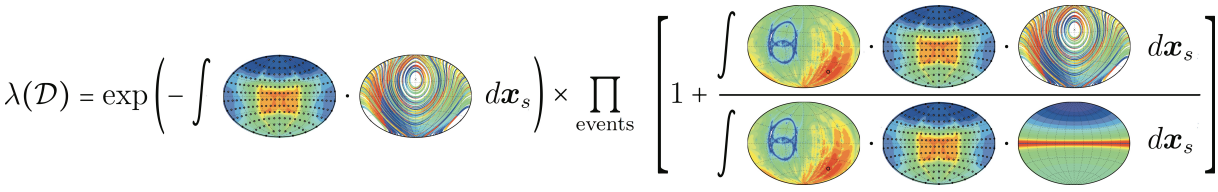


Figure 14: *An illustration of the test statistic $\lambda(\mathcal{D})$. Besides the probability landscapes (displayed in simplified form with), this integral depends on the detection efficiency ( ) and the expected event signature flux from backgrounds ( ) and from cosmic neutrino sources ( ). In reality, these functions are eight-dimensional.*

The $\mathcal{H}_1$ hypothesis can be accepted if the test statistic is greater than a certain value $\lambda_{\text{crit.}}$. This can be the case if a few events are caused almost certainly by cosmic neutrinos from $\mathcal{H}_1$, or if there are many events with a relatively small probability to originate from

$\mathcal{H}_1$. The value of $\lambda_{\text{crit.}}$ depends on the $\mathcal{H}_1$ hypothesis being tested, and is chosen such that the probability of a false discovery ($\mathcal{H}_1$ is accepted while $\mathcal{H}_0$ is true) is less than 0.000057%.

Advantages and disadvantages The most significant advantages of the new method over existing methods are:

- The new method includes maximum information, therefore maximising the sensitivity to a potential neutrino source:
  - In conventional methods, only the reconstructed, i.e. the most probable, neutrino properties are used in the analysis. The new method uses all possible solutions.
  - The neutrino properties cannot be determined with absolute precision. In conventional methods, the inaccuracies of the determined neutrino properties are approximated with relatively simple parameterisations. In the new method, the most accurate way to parameterise these inaccuracies is used, since they are described by the event likelihood landscapes.
- The new method is very flexible, in the sense that each of the ingredients in the test statistic is easily interchangeable with another/improved function, without the need to update the other ingredients:
  - Any potential source (collection) can be tested, by using the corresponding event signature flux in the test statistic integral.
  - Data collected with different configurations of KM3NeT/ARCA can be combined. To this extent, only the detection efficiency has to be changed. This makes it possible to use events detected during the years KM3NeT is still under construction.
  - Even more event information can be included in the event likelihood landscapes by using the number of detected L0s and the ToT information of the L0s. This option is discussed in appendix I.
- Selection criteria are used in conventional methods to exclude poorly and incorrectly reconstructed events from the analysis. Since the new method uses event likelihood landscapes rather than reconstructed neutrino parameters, there are no misreconstructed events. This eliminates the need for selection criteria, allowing all events to contribute to the identification of (a) neutrino source(s).
- In conventional methods, shower- and track-like event signatures are reconstructed with separate algorithms. A 'flavour identification' step is used to estimate the most likely signature. In the new method, both showers, tracks and a combination of a these two are described. Therefore, more information can be extracted from each event. It is relatively straightforward to extend the new method to include more exotic signatures.

The new method has one major drawback; a relatively large amount of computing power is needed to calculate each of the functions in the integral (especially the event likelihood landscapes), as well as the integral itself. This disadvantage was already anticipated at an early stage of the design, and has been deliberately accepted to see to which extent the new method is applicable.

The test statistic integral $\mathcal{L}(\mathcal{E}|\mathcal{H})$ has been calculated for 60 simulated events and (in addition to the $\mathcal{H}_1$hypothesis) for 121 different $\mathcal{H}_1$ hypotheses. The fact that these integrals can be evaluated shows that the new method can be used to test any potential source. At the moment, however, the 60 evaluated events are not sufficient to compare the performance of the new method with conventional methods.

Conclusions.

Techniques have been developped to calibrate the KM3NeT detectors, exploiting the backgrounds from ^{40}K decays and atmospheric muons. The time calibration is more accurate than one nanosecond, which is accurate enough to reconstruct (neutrino) events with good accuracy. The calibration of the PMT efficiencies is precise enough to distinguish small differences between the PMTs. Several cross-checks of data and simulations (using the determined PMT efficiencies), including a measurement of the depth dependence of the atmospheric muon flux with KM3NeT/ORCA and KM3NeT/ARCA, show that the detector is working as expected.

A new method to identify and study cosmic neutrino sources has been developped. This method is designed to approach the theoretically best possible sensitivity. This is done with event likelihood landscapes, detailed parameterisations of the expected flux of events and a new way to describe event signatures. An advantage of this approach is that the new method is very flexible: all possible cosmic sources can be tested, all data can be used and more complex signatures than showers and tracks are fairly easy to implement.

Unfortunately, this method is computationally demanding. To meet limitations in the available computing power, a hybrid method can be recommended. In this hybrid method, the conventional method is used for low-energy 'standard' events (well described with conventional reconstruction methods), while the new method is applied to events that 1) would otherwise not be included in the analysis (most notably the poorly reconstructed events), 2) the most high-energy events (for example, events with many L0s) and 3) events with a reconstructed direction close to a potential cosmic neutrino source(s).

Part I

Introduction

Abstract

Neutrinos interact only through the weak force. The corresponding cross section is very small. This makes them ideal 'messengers' to study the universe, a field called *neutrino astronomy*. One of the questions that could by answered by neutrino astronomy is the long-standing mystery of the cosmic origin of the highest energy cosmic rays.

The small interaction probability makes neutrinos notoriously hard to detect. A neutrino telescope, called KM3NeT/ARCA, is currently under construction to detect high-energy cosmic neutrinos. KM3NeT detects the light induced by secondary particles produced in neutrino interactions with a grid of optical sensors in the deep waters of the Mediterranean Sea. Computer algorithms are used to 'reconstruct' the parameters of interacting neutrinos from the detected light patterns.

The motivation and recent highlights of neutrino astronomy will be introduced in this part, followed by a description of the detection principle, simulations and reconstruction algorithms used in KM3NeT.

Chapter 1

Messengers from the Universe

The Universe has intrigued humanity for many centuries, as is clear from the alignment of large-scale structures like Stonehenge and early maps of the night sky like the Nebra disc[35] (figure 1.1), which dates back to 1600 BC. Alongside the moon in various phases, it shows several stars, including the Pleiades star cluster. The Pleiades cluster is located approximately 400 light-years away from the Earth, which means the light from the Pleiades recorded on the Nebra disc must have been emitted around 2000 BC. The most distant object that can be observed by the naked eye is the Andromeda Galaxy, located approximately 2.5 million light-years away.

Figure 1.1: *The Nebra sky disc. It is the oldest known map of the night sky, dated 1600 BC. Shown are the full moon, a waxing moon, several stars and the Pleiades star cluster[35].*

Photons With modern-day technology, it is possible to observe photons which were emitted 13.7 billion years ago (379,000 years after the big bang). These photons, known as the *cosmic microwave background*, give a unique insight into the processes that happened in the early Universe. On the high-energy side of the electromagnetic spectrum, γ-rays allow us to study high-energy processes happening in the more recent Universe.

A disadvantage of using photons to study the Universe lies in the visible Universe

having a limited range. For the cosmic microwave background (and any photon with an energy lower than 400 GeV), the visible horizon is limited by the moment when electrons bound to protons to form hydrogen (the *recombination* stage), making the Universe transparent for the first time. This moment in time corresponds to a visible horizon at about 10 Gpc. At photon energies above approximately 400 GeV, interactions of photons with matter, as well as photon-photon interactions with the cosmic microwave background, result in a significantly smaller mean free path. This visible limit is shown in figure 1.2. For photon energies between 10^5-10^8 GeV, even the closest major galaxy, Andromeda, is barely observable. Fortunately, photons are not the only messengers from the Universe.

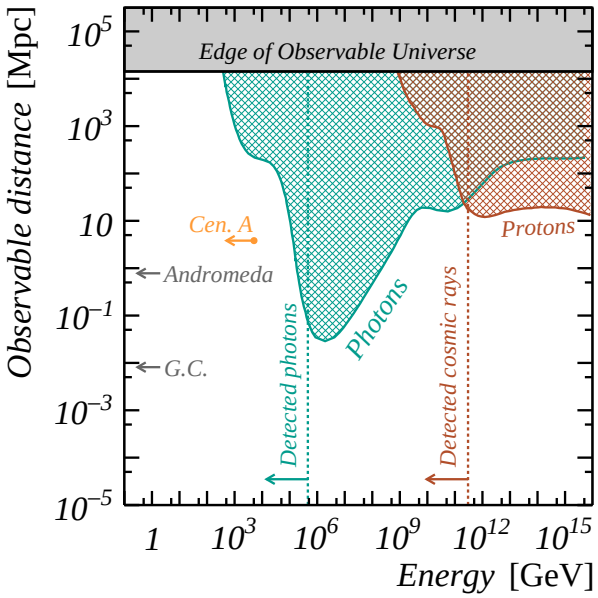


Figure 1.2: *The observable distance as function of the particle energy for photons and protons (reproduced from K. Hoffman[5], in which redshift effects are not included). For comparison, the distance to the galactic centre (GC) and Andromeda Galaxy (or M31, the nearest major galaxy to the Milky Way) are given. The intrinsic upper limit on the observable Universe, corresponding to the moment of recombination, is indicated by the shaded area. Vertical dotted lines give the highest energy photon and cosmic ray, detected by the Tibet ASγ and HiRes experiments respectively[36, 37]. The measured energy spectrum from the Centaurus A galaxy is given by the arrow labelled 'Cen. A' [38].*

Cosmic Rays Between 1911 and 1913, Victor Hess showed with a series of high altitude balloon flights, that a flux of down going particles traverses the Earth's atmosphere [39]. These particles were studied in more detail by Rossi and Auger, who proved that these particles originate from high-energy charged particles or cosmic origin. These particles, named *cosmic rays*, interact in the upper atmosphere, producing a 'cascade' of secondary particles called a *cosmic ray air shower*.

It has been observed that low-energy cosmic rays consist mainly of protons, with small contributions from heavier nuclei, electrons and antiparticles [4]. Recent observations suggest that the highest energy cosmic rays consist of heavier nuclei such as iron[40, 41]. Figure 1.3 shows the observed energy spectrum of cosmic rays. The rate of observed cosmic rays ranges from one interaction per square decimetre per second to one interaction per square kilometre per millennium. In order to probe this large range, cosmic rays are studied using a variety of detectors, ranging from satellite experiments to detect the low-energy cosmic rays, to large ground-based arrays of detectors to cover the highest energies. The cosmic ray with the highest energy ever detected had an energy of approximately 300

EeV ($3 \cdot 10^{20}$ eV, or 51 J)[14].

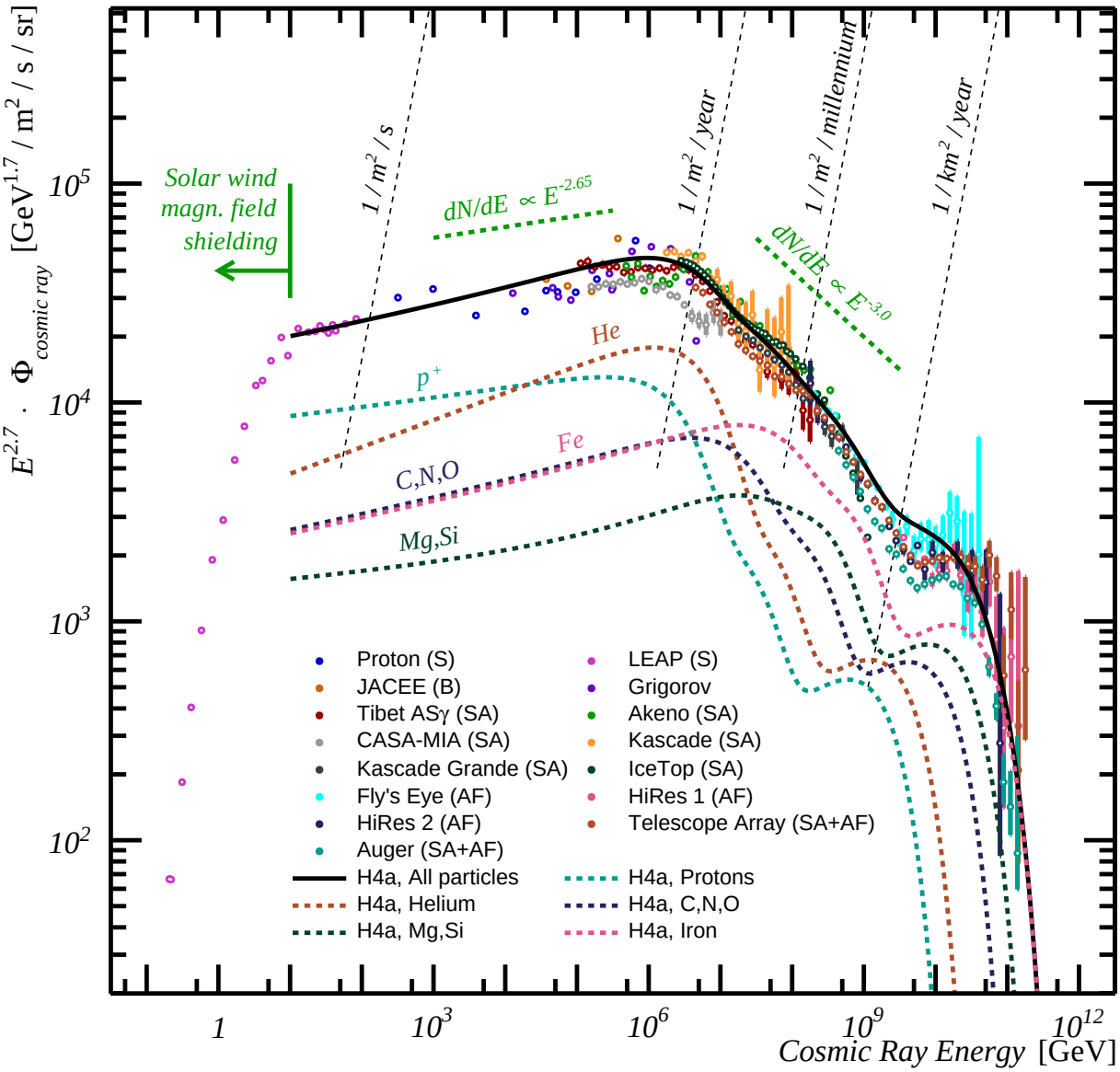


Figure 1.3: *The detected flux of cosmic rays as function of the energy (E) as observed by various experiments [42–59]. The shown flux (defined as the number of particles per unit energy, area, time and solid angle) is multiplied by $E^{2.7}$. The abbreviations used in the legend correspond to the detection technique(s) used; S: Satellite, B: Balloon, SA: surface array and AF: Air-Fluorescence. The lines give the predicted flux of the H4a model [60], split according to the contributions of different elements (coloured dotted lines) as well as the total predicted cosmic ray flux (black solid line). To guide the eye, the dotted black lines correspond to several fluxes of down-going particles, ranging from one particle per square metre per second to one particle per square kilometre per year.*

At energies below 10 GeV, the magnetic field and plasma outflow of the sun acts as

a shield for cosmic rays, deflecting the cosmic rays away from the Earth[61]. At higher energies, the cosmic ray energy spectrum can be categorised in a few regions[16]; up to an energy of approximately 3 PeV (a point called the 'knee'), the spectral index (γ in $dN/dE = E^{-\gamma}$) is approximately 2.7. The spectrum steepens to $dN/dE \approx E^{-3}$ at an energy of about 100 PeV (the 'second knee'). Above an energy of 3 EeV (the 'ankle'), the spectrum seems to flatten again, followed by a sharp 'cutoff' at approximately 30 EeV.

A possible explanation for the shape of the cosmic ray energy spectrum is based on the assumption that three populations of cosmic ray sources exist, each producing cosmic rays up to a certain energy per nucleon, causing the steepening of the spectrum at the 'knee', 'second knee' and 'cutoff' respectively. A model (called 'H4a') based on these assumptions is given by Gaisser[60]. In this model, the populations producing cosmic rays with energies up to the 'knee' and 'second knee' are related to galactic sources, while the highest energy cosmic rays are produced in extra-galactic sources. In figure 1.3, this model is shown.

An alternative explanation for the cutoff of the cosmic ray spectrum at the highest energies is given by the interactions of cosmic rays with the cosmic microwave background. Due to these interactions, the mean free path of protons with energies above 50 EeV is smaller than about 50 Mpc (see figure 1.2). Only a few objects within this distance are believed to be capable of producing cosmic rays with these energies, thus setting an upper limit on the flux of protons with energies above 50 EeV. This limit is called the *Greisen-Zatsepin-Kuzmin limit* (GZK limit). The GZK limit does not apply to heavier nuclei (such as iron). Instead, interactions with the cosmic microwave background at these energies cause a break up into lighter nuclei (*photodisintegration*), resulting in a similar limit (see e.g. Batista et al.[62]).

Despite being the subject of an active field of research, the exact origin of cosmic rays is currently unknown. The identification of cosmic ray sources is challenging since (due to their charge) cosmic rays are deflected by interstellar magnetic fields on their journey to Earth, and therefore do not point back to their origin(s).

Neutrinos Neutrinos offer an alternative means of studying the Universe. Neutrinos are elementary particles interacting only under influence of the weak force. Their existence was postulated in 1930 by Wolfgang Pauli, to explain the observed energy distribution of electrons produced in β-decays [63]. Since neutrinos only interact through the weak force, their interaction cross-section is so low that Pauli feared neutrinos would never be detected. Nevertheless, the first neutrinos were observed by Cowan and Reines in 1956, using a detector situated eleven metres away from a nuclear reactor* [64].

Neutrinos come in three flavours: electron (ν_e), muon (ν_μ) and tau (ν_τ) neutrinos, each with their corresponding antiparticle ($\bar{\nu}_e, \bar{\nu}_\mu, \bar{\nu}_\tau$). The neutrinos detected by Cowan and Reines were electron-antineutrinos. The muon neutrino was first detected by Ledermann, Schwartz and Steinberger† [65], and the tau neutrino by Kodama et al.[66]. No more than three neutrino flavours exist that interact through the weak interaction[67]. There are theories that postulate the existence of one or more additional neutrino types (named

*Nobel Prize in Physics (1995), "For the detection of the neutrino".

†Nobel Prize in Physics (1988), "For the neutrino beam method and the demonstration of the doublet structure of the leptons through the discovery of the muon neutrino"

sterile neutrinos), which do not interact through any of the forces in the standard model.

Neutrinos do not have electric charge. Therefore, their trajectories are unaffected by magnetic fields, and they propagate through the Universe in straight trajectories. Additionally, their low interaction cross section allows them to propagate through dense regions which are opaque to light and other particles. For these reasons, neutrinos are ideal to probe the furthest and densest regions in the Universe.

Gravitational Waves The recent discovery of gravitational waves gives direct evidence of the general theory of relativity[68], while opening a new window into the Universe. Gravitational wave detectors are particularly sensitive to the merging of two massive objects. From the observed signal, a large variety of properties of the merging objects can be deduced[69].

1.1 Connection between Messengers

Neutrinos and photons are expected to be produced in interactions of cosmic rays with matter and photons surrounding their sources, as will be discussed in this section. Because of this connection, information from neutrinos and photons can be used to study sources of cosmic rays and vice-versa.

Hadronuclear interactions The main interaction channels of protons with matter are:

$$p + p \rightarrow \begin{cases} p + n + \pi^+, \\ p + p + \pi^0, \end{cases} \tag{1.1}$$

and

$$p + n \rightarrow \begin{cases} p + n + \pi^0, \\ p + p + \pi^-. \end{cases} \tag{1.2}$$

An (approximately) equal ratio of produced charged pions (π^+ and π^-) and neutral pions (π^0) is expected. The production of pions via these interactions is called the *hadronuclear* scenario.

Photohadronic interactions In addition to hadronuclear interactions, cosmic rays can interact with photons. In this *photohadronic* scenario, the leading interactions take place via the production of a (virtual) Δ^+ resonance:

$$p + \gamma \rightarrow \Delta^+ \rightarrow \begin{cases} p + \pi^0 & (\approx 67\ \%), \\ n + \pi^+ & (\approx 33\ \%). \end{cases} \tag{1.3}$$

Pion decays Pions produced in hadronuclear and photohadronic interactions decay. While neutral pions generally decay to two photons, charged pions decay predominantly

to a (anti)muon and muon (anti)neutrino. In turn, the (anti)muon decays to an (anti)electron and two neutrinos:

$$\begin{aligned} \pi^0 &\rightarrow \gamma + \gamma, \\ \pi^- &\rightarrow \mu^- + \bar{\nu}_\mu, \\ \mu^- &\rightarrow e^- + \nu_\mu + \bar{\nu}_e, \\ \pi^+ &\rightarrow \mu^+ + \nu_\mu, \\ \mu^+ &\rightarrow e^+ + \bar{\nu}_\mu + \nu_e. \end{aligned} \tag{1.4}$$

In both the hadronuclear and photohadronic scenario, with subsequent pion decay as in equation 1.4, the produced neutrinos carry about five per cent of the primary proton energy [70, 71]. Therefore, the energy spectrum of the produced neutrinos follows the energy spectrum of the accelerated cosmic rays. In comparison to neutrinos, the energy spectrum of the photons produced in both scenarios is less tightly connected to the cosmic ray energy spectrum, as photons lose energy through interactions and the production of e^+e^- pairs. In dense environments, they may even be blocked completely.

A major difference between the photohadronic and hadronuclear scenarios is the lack of negative pions produced in photohadronic interactions. Consequently, the ratio between produced neutrinos and antineutrinos is different. A detector capable of distinguishing neutrinos from antineutrinos[‡] would be capable of distinguishing these two scenarios.

The relation between cosmic rays, γ-rays and neutrinos underlines the advantage of combining the knowledge from each of these messengers. In the following two sections, the current knowledge on sources of γ-ray and cosmic rays sources will be summarised. The field of neutrino astronomy will be discussed in the next chapter.

1.2 Gamma-ray Sources

In figure 1.4, the directions of observed γ-rays with energies ≥ 1 TeV are shown. The majority of the γ-ray sources are located within in the galactic plane. A second, isotropically distributed, set of sources can be identified away from the galactic plane. Both populations correspond to different types of sources. In this section, these sources will be discussed. The relation between them is outlined in figure 1.5.

Supernova Remnants The majority of the identified sources of gamma rays within the Milky Way originate from explosions caused by the collapse of massive stars, called ***supernovae***. Depending on the mass of the progenitor star, a distinction between two types of supernovae can be made. In case the star has a mass less than approximately ten solar masses, it collapses to a ***white dwarf***. A white dwarf consist mainly of a plasma of electrons and has a size comparable to the Earth.

The collapse from a star to a white dwarf is not particularly spectacular, and does not qualify as a supernova. However, in case the white dwarf accretes sufficient surrounding matter (for example from a neighbouring heavy star) to surpass about 1.4 solar masses

[‡]For example by exploiting the enhanced antineutrino-electron cross section $\bar{\nu}_e + e^- \rightarrow W^- \rightarrow X$, see section 3.2.2

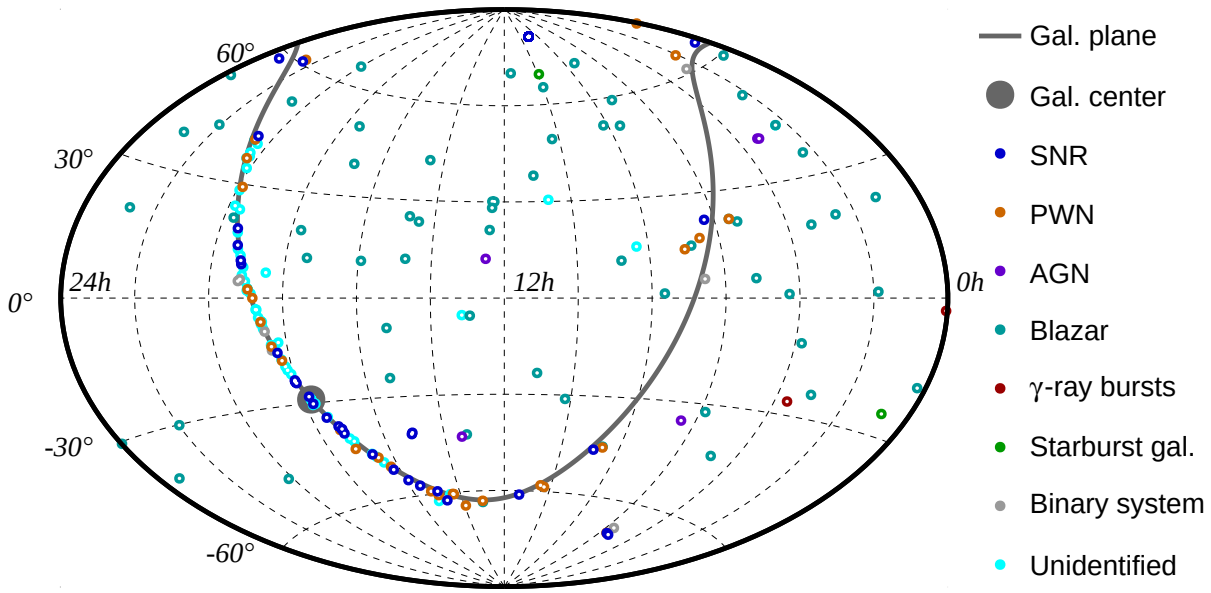


Figure 1.4: *The locations of ≥ 1 TeV γ-ray sources, in equatorial coordinates. Data obtained by the HESS, MAGIC and VERITAS collaborations, bundled by Wakely et al.[72]. Shown are the galactic plane (grey line) and the galactic centre (grey marker). Abbreviations are used for the following sources: supernova remnants (SNR), pulsar wind nebulae (PWN) and active galactic nuclei (AGN).*

(the 'Chandrasekhar limit'), the outward pressure of the electrons cannot withstand the inward pressure of the gravitational force, causing the white dwarf to collapse in a ***type Ia supernova***.

In case the progenitor star weighs more than ten solar masses, the star collapses without the intermediate white dwarf stage. These supernovae are labelled ***type II supernovae***.

In a supernova, a shockwave of material is ejected into space. While sweeping up surrounding material, this shockwave is hardly slowed down, expanding for periods up to thousands of years. This shockwave forms the ***supernova remnant***, together with the remainder of the progenitor star. Depending on the mass of the collapsing star, this can either be a black hole or a neutron star.

Neutron stars As their name suggests, *neutron stars* consist mainly of neutrons. In contrast to white dwarfs, there are nearly no electrons to withstand the inward gravitational force. Consequently, neutron stars are some of the densest objects in the Universe, only surpassed by black holes. Conservation of angular momentum causes them to be some of the fastest spinning cosmic objects, making up to several hundreds of rotations per second.

Based on indirect proof from measurements of emitted γ-rays, neutron stars are believed to exhibit the strongest magnetic fields (up to $\approx 10^{10}$ tesla[73]), which is about 10^7 times stronger than some of the strongest human-made magnetic fields[74]. The mechanism giving neutron stars their magnetic fields is not yet understood[73]. In some neutron

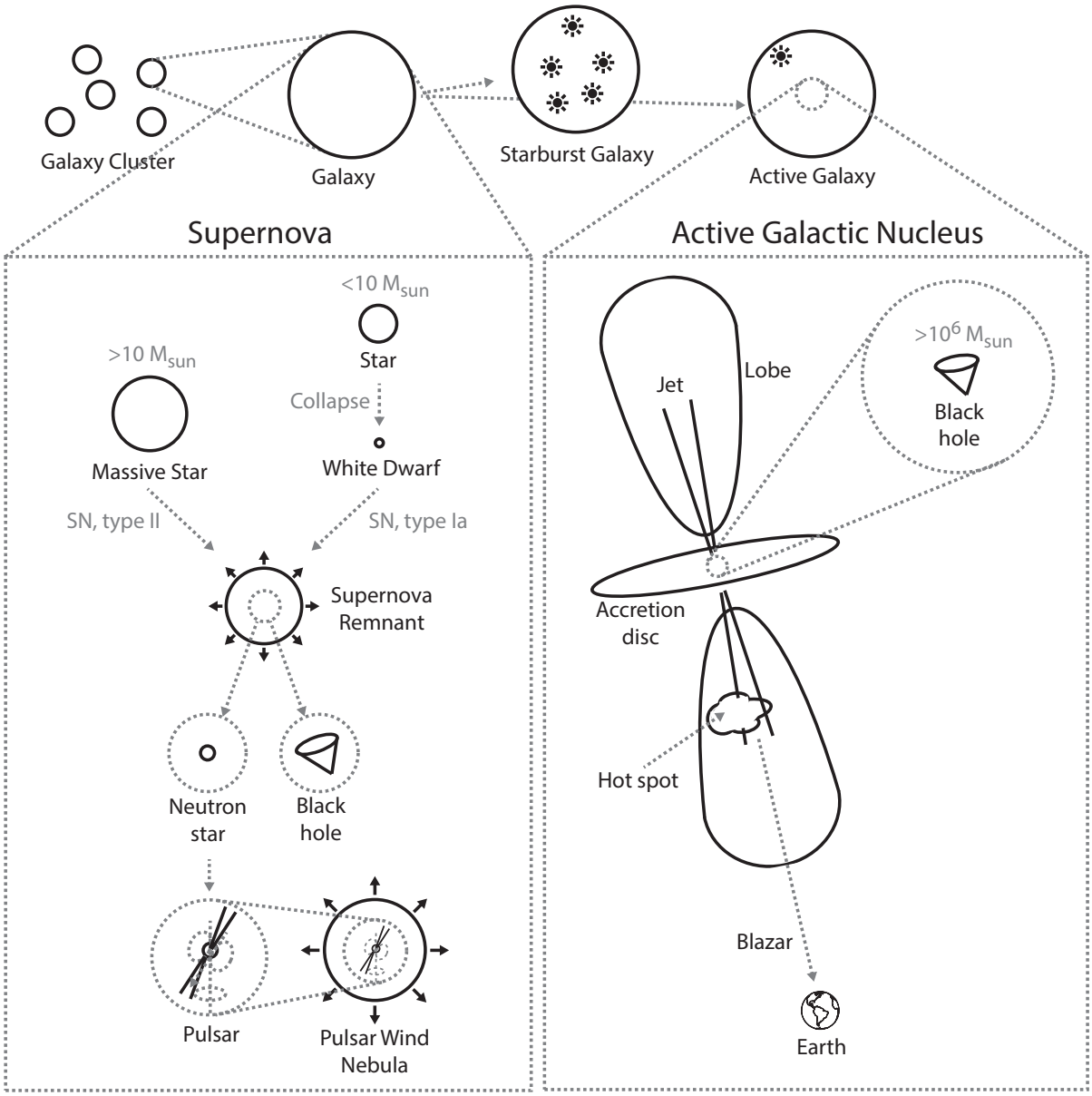


Figure 1.5: *Sketch of the relations between several γ-ray sources. Please see the text for a discussion of the different features.*

stars, called ***pulsars***, strongly fluctuating magnetic fields are a likely source of γ-ray beams radiated from their magnetic poles. Pulsars of which the beams are not aligned with the rotation axis have been detected using the characteristic lighthouse-like signature caused by the electromagnetic beams sweeping the Universe.

As special name is given to supernova remnants with a pulsar in its centre; ***pulsar wind nebulae***. A famous example of a pulsar wind nebula is the Crab nebula.

Gamma-ray bursts Some supernovae are accompanied by extremely luminous bursts of γ-rays. These ***γ-ray bursts*** last from a few milliseconds up to minutes. The long bursts

($\geq$ 2 seconds) are believed to be connected to the collapse of a massive star in a supernova of type II, while short burst ($\leq$ 2 seconds) are caused by the merger of binary systems of two neutron stars.

Starburst galaxies About 7% of all galaxies have a star formation rate exceeding typical galaxies by a factor of 100 or more. These galaxies are called starburst galaxies. The accretion of matter and collisions of the newly formed stars leads to a high rate of supernovae and thus a sizeable flux of emitted radiation.

Active galactic nuclei Some galaxies contain a very luminous core that outshines the outer regions by orders of magnitude. Such a galaxy is called an 'active galaxy', its centre an ***active galactic nucleus*** (AGN). Most AGNs contain a supermassive ($\geq 10^6$ solar masses) black hole. The most luminous AGNs are known as *quasars*

An AGN consists of a few regions (see figure 1.5). The supermassive black hole in the core forms, together with the disc from which it accretes matter, the ***AGN core region***. Some AGN are accompanied by two ***AGN jets***, consisting of ultrarelativistic particles ejected along the rotation axis of the system. The regions around these, as well as the points where jets interact with intergalactic matter, can cause significant radio emission. These regions are called ***AGN lobes*** and ***AGN hot spots***, respectively.

A special case of active galaxies are those with one of the AGN jets pointed towards the Earth. These objects are luminous objects with largely varying intensity, called ***blazars***. Even though blazars are relatively rare, their characteristic emission profile makes them relatively easy to identify. Most of the extra-galactic TeV γ-ray sources are classified as blazars.

1.3 Cosmic Ray Sources

As discussed in section 1.1, sources of γ-rays are likely candidates for the sources of cosmic rays. Cosmic rays with energies below $\approx$ 3 PeV (the 'knee') are likely to be accelerated by supernova remnants in the galactic plane (see section 1). At the moment of writing, two supernova remnants (IC 443 and W44) have been identified as likely sources of cosmic rays [15]. This was discovered indirectly, by observing the $\pi^0 \rightarrow \gamma + \gamma$ decay signature in the γ-ray spectra of these sources.

Direct identification of cosmic ray sources is difficult, due to the curved trajectories of cosmic rays. Only for the highest energy cosmic rays, the deflections are reasonably small (of the order of a few degrees), and the arrival directions of cosmic rays can be associated to their original sources. Figure 1.6 shows the direction of 27 cosmic rays with reconstructed energies higher than 57 EeV, as detected by the Pierre Auger observatory[40].

A galactic origin of these high-energy cosmic rays can be excluded from the fact that they are not distributed close to the galactic disc. Additional proof is given in a more recent study by Aab et al.[78]. Instead, the correlation with the locations of 472 close-by active galactic nuclei makes these likely sources. The most significant cluster of cosmic rays is located close to Centaurus A; a close-by galaxy that is particularly prominent in the radio regime of the electromagnetic spectrum.

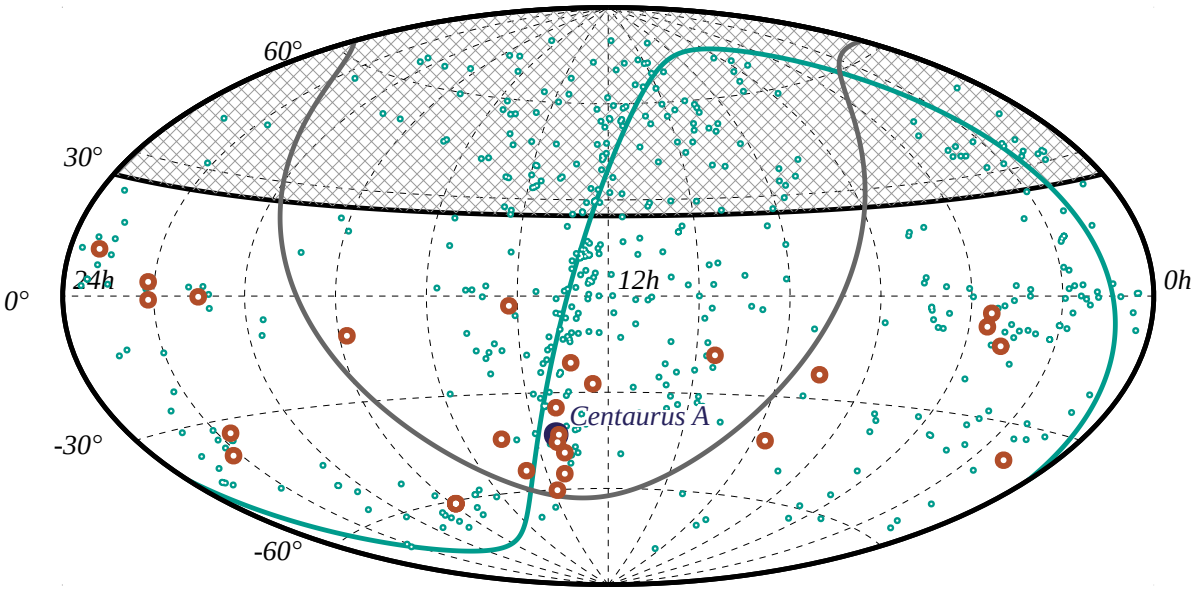


Figure 1.6: *Locations of 27 cosmic rays with reconstructed energy >57 EeV (orange markers), as well as the locations of known nearby AGNs (within 75 Mpc) given by Véron et all. [75], extracted from the Vizier catalogue[76], in blue. This figure is reproduced from the analysis discussed in the paper by the Auger Collaboration[40, 77]. The blue-green line indicates the super-galactic plane; a plane in which many close-by galaxies are found. The grey shaded area indicates the region where the detector cannot see anything.*

Even though supernova remnants and active galactic nuclei are likely to be sources of cosmic rays, these objects do not radiate sufficient thermal energy to produce the highest energies of cosmic rays. A possible explanation has been proposed by Enrico Fermi, in which cosmic rays are emitted with relatively low energies, and are subsequently accelerated to higher energies [79].

In its original form (now known as 'Second order Fermi acceleration'), charged particles are accelerated by randomly moving clouds of ionised gas. Deflections due to surrounding magnetic fields cause the particles to cross the clouds multiple times. In each crossing, the energy gained by a particle with energy $\bar{E}_i$ is proportional to the square of the velocity of the cloud [79]:

$$\frac{\bar{E}_{i+1} - \bar{E}_i}{\bar{E}_i} \propto \left(\frac{v}{c}\right)^2 \equiv \beta^2, \tag{1.5}$$

where $\bar{E}_{i+1}$ denotes the average energy of the charged particle after the crossing, v the speed of the moving cloud, c gives the speed of light and β denotes the ratio between them. Estimated values of β are of the order of 10^{-2} for typical acceleration sites like SNRs[80].

Even though the second order Fermi acceleration mechanism elegantly describes the observed power law spectrum, with a spectral index that is close to the observed cosmic ray spectrum, it is not a very efficient way to accelerate cosmic rays. Because of the small

value of β^2, a large number of crossings would be needed to reach the highest cosmic ray energies. A more efficient mechanism is a variation on this mechanism, called 'first order Fermi acceleration'. This process is believed to occur in the expanding shockwaves of supernova remnants, as well as in the jets from γ-ray burst, pulsars and active galactic nuclei.

1.3.1 First Order Fermi Acceleration

Suppose a shockwave of ionised gas moves through space with a sub-relativistic velocity v_S ($v_S/c \equiv \beta \ll 1$). Being deflected by the magnetic fields caused by (and moving along with) the ionised gas, charged particles can traverse the shockwave multiple times. In case the particle is already sufficiently relativistic (such that $E \approx p \cdot c$), and no energy is lost in the magnetic deflections ('elastic' deflection), the energy before (E_i) and after (E_{i+1}) a two-way transition of the shock front is given by:

$$E'_i = E_i \cdot \gamma \cdot (1 - \beta \cdot \cos(\theta_i)), \tag{1.6}$$

$$E'_{i+1} = E'_i, \qquad \text{(Elastic deflection)} \tag{1.7}$$

$$E_{i+1} = E'_{i+1} \cdot \gamma \cdot (1 + \beta \cdot \cos(\theta'_{i+1})), \tag{1.8}$$

$$= E_i \cdot \gamma^2 \cdot (1 - \beta \cdot \cos(\theta_i)) \cdot (1 + \beta \cdot \cos(\theta'_{i+1})), \tag{1.9}$$

where an apostrophe denotes a quantity in the rest frame of the moving shock, θ_i and θ_{i+1} correspond to the angle of the particle with respect to the shock plane and $\gamma \equiv 1/\sqrt{1-\beta^2}^{-1}$ is the Lorentz factor of the shockwave in the rest frame of the observer.

If the shockwave would be stationary, θ_i and θ'_{i+1} would be distributed isotropically and on average no energy would be gained. However, since the shockwave is moving, the probability of a particle crossing the shockwave is proportional to its velocity component perpendicular to the shockwave:

$$\frac{\partial P}{\partial \cos(\theta'_{i+1})} = 2\cos(\theta'_{i+1}), \qquad 0 \le \cos(\theta'_{i+1}) \le 1, \tag{1.10}$$

$$\frac{\partial P}{\partial \cos(\theta_i)} = -2\cos(\theta_i), \qquad -1 \le \cos(\theta_i) \le 0. \tag{1.11}$$

Using these relations, the average energy (denoted by $\bar{E}_{i+1}$) after each crossing (back and forth) is given by:

$$\begin{aligned} \bar{E}_{i+1} &= \int_{-1}^{0}\int_{0}^{1} \bar{E}_{i+1} \cdot \frac{\partial P}{\partial \cos(\theta'_{i+1})} d\cos(\theta'_{i+1}) \cdot \frac{\partial P}{\partial \cos(\theta_i)} d\cos(\theta_i), \\ &= \bar{E}_i \cdot \gamma^2 \cdot \left(\frac{2}{3}\beta + 1\right)^2 = \frac{E_i}{1-\beta^2} \cdot \left(\frac{4}{9}\beta^2 + \frac{4}{3}\beta + 1\right). \end{aligned} \tag{1.12}$$

Since the shockwave is moving at sub-relativistic speeds ($\beta \ll 1$), $\frac{4}{9}\beta^2 \ll \frac{4}{3}\beta + 1$ and $1 - \beta^2 \approx 1$. Consequently, the energy gain per shock crossing (back and forth) is given by:

$$\epsilon \equiv \frac{\bar{E}_{i+1} - \bar{E}_i}{\bar{E}_i} = \frac{1}{1-\beta^2}\left(\frac{4}{9}\beta^2 + \frac{4}{3}\beta + 1\right) - 1 \approx \frac{4}{3}\beta. \tag{1.13}$$

In contrast to the second order Fermi acceleration mechanism (equation 1.5), the average energy gain per crossing is in this case proportional to β, rather than β^2 (hence its name). This makes the first order mechanism a much more efficient mechanism, as a smaller number of crossings is needed to reach a certain cosmic ray energy. For first order Fermi acceleration, the average number of required crossings n to reach an average cosmic ray energy $\bar{E}$ is readily derived:

$$\bar{E} \approx \bar{E}_0 \cdot (1+\epsilon)^n \quad \Rightarrow \quad n \approx \frac{\log\left(\bar{E}/\bar{E}_0\right)}{\log(1+\epsilon)}. \tag{1.14}$$

Using this relatively simple mechanism, the $dN/dE \propto E^{-\gamma}$ cosmic ray energy spectrum is successfully explained. However, it does not (yet) predict the value of the cosmic ray energy spectral index γ.

After n crossings, the number of particles N with energy $\bar{E}$ is given by:

$$N = N_0 \cdot (1 - P_{esc.})^n, \tag{1.15}$$

$$= N_0 \cdot \left(\bar{E}/\bar{E}_0\right)^{\frac{\log(1-P_{esc.})}{\log(1+\epsilon)}}, \tag{1.16}$$

where N_0 denotes the number of injected particles, $P_{esc.}$ is introduced to denote the probability that a particle escapes from the acceleration region without crossing the shockwave any more, and equation 1.14 is used to substitute n. Consequently, the energy spectrum is:

$$\frac{\partial N}{\partial E} \propto E^{-\gamma}, \quad \gamma \equiv 1 - \frac{\log(1-P_{esc.})}{\log(1+\epsilon)}. \tag{1.17}$$

For typical astrophysical shocks believed to be efficient cosmic ray accelerators, the escape probability follows from kinetic gas theory [16]:

$$P_{esc.} \approx \frac{4}{3}\beta, \tag{1.18}$$

resulting in a spectral index:

$$\gamma \approx 1 - \frac{\log\left(1-\frac{4}{3}\beta\right)}{\log\left(1+\frac{4}{3}\beta\right)} \approx 1 + 1 + \frac{4}{3}\beta + \frac{1}{2}\left(\frac{4}{3}\beta\right)^2 + \ldots \approx 2. \tag{1.19}$$

where equations 1.13, 1.18 and 1.17 are combined. The derived value of $\gamma \approx 2$ does not include the effects of energy losses in collisions with surrounding matter. More detailed calculations including these effects indicate slightly harder spectra ($2.1 \leq \gamma \leq 2.4$)[81].

The observed cosmic ray energy spectrum is slightly steeper than the derived $dN/dE \propto E^{-2}$ spectrum (see figure 1.3). Similar to the explanation of the 'knee' in the cosmic ray energy spectrum (see section 1), this can be explained by considering the effect of the galactic magnetic field on cosmic ray propagation. A cosmic ray originating from a source within the Milky Way has a negligible probability to be observed at Earth in case its gyroradius (under influence of the galactic magnetic field) is larger than the distance from the source to Earth. As a result, the number of galactic sources 'in sight' decreases with the cosmic ray energy, causing a steeper observed spectrum.

As will be discussed in the next subsection, the gyroradius of cosmic rays can also be used to set a limit on the maximum energy of cosmic rays accelerated in the vicinity of various sources.

1.3.2 The Hillas Criterion

In order to be accelerated via the process of first order Fermi acceleration, a cosmic ray has to traverse the shock front multiple times. Therefore, the radius of the acceleration site r_S has to be larger than the gyroradius $r_g = p/(q \cdot B_S)$ of the particles it accelerates (where p and q denote the momentum and charge of the particle, and B_S is the magnetic field strength of the cosmic object). In the ultrarelativistic regime (where $E \approx p \cdot c$), this condition sets an upper limit on the energy $E_{max.}$ of accelerated cosmic rays:

$$\frac{E_{max.}}{q} = c \cdot r_S \cdot B_S. \tag{1.20}$$

This upper limit is known as the *Hillas criterion*. A graphical representation of the Hillas criterion is given in figure 1.7. In this figure, the (estimated) magnetic field strength and radii of several cosmic sources are given (data from Hillas and Fraschetti et al.[82, 83]). The diagonal lines correspond to limits of the energy per charge ($E_{max.}/q$) of produced cosmic rays. Several objects fulfil the Hillas criterion for energies larger than the ankle in the cosmic ray energy spectrum.

.

Motivation & Outline

The interactions of emitted matter and/or photons with surrounding matter should result in a detectable flux of neutrinos that, as a result of their low interaction cross section and lack of electric charge, can be considered as ideal messengers to answer the following questions:

1. Which objects exist beyond the observable distance?
2. What are the sources of (the highest energy) cosmic rays?
3. Which interactions play a role in the vicinity of these sources?
4. What is the energy spectrum of cosmic rays at their sources?

Only a fraction of the cosmic ray, γ-ray and astronomy research fields has been covered in this introduction. Various excellent books and papers have been written on these subjects, of which the recently published book by Gaisser, Engel and Resconi[16] ('Cosmic Rays and Particle Physics') deserves some particular recommendation.

The remainder of this thesis is structured in three parts. An introduction to neutrino astronomy, as well as the basic principles of neutrino detection will be discussed in the first part. This introduction will be focus mainly on the KM3NeT neutrino detector, used for the analyses discussed in parts two and three. In the second part, the performance of the KM3NeT detector is characterised, using the first data taken by the KM3NeT detectors. In the final part, a novel statistical method to identify cosmic neutrino sources is presented. At the expense of a larger computational effort, this method is proven to give the highest sensitivity for the identification of neutrino sources.

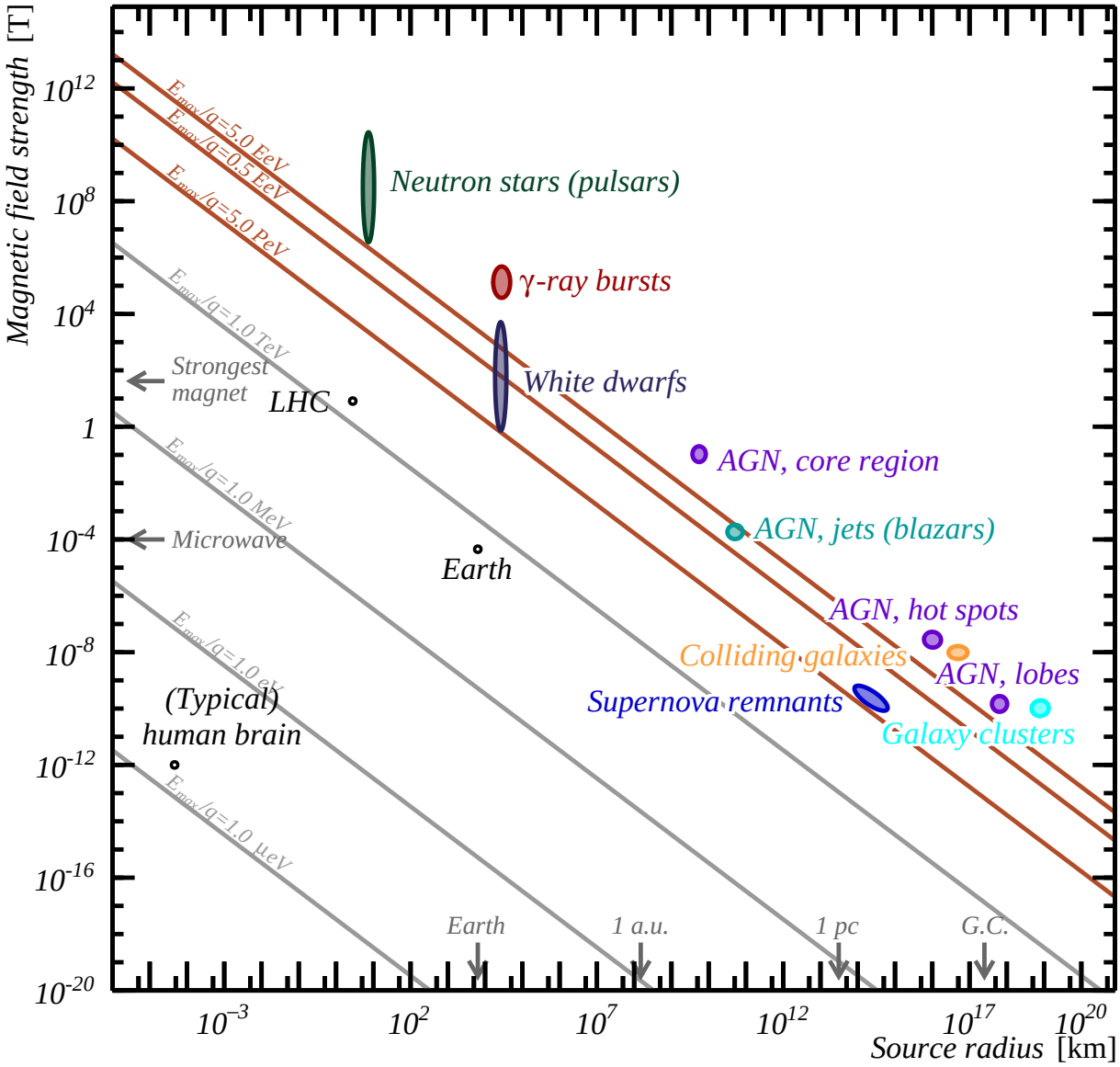


Figure 1.7: *A graphical representation of the Hillas criterion. All objects above a diagonal line are capable of accelerating cosmic rays with the energy per charge $E_{max.}/q$ specified next to the line. The orange lines correspond to the energy levels of the knee, second knee and ankle in the cosmic ray energy spectrum (see text), assuming a cosmic ray with q = 1 (i.e. a proton). For heavier nuclei, the corresponding lines lie lower. The radii and magnetic field strengths of the shown cosmic objects are taken from Hillas[82] and Fraschetti et al.[83]. A few non-cosmic objects are given to put these into perspective.*

Chapter 2

Neutrino Astronomy

As discussed in the previous chapter, neutrinos play a fundamental role in the search for the origin of (high-energy) cosmic rays. The field of astronomy that uses neutrinos to study the cosmos is called neutrino astronomy. The total flux of neutrinos detected at Earth is, however, dominated by low-energy neutrinos from various other sources. An overview of the current status of neutrino astronomy, as well as the predicted neutrino fluxes from background and signal sources will be discussed in this chapter.

2.1 Background Neutrino Sources

The expected flux of neutrinos from various background sources are shown in figure 2.1. The following sources are included:

- **Cosmic neutrino background (CνB)**: It is believed that neutrinos decoupled from matter about one second after the big bang (much earlier than photons, which decoupled from matter approximately 379000 years after the big bang). An indirect proof of the existence of this *cosmic neutrino background* (CνB) has been derived from anisotropies observed in the cosmic microwave background[85]. These measurements indicate a density of the order of 100 CνB neutrinos plus antineutrinos per cubic centimetre. The corresponding flux depends on the sum of the (unknown) neutrino masses. The flux of CνB is maximal in case the one of the neutrino mass eigenstates is massless. In this case, the CνB flux is about $3.4 \cdot 10^{12}$ neutrinos and antineutrinos per square centimetre per second[84]. At the moment the CνB decoupled from matter, the temperature of the Universe was about 10^{10} Kelvin, or ≈ 1 MeV[86]. However, due to the large redshift, the neutrinos still existing today have a significant lower energy; about 10^{-4} eV in the lightest neutrino mass scenario. The minuscule interaction cross section at these energies makes it very challenging to detect the CνB. An experiment (named PTOLEMY) to undertake this challenging project is currently in the design phase[87].

- **Solar neutrinos**: The Sun produces the second highest flux of neutrinos on Earth, exceeded only by the CνB. Neutrinos are produced in the nuclear fusion chain in the Sun, most notably by the process $p + p \rightarrow d^+ + e^+ + \nu_e$. In this chain, only electron

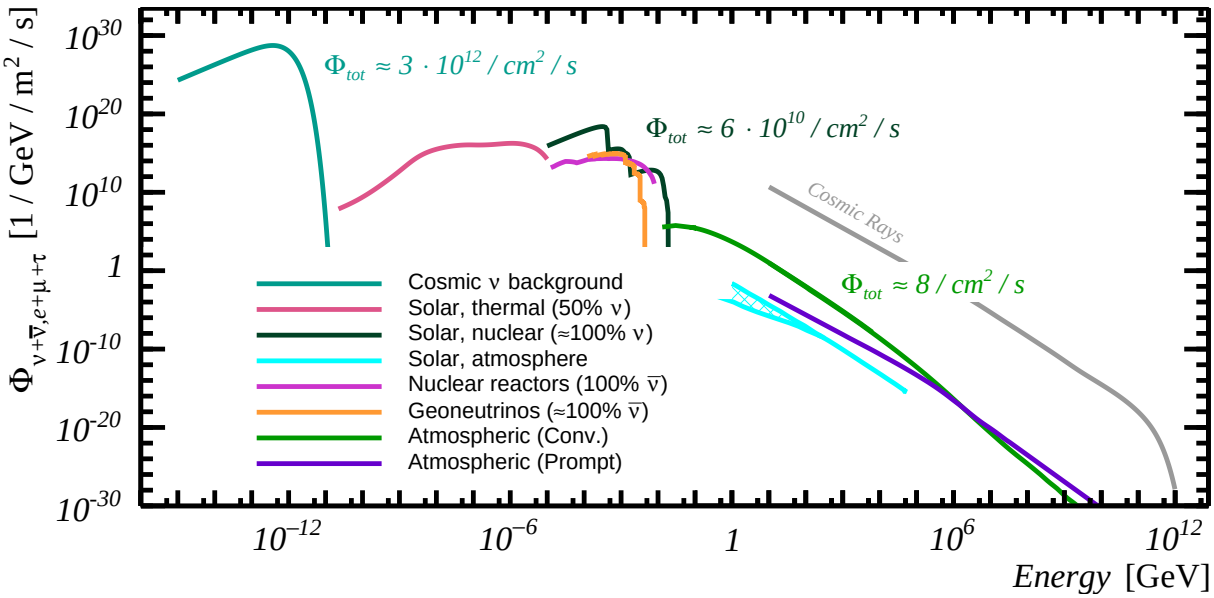


Figure 2.1: *The predicted energy spectrum of neutrinos and antineutrinos, summed over all flavours and integrated over the full sky [84]. For the fluxes from the sun, the considered plane is perpendicular to the line of sight. The quoted Φ_{tot} values correspond to the energy-integrated neutrino flux. Shown (as a reference) is the cosmic ray energy spectrum, as given in figure 1.3.*

neutrinos are produced*, resulting in a flux at Earth of about $6 \cdot 10^{10}$ neutrinos per square centimetre per second. This flux was first detected by the famous 'Homestake' experiment led by Davis[88]. The measured flux turned out to be about one third of the expected neutrino flux. As was shown later, this deficit was the first observation of neutrino oscillations. In addition to neutrinos produced in fusion (nuclear solar neutrinos), (anti)neutrinos are produced in interactions of electrons and photons in the core of the sun [89]. These are called thermal solar neutrinos, constituting the dominant flux of (anti)neutrinos in the eV-keV energy regime.

- **Geoneutrinos**: Decays of radioactive isotopes in the Earth produce a flux of neutrinos with energies in the MeV region. The main contributions are the decays of the potassium-40 isotope ($^{40}K \rightarrow {}^{40}Ca + e^- + \bar{\nu}_e$), the uranium-238 decay chain ($^{238}U \rightarrow {}^{206}Pb + 8\alpha + 6e^- + 6\bar{\nu}_e$) and the thorium-232 decay chain ($^{232}Th \rightarrow {}^{208}Pb + 6\alpha + 4e^- + 4\bar{\nu}_e$) [84]. In these, only electron antineutrinos are produced.

- **Nuclear Reactors**: The most significant human-made sources of neutrinos are nuclear reactors. Antineutrinos produced in the β-decays of fission products in nuclear reactors were the first neutrinos ever detected [64]. On average, the flux of reactor neutrinos at Earth is approximately $4 \cdot 10^7$ $\bar{\nu}$ $/cm^2/s$. The most significant contribution to this flux comes from the splitting of ^{235}U, ^{239}Pu, ^{238}U and ^{241}Pu.

*A negligible flux of electron antineutrinos is expected from the decays of (mostly) ^{40}K, ^{232}Th, and ^{238}U isotopes

- **Atmospheric neutrinos**: Neutrinos produced in cosmic ray air showers in the Earths' atmosphere dominate the neutrino flux above ≈ 20 MeV. As these *atmospheric neutrinos* form the main background in neutrino astronomy, they will be discussed in more detail in the next section. A special subclass of atmospheric neutrinos are those produced in cosmic rays interacting in the atmosphere of the Sun. The flux of these *solar atmospheric neutrinos* is easily distinguishable from the flux of cosmic neutrinos.

2.2 High-Energy Neutrinos

The flux of cosmic neutrinos is expected to exceed the background at high energies. The (background and signal) neutrino sources in this energy regime will be discussed in this section.

2.2.1 Atmospheric Neutrinos

The same interactions that produce neutrinos in the vicinity of cosmic ray sources (see section 1.1) happen in cosmic ray air showers produced by cosmic ray interactions in the Earths upper atmosphere. Decays of charged pions (and kaons) produce neutrinos in a flavour ratio of $\nu_e : \nu_\mu : \nu_\tau \approx 1 : 2 : 0$ (see section 1.1 and equation 1.4). The corresponding neutrino component is called the *conventional atmospheric neutrino* flux. In addition to pions and kaons, heavier mesons containing a charm quark (*charmed mesons*) are produced in cosmic ray air showers. The flux of neutrinos produced in the decays of these is called the *prompt atmospheric neutrino* flux.

The energy spectrum of the conventional atmospheric neutrino flux is related to the energy spectrum of the primary cosmic rays ($dN/dE \propto E^{-2.7}$, see section 1). As a result of energy losses of the pions and kaons during propagation, the spectral index of the muon (anti)neutrino flux is slightly steeper than −2.7. Energy losses of muons before subsequent decay result in an even steeper spectrum for the electron (anti)neutrinos.

The heavier charmed mesons in cosmic ray air showers decay almost instantly, before energy losses start to play a role. Therefore, the corresponding energy spectrum is expected to be harder than the conventional flux, dominating the total atmospheric neutrino flux above energies of $E_\nu \gtrsim 1$ PeV. In figure 2.2 (left), the measured atmospheric neutrino flux is given, as well as a model of the flux given by Honda at al.[60, 90].

The directions of prompt atmospheric neutrinos are isotropically distributed. This is not the case for the conventional atmospheric neutrino flux. The difference is caused by the energy losses of the pions and kaons before they decay. Since vertical down-going particles reach the denser parts of the atmosphere earlier than more horizontal moving ones, the resulting neutrino energy spectrum is steeper. The direction distribution of atmospheric neutrinos is given in figure 2.2 (right).

2.2.2 Cosmogenic Neutrinos

While propagating through the Universe, cosmic rays can interact with photons from the cosmic microwave background. These interactions are significant only for the highest

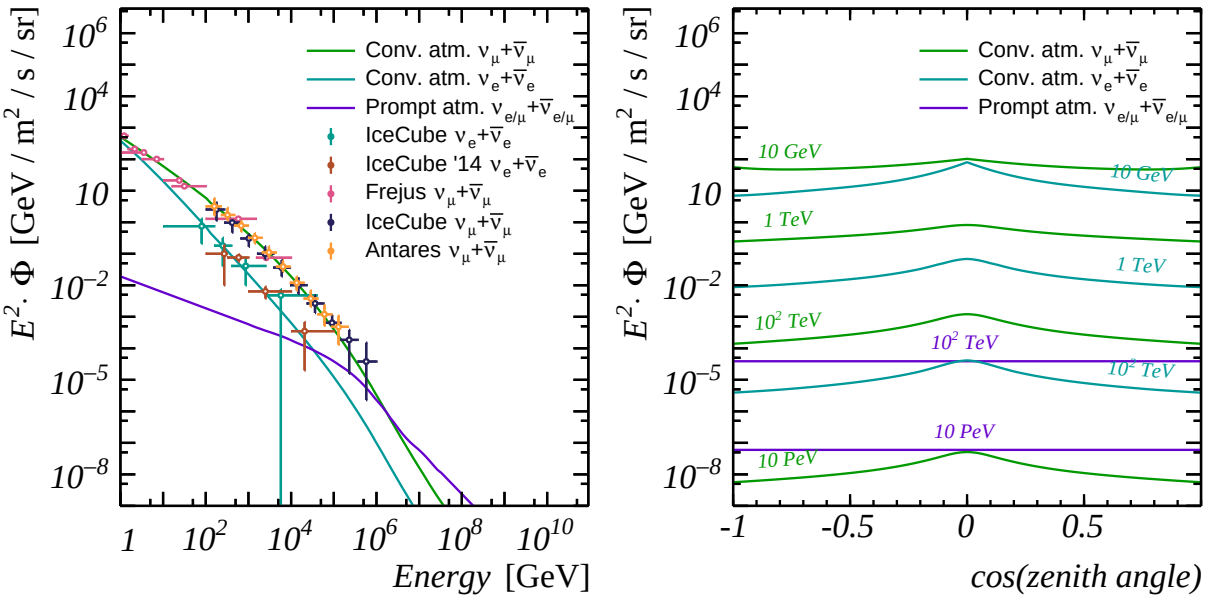


Figure 2.2: ***Left:*** *The energy spectrum of atmospheric neutrinos. The three lines give the flux of the conventional component for electron (anti)neutrinos and muon (anti)neutrinos[60, 90], as well as the prompt component (which is the same for electron and muon flavour neutrinos)[60, 91, 92]. The conventional atmospheric flux measured by the IceCube, Frejus and Antares collaborations are shown[93–97].* ***Right:*** *Same as left figure, but shown as function of the cosine of the neutrino zenith angle.*

energy cosmic rays ($\geq 5 \cdot 10^{10}$ GeV), resulting in the GZK limit (see chapter 1). The main interaction channels are the same as those occurring in the vicinity of cosmic ray sources (see the discussion of the photohadronic scenario in section 1.1). The produced neutrinos are called *cosmogenic neutrinos*, and have a typical energy of about 1 EeV (10^{18}eV). Even the most distant cosmic ray sources contribute to the cosmogenic neutrino flux, thus resulting in an isotropic flux.

2.2.3 Neutrinos from Cosmic Ray Sources

In chapter 1, the characteristics and main production channels of a set of potential cosmic ray sources were discussed. The findings are summarised in table 2.1.

The observed flux of cosmic rays can be used to estimate the maximum neutrino flux from these sources. A rough estimate is given by the Waxman-Bahcall bound.

The Waxman-Bahcall bound In 1998, Waxman and Bahcall derived an upper bound on the flux of neutrinos from cosmic ray sources[98], using the relation between neutrinos and cosmic rays as discussed in section 1.1. While this upper bound was derived only for neutrinos produced in photohadronic interactions, a follow-up paper gives a similar relation for the hadronuclear interaction scenario[99]. The upper bound on the neutrino flux was computed under the following assumptions: 1) The cosmic ray energy spectrum at the source follows $dN/dE \propto E^{-2}$, as expected in first order Fermi acceleration (section 1.3.1). 2) Cosmic rays above 10 EeV are accelerated in homogeneously distributed extra-

Object	Location	$E_{max.}$ Fe^{26+}	Main int.	Time profile
Supernovae (collapse)	Galactic	-	pp	Burst
Supernovae (remnants)	Galactic	832 PeV	pp	Semi-const.
Pulsar wind nebulae	Galactic	832 PeV	pp	Semi-const.
Neutron stars (pulsars)	Galactic	1.6 YeV	pγ	Constant (pulsed)
White dwarfs	Galactic	1.1 ZeV	pγ	Constant
AGN, core	Extra-gal.	9.8 ZeV	pγ	Const.&Flares
AGN, jets (blazars)	Extra-gal.	155 EeV	pγ	Const.&Flares
AGN, hot spots	Extra-gal.	5.0 ZeV	pp&pγ	Const.&Flares
AGN, lobes	Extra-gal.	1.6 ZeV	pp&pγ	Const.&Flares
Colliding gal.	Extra-gal.	8.1 ZeV	pp&pγ	Burst
Gal. clusters	Extra-gal.	14.7 ZeV	pp&pγ	Semi-const.
Starburst gal.	Extra-gal.	1.0 ZeV	pp	Constant
γ-ray bursts	Extra-gal.	116 ZeV	pγ	Burst

Table 2.1: *A list of characteristics of γ-ray and cosmic ray sources. The second column gives the location of the main population observed at Earth. The third column gives the maximum energy of accelerated cosmic rays (iron only), estimated using the Hillas criterion and figure 1.7. The fourth and fifth columns indicate the main interaction channel and the time profile of produced neutrinos.*

galactic sources only. This condition is in agreement with the results discussed in section 1.3. 3) Cosmic rays above 10 EeV are only protons, and 4) A source of high-energy neutrinos is a source of high-energy cosmic rays.

The first two assumptions can be combined with the observed cosmic ray flux to determine the rate of 10 EeV-1 YeV cosmic rays produced in the Universe. Using assumptions three and four, the upper bound on the cosmic neutrino flux is given by (see Waxman[100]):

$$E_\nu{}^2 \cdot \Phi_{\nu+\bar{\nu},e+\mu+\tau} \lesssim 4.5 \cdot 10^{-8} \quad [\text{GeV / cm}^2 \text{ / s / sr}]. \tag{2.1}$$

The third assumption upon which the Waxman-Bahcall limit is based seems to contradict recent measurements (see section 1 and the papers by the Pierre Auger collaboration[40, 41]). A generalisation of the Waxman-Bahcall bound for heavier elements is given by Murase et al.[101] and Anchordoqui et al.[102], and results in a similar upper bound. The fourth assumption is only valid in case the accelerated cosmic rays have a high probability to escape the region around the source. This is, for example, not the case in the core region of active galactic nuclei (a *hidden core*). The neutrino flux from these sources is not constrained by the Waxman-Bahcall limit[103].

Flux Predictions A variety of models exist for the expected flux of high-energy astrophysical neutrinos. In figure 2.3, the expected energy spectra of some benchmark models for neutrino sources are shown. These will be compared with measurements in section 2.3.3. The expected background flux of atmospheric neutrinos (as discussed in section 2.2.1) is given for comparison. The shown atmospheric neutrino flux corresponds to the

background level for an analysis trying the prove the existence of a source in an area of $1° \times 1°$. The horizontal line indicates the Waxman-Bahcall limit. The model of AGN cores is based on the assumption of a hidden core. Therefore, the expected flux is not constrained by the Waxman-Bahcall limit.

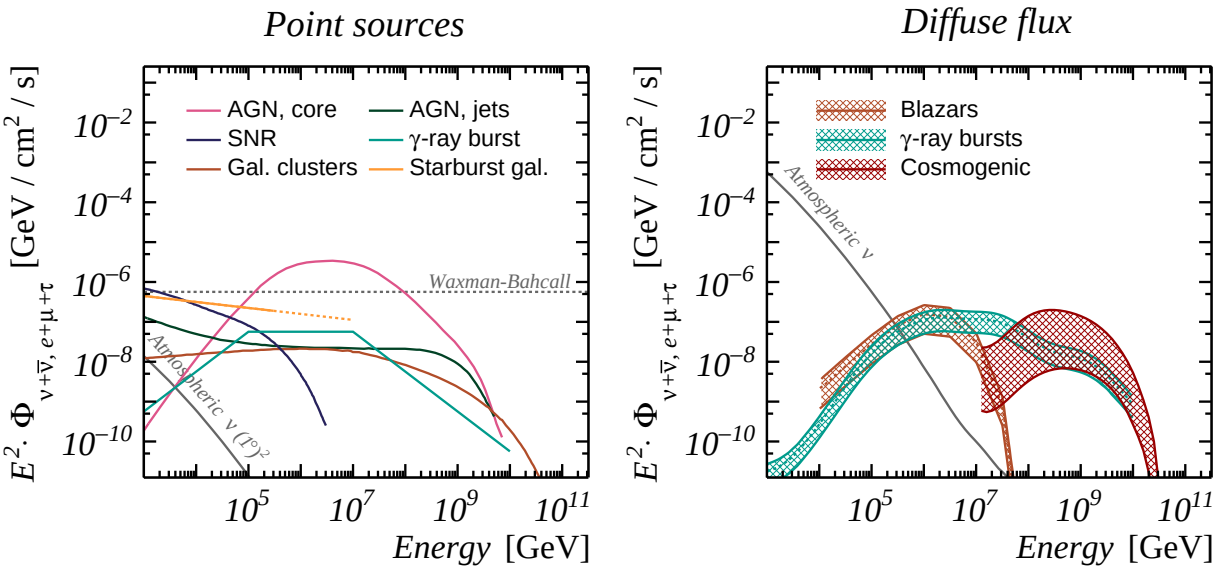


Figure 2.3: *The expected $\nu + \bar{\nu}$ flux from several neutrino sources, summed over the three neutrino flavours (when applicable assuming a $\nu_e : \nu_\mu : \nu_\tau \approx 1 : 1 : 1$ flavour ratio, as discussed in section 2.2.4). Models are given for point sources[103–110] (**left panel**), as well as for diffuse neutrino sources[111–113] (**right panel**). The Waxman-Bahcall upper bound for point sources is indicated by the horizontal line[100]. For the diffuse signal, the atmospheric neutrino background is integrated over the full sky, for point sources it is given for a region in the sky of $1° \times 1°$.*

2.2.4 Neutrino Flavour Ratio

The ratio between the flavours of the (anti)neutrinos produced in cosmic ray interactions is given by $\nu_e : \nu_\mu : \nu_\tau \approx 1 : 2 : 0$ (see section 1.4). In this *pion decay* scenario, the assumption is made that muon energy losses do not play a significant role. This is not necessarily the case in environments with strong magnetic fields, where muons lose energy due to synchrotron radiation. Consequently, the electron (anti)neutrino has a lower energy. This scenario is called the *muon-damped* scenario. In the most extreme cases, only low-energy electron (anti) neutrinos are produced, leading to a $\nu_e : \nu_\mu : \nu_\tau = 0 : 1 : 0$ neutrino flavour ratio at high energies. More realistic is a mixed scenario, in which the flavour ratio depends on the energy of the pions and muons, and is thus related to the energy of the produced neutrinos[114].

A third, almost unrealistic†, scenario (called the *neutron beam* scenario) is based on the assumption of a (extreme) source emitting high-energy neutrons only. These neutrons

†It is only included for sake of completeness

decay, each producing a proton, electron and corresponding electron antineutrino ($n \rightarrow p + e^- + \bar{\nu}_e$), thus resulting in a $\nu_e : \nu_\mu : \nu_\tau = 1 : 0 : 0$ flavour ratio.

A measurement of the neutrino flavour ratio can provide information on the interactions and processes happening close to a cosmic ray source. However, since neutrinos oscillate while travelling through the Universe, the neutrino flavour composition at the Earth is not equal to the flavour ratio at the source(s). The probability $P_{\nu_\alpha \rightarrow \nu_\beta}$ to observe a neutrino of energy E produced in flavour eigenstate α as a neutrino with flavour β after propagating L meters is given by

$$\begin{aligned} P_{\nu_\alpha \rightarrow \nu_\beta}(L/E) = \delta_{\alpha,\beta} - 2 \sum_{k \neq j} \left[\mathrm{Re}\left[U_{\beta,j} \cdot U^*_{\alpha,j} \cdot U^*_{\beta,k} \cdot U_{\alpha,k} \right] \cdot \sin^2 \left(\frac{c^3}{4\hbar} \cdot \frac{\Delta m^2_{j,k} \cdot L}{E} \right) \right] \\ + \sum_{k \neq j} \left[\mathrm{Im}\left[U_{\beta,j} \cdot U^*_{\alpha,j} \cdot U^*_{\beta,k} \cdot U_{\alpha,k} \right] \cdot \sin \left(\frac{c^3}{4\hbar} \cdot \frac{\Delta m^2_{j,k} \cdot L}{E} \right) \right], \end{aligned} \tag{2.2}$$

with U an unitary 3×3 (complex) matrix U, called the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix[115].

Cosmic neutrino sources are typically located $\geq \mathcal{O}(10^{14})$ km away from the Earth. Cosmic neutrinos ($\mathcal{O}(10^6)$ GeV) from these sources cover many periods of the derived oscillation pattern before reaching the Earth. Unless the used neutrino detector has a perfect energy resolution, and the source extends not more than one oscillation period, the oscillation probability 'averages out' over these distances. This is known as the classical limit:

$$\begin{aligned} P_{\nu_\alpha \rightarrow \nu_\beta} &= \delta_{\alpha,\beta} - 2 \sum_{k \neq j} \left[\mathrm{Re}\left[U_{\beta,j} \cdot U^*_{\alpha,j} \cdot U^*_{\beta,k} \cdot U_{\alpha,k} \right] \cdot \frac{1}{2\pi} \cdot \int_0^{2\pi} \sin^2(x)\, dx \right] \\ &\quad + \sum_{k \neq j} \left[\mathrm{Im}\left[U_{\beta,j} \cdot U^*_{\alpha,j} \cdot U^*_{\beta,k} \cdot U_{\alpha,k} \right] \cdot \frac{1}{2\pi} \cdot \int_0^{2\pi} \sin(x)\, dx \right], \end{aligned} \tag{2.3}$$

$$= \delta_{\alpha,\beta} - \sum_{k \neq j} \left[\mathrm{Re}\left[U_{\beta,j} \cdot U^*_{\alpha,j} \cdot U^*_{\beta,k} \cdot U_{\alpha,k} \right] \right], \tag{2.4}$$

$$= \sum_{j=1}^{3} \left[|U_{\alpha,j}|^2 \cdot |U_{\beta,j}|^2 \right], \tag{2.5}$$

which gives, assuming normal neutrino mass ordering and using constraints on the neutrino oscillation parameters obtained by several experiments[116, 117]:

$$\begin{bmatrix} f(\nu_e) \\ f(\nu_\mu) \\ f(\nu_\tau) \end{bmatrix}^{det.} \approx \begin{bmatrix} 0.55 & 0.24 & 0.21 \\ 0.24 & 0.38 & 0.38 \\ 0.21 & 0.38 & 0.41 \end{bmatrix} \cdot \begin{bmatrix} f(\nu_e) \\ f(\nu_\mu) \\ f(\nu_\tau) \end{bmatrix}^{source}, \tag{2.6}$$

where the first vector (with superscript *det.*) gives the detected neutrino flavour ratios and the last vector (with superscript *source*) the flavour ratio at the source. For the discussed scenarios, the expected neutrino flavour ratios at Earth are given in table 2.2.

2.3 Recent Highlights in Neutrino Astronomy

In this section, some highlights in the field of neutrino astronomy will be discussed. These results are categorised according to flux parameter(s) used to differentiate the cosmic

	$\nu_e : \nu_\mu : \nu_\tau$	
Scenario	At source	At Earth
Pion decay	1 : 2 : 0	0.34 : 0.33 : 0.32
Muon damped	0 : 1 : 0	0.24 : 0.38 : 0.38
Neutron beam	1 : 0 : 0	0.55 : 0.24 : 0.21

Table 2.2: *Neutrino flavour ratios at their source(s) and at Earth for three neutrino production scenarios (see text).*

neutrino flux from the atmospheric neutrino background; namely the arrival time, arrival direction, energy spectrum and flavour ratio.

2.3.1 Multi-Messenger Astronomy

For time-varying candidate neutrino sources such as supernovae, gamma ray bursts, blazars and binary system mergers, the time-dependent variations in the neutrino flux can be used to (significantly) reduce the number of background neutrinos. Consequently, a relatively small signal flux can already be sufficient to identify the sources. The background reduction is even more effective in case the arrival direction of the source is known, and/or if the arrival time of the neutrino signal is correlated with the arrival times of cosmic rays, photons and/or gravitational waves from the same source. This field is called *multi-messenger* astronomy.

A textbook example of the identification of an astrophysical neutrino source using the multi-messenger approach is the observation of the supernova SN1987A on 23 February 1987. Within a time period of 13 seconds, twelve and eight neutrinos were detected in the Kamiokande II and Irvine-Michigan-Brookhaven (IMB) detectors[18, 19]. About half a minute later, five additional antineutrinos were detected in the Baksan Neutrino Observatory [20] (the absolute time accuracy of this detector was about one minute). In figure 2.4, the reconstructed energy of each neutrino is given as function of its arrival time. Approximately 2.5 hours after the emission of neutrinos, the first photons escaped from this supernova, and were detected by Shelton and Duhalde, and independently by A. Jones[21]. From the number of detected events, it is estimated that about 10^{58} neutrinos were emitted in this supernova[118].

A second example of a result obtained using a multi-messenger approach is given by the correlated detection on September 22, 2017 of a ≈ 290 TeV muon neutrino (an event named IceCube-170922A)[22], followed by the observation of γ-rays from approximately the same location in the sky. The γ-rays were emitted in a flare of the blazar TXS 0506+056[23]. The source of the detected neutrinos, on the other hand, is uncertain, and coincides with the location of TXS 0506+056 within a confidence level of only 50%[24].

The coincidence of this single neutrino event with the increased activity does not prove that the neutrino was emitted by TXS 0506+056, as it may be a background neutrino produced in the atmosphere or from an unknown source. Nevertheless, this result triggered a follow-up research looking for an excess in the flux of ≥ 100 GeV neutrinos from (approximately) the same region in the sky in 9.5 years of archived data.

An excess of neutrino events was found between October 2014 and February 2015[119].

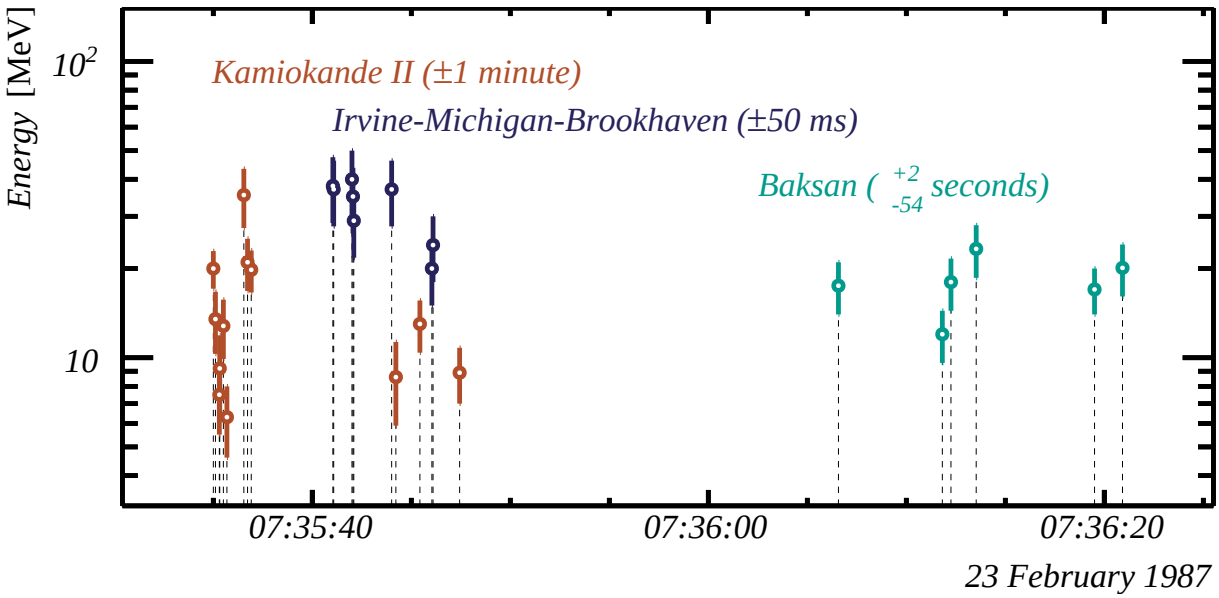


Figure 2.4: *The detection times of observed neutrinos from supernova SN1987A by the Kamiokande II [18], Irvine-Michigan-Brookhaven [19] and Baksan [20] detectors. The vertical axis corresponds to the reconstructed neutrino energy. The uncertainty on the absolute time of each detector is indicated between brackets.*

The possibility that this neutrino flare is caused by a random coincidence from a diffuse (atmospheric and high-energy) flux is excluded by 3.5σ. However, during that period no particular $\gamma - ray$ activity of TXS 0506+056 has been observed[120] (see figure 2.5). Therefore, TXS 0506+056 is an unlikely source of the neutrino flare. To complicate the picture even more, a close-by ($\approx 1.2\ ^\circ$) quasar named PKS 0502+049 showed an increased γ-ray intensity just before and after the neutrino flare[25]. The hypothesis that the neutrino flare is actually caused by PKS 0502+049, rather than TXS 0506+056 is rejected by the IceCube collaboration, with the argument that PKS 0502+049 lies outside of the 90% confidence interval of the origin of IceCube-170922A. Padovani et al.[25] provide an in-depth discussion.

2.3.2 Arrival Direction

Three commonly used ways to identify sources of neutrinos using the distribution of their arrival directions are:

- **Full-sky point source**: Search for a statistically significant clustering of events anywhere in the sky.

- **Catalogue** (alternatively named **stacked source**, or **candidate list**) **searches**: Analysis in which the locations of the neutrino sources are fixed at a list of known cosmic objects (for instance all AGN). Only the neutrino energy spectrum and optionally the flavour ratio at the source(s) are fitted.

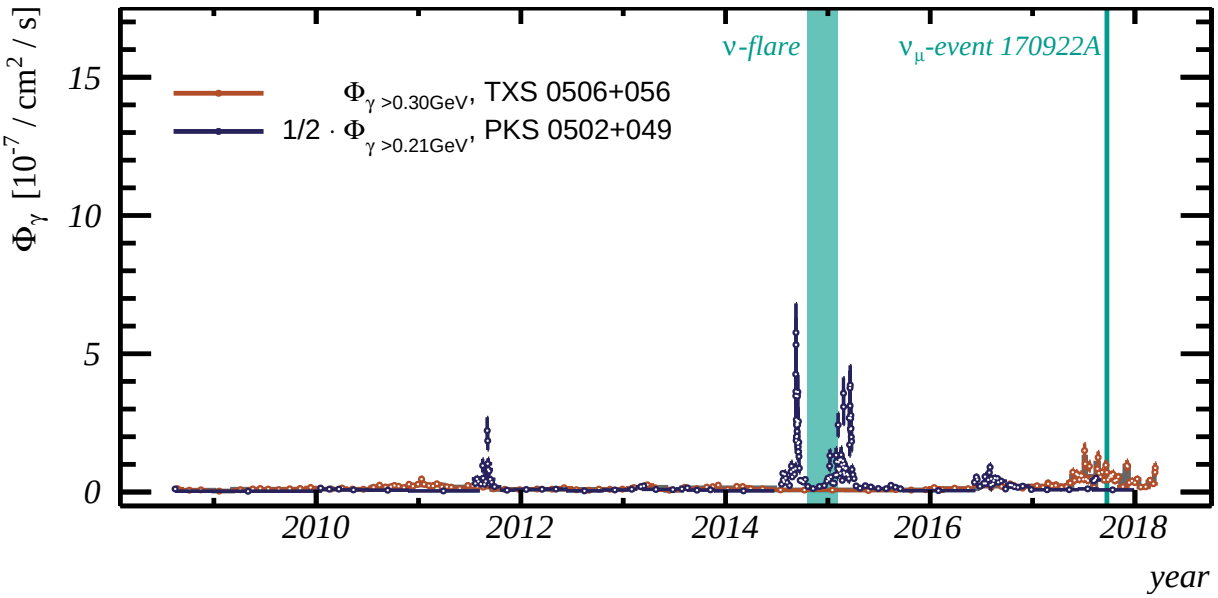


Figure 2.5: *The measured flux of γ-rays >300 MeV from blazar TXS 0506+056 (taken from a paper by the Fermi-LAT, ASAS-SN and IceCube collaborations [120]) and >214 MeV from quasar PKS 0502+049 (from Sahakyan et al.[121]), both using data from Fermi-LAT[122]. Please note that the flux from PKS is scaled by 1/2. Indicated by the vertical line and shaded area are the detection of a ≈ 290 TeV muon neutrino event by the IceCube experiment from a direction close to TXS 0506+056[22] and the time of an increased neutrino flux (corresponding to two times the standard deviation of the 'Gaussian shaped' analysis used in the corresponding paper[119]) from a region surrounding TXS 0506+056.*

- **Region in the sky**: Similar to the catalogue search, but (rather than a set of point sources) a certain region in the sky is chosen, for example the region surrounding the galactic centre[123].

A recent full-sky point source search is discussed in a paper by the IceCube collaboration[124]. The probability of a neutrino source located anywhere in the sky is shown in figure 2.6. No significant clustering of events close to the galactic plane is observed, excluding a galactic origin of the detected neutrino flux [125]. Two clusters of events are found, the most significant one located close to the active galaxy NGC 1068. However, the detected number of events is not sufficient to claim a discovery.

The non-observation of a neutrino source can be used to put constraints on the astrophysical flux from the tested locations in the sky. These constraints are strongly correlated to the number of expected background events, which is generally parameterised as a function of the declination of the assumed source. In figure 2.7, the upper limit on the possible neutrino flux from several assumed sources, the flux needed to get a 50% probability to observe a signal with a 90 % confidence level (the *sensitivity*) or with a 5σ confidence level (the *discovery potential*) are shown for IceCube and Antares‡, for two assumed neutrino

‡The statistical concepts sensitivity, discovery potential and discovery potential will be discussed in more detail in chapter 10

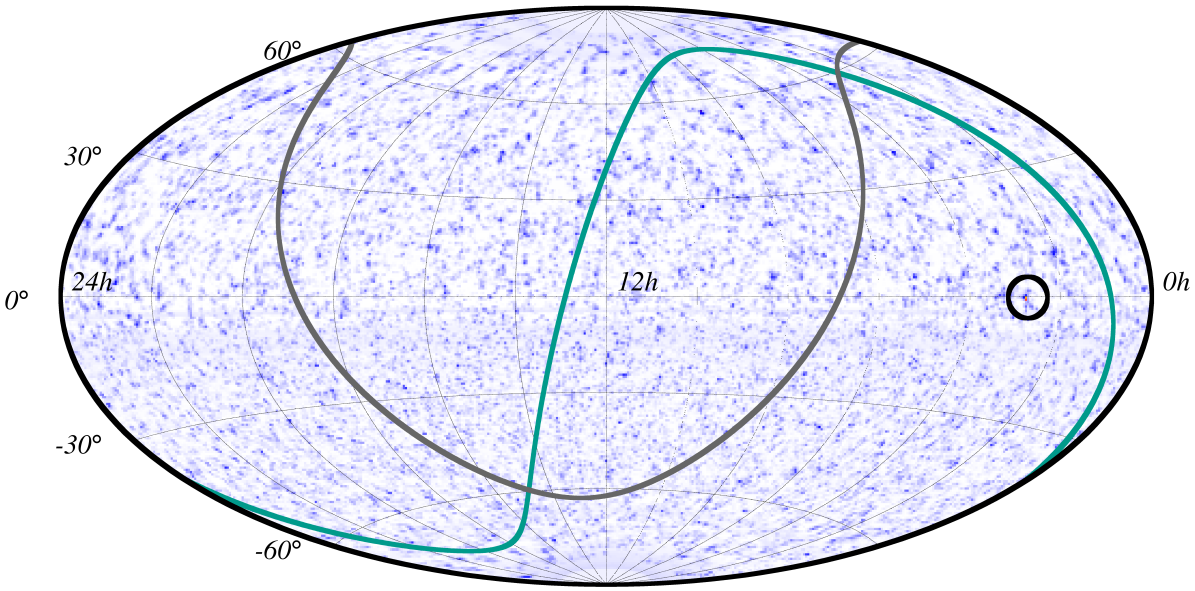


Figure 2.6: *The locations of clusters of neutrino events in ten years of data taken by the IceCube detector[124]. A darker shade indicates a higher probability of the existence of a neutrino source. The location of the most significant cluster of events is indicated by the black circle. In addition, the galactic plane and super-galactic plane are indicated by the grey and blue lines, respectively.*

energy spectra[124, 126].

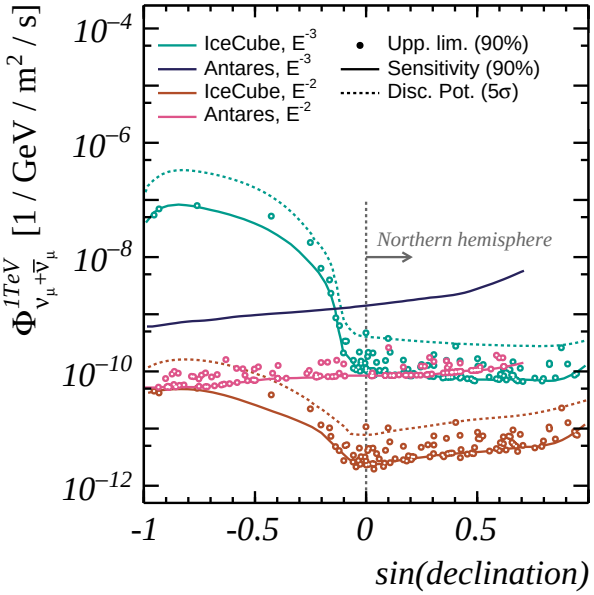


Figure 2.7: *The upper limit, sensitivity and discovery potential (see text for explanation of these terms) for point sources of neutrinos, given as function of the declination of the assumed source. Shown are the results from ten years of IceCube data[124] and eleven years of data taken by the Antares detector[126]. The muon neutrino and antineutrino flux at 1 TeV is given for two typical energy spectra; $dN/dE \propto E^{-2}$ and $dN/dE \propto E^{-3}$.*

The instrumented volume of the ANTARES detector is about 50 times smaller than the IceCube detector. Therefore, the detected number of neutrinos is significantly smaller. As a result, the sensitivity of Antares to neutrino point sources is smaller than IceCube. Nevertheless, as a result of its location on the Northern hemisphere, Antares can still

compete with IceCube for sources in the Southern hemisphere.

2.3.3 Energy Spectrum

The non-observation of a significant cosmic neutrino flux from one (or more) source(s) leaves the possible existence of a diffuse cosmic neutrino flux. A generally considered parameterisation of a diffuse astrophysical neutrino flux component is given by a single power law:

$$\frac{d}{dE_\nu}\left[\Phi_{\nu+\bar{\nu},e+\mu+\tau}\right] \equiv \Phi^{100TeV}_{\nu+\bar{\nu},e+\mu+\tau} \cdot \left(\frac{E_\nu}{[100\ \mathrm{TeV}]}\right)^{-\gamma} \quad [1/\mathrm{GeV}/\mathrm{cm}^2/\mathrm{s}/\mathrm{sr}], \tag{2.7}$$

where $d\Phi_{\nu+\bar{\nu},e+\mu+\tau}/dE_\nu$ gives the all-flavour flux of neutrinos and antineutrinos, $\Phi^{100TeV}_{\nu+\bar{\nu},e+\mu+\tau}$ the flux at 100 TeV and γ the *spectral index*. Relatively recently, such a diffuse flux of high-energy neutrino events exceeding the atmospheric neutrino background has been discovered[127].

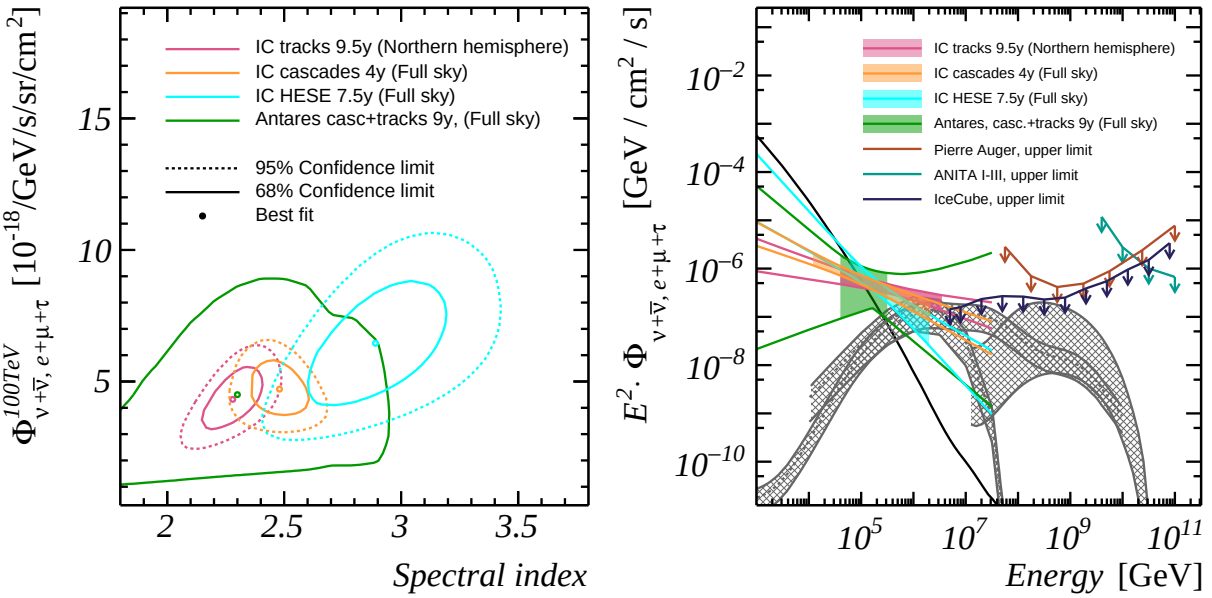


Figure 2.8: ***Left:*** *The best fit 'single power law' energy spectra (see text) of a diffuse astrophysical $\nu + \bar{\nu}$ flux. Shown are the best fit and confidence limits obtained from track (9.5 years)[26], cascade (4 years)[27] and HESE (7.5 years)[28] events detected by the IceCube collaboration, as well as cascades plus tracks (9 years) detected by the Antares experiment[29] (please note that the 95% confidence limit of this study is not shown in the figure). The flux normalisation corresponds to an all-flavour neutrino and antineutrino flux, assuming a neutrino flavour ratio of $\nu_e : \nu_\mu : \nu_\tau \approx 1 : 1 : 1$ at Earth.* ***Right:*** *68% confidence interval of single power law spectra corresponding to the left figure, shown as function of neutrino energy. The shaded areas indicate the energy range of events used in the fit. The grey lines in the background correspond to the diffuse neutrino flux predictions shown in figure 2.3. Upper limits on the diffuse neutrino flux at high energies [128–130] are indicated by solid lines with arrows pointing down.*

The results of a single power law fit to the flux of detected astrophysical neutrinos by the IceCube[26–28] and Antares detectors [29] are summarised in figures 2.8 (left) and 2.8 (right). The left figure shows the best fit and confidence level of the fitted spectrum, the right figure gives the corresponding spectra (at 68% confidence level) and indicates the energy region of events used in the fit. Each of the shown analyses uses a particular event signature in the detector (more details will be given in section 3.3):

- Tracks: Light induced along the path of a (high-energy) muon. Generally, only a fragment of the track (that can range up to tens of kilometres) is observed by the detector, making an accurate estimate of the neutrino energy difficult.
- Cascades: Spherical bursts of light, generally caused by all neutrino interactions other than those leading to tracks. An exception to this are particular signatures caused the interactions of tau (anti)neutrinos, which may be classified as High-Energy Starting Events.
- *High-energy starting event* (HESE): High-energy track- and shower-like events with the interaction position within the instrumented volume are called high-energy starting events events. These events form a small fraction of the total number of events, but the accuracy of the estimated energy is better than the respective resolution of the other categories.

The neutrino with the highest energy ever detected has an estimated energy of 4.5 EeV[131]. The non-observation of events above this energy can be used to constrain the astrophysical neutrino flux at even higher energies. In figure 2.8 (right), the upper limit on the diffuse neutrino flux (90% confidence level) as determined by the IceCube collaboration[128], as well as the upper limits by the Pierre Auger and Anita experiments are given.

In addition to the relatively simple single power-law spectrum, more sophisticated models have been considered in the discussed analyses. None of these other models is *"significantly preferred compared to the single power-law description."*[28].

2.3.4 Flavour Ratio

The event ratio between the three event signatures discussed in the previous section have been used to fit the neutrino flavour ratio. As discussed in section 2.2.4, the observed flavour ratio can be translated back to the flavour ratio at the source, thus providing information on the leading interactions happening in the vicinity of neutrino sources (see section 2.2.4).

A first measurement of the flavour ratio of high-energy neutrino events was performed by the IceCube detector[132]. In figure 2.9, the best fit and confidence levels on the flavour ratio are compared to the flavour ratios corresponding to the scenarios discussed in section 2.2.4. No events compatible with interactions of tau neutrinos have been identified, and the best fit neutrino flavour ratio is $\nu_e : \nu_\mu : \nu_\tau \approx 1 : 1 : 0$. The muon-damped scenario is slightly preferred in comparison with the pion decay scenario, the neutron beam scenario is disfavoured by about 99%[132].

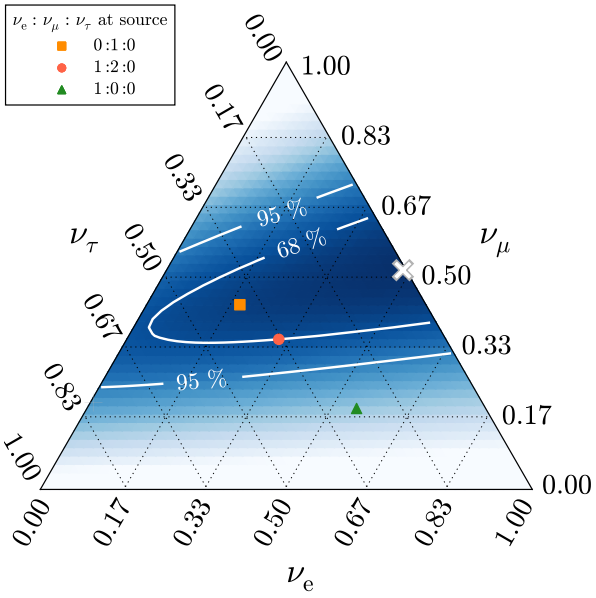


Figure 2.9: *The high-energy cosmic neutrino flavour ratio (at Earth), measured by the IceCube collaboration (figure taken from Aartsen et al.[132]). The white cross indicates the best fit neutrino flavour ratio. The coloured square, circle and triangle represent the predicted flavour ratios after oscillations of three source interaction hypotheses (see section 2.2.4).*

Chapter 3

Neutrino Detection in KM3NeT

Currently, the KM3NeT (*'Cubic Kilometre Neutrino Telescope'*) neutrino research infrastructure is under construction on the seabed of the deep waters of the Mediterranean Sea. In this chapter, the scientific goals, detection principle, detector outline, backgrounds and data acquisition systems of the KM3NeT detectors will be presented.

3.1 Research Goals of the KM3NeT Project

KM3NeT consists of two detectors, with a foreseen total instrumented volume of approximately 1.1 km^3. These detectors, called KM3NeT/ARCA and KM3NeT/ORCA, are deployed off-shore Sicily (Italy) and Toulon (France), respectively.

KM3NeT/ARCA The KM3NeT/ARCA (*'Astroparticle Research with Cosmics in the Abyss'*) detector aims at the detection of high-energy ($\geq$ 1 TeV) astrophysical neutrinos and to study their sources. The origins of detected neutrinos are expected to be resolved with an unprecedented angular resolution (discussed in more detail in chapter 5).

KM3NeT/ORCA With the observation of neutrino oscillations [88, 133, 134], neutrinos are proven to have mass. However, no mechanism is present in the standard model to explain non-zero neutrino masses, and fundamental properties like the ordering of the neutrino mass eigenstates are still unknown.

In parallel with KM3NeT/ARCA, a second neutrino detector is being built by the KM3NeT collaboration. This detector, named KM3NeT/ORCA (*'Oscillation Research with Cosmics in the Abyss'*) is foreseen to determine the neutrino mass ordering within a few years after completion[135]. By measuring the flux of atmospheric neutrinos propagating through the Earth, a large fraction of the neutrino oscillation pattern will be probed. Studies have shown that neutrinos passing through the core of the Earth with energies $\lesssim$ 20 GeV provide the most sensitivity[136].

Additional Research with KM3NeT In addition to neutrino astronomy and the study of neutrino properties, the detectors of KM3NeT will be used to study a variety of other topics. These include cosmic ray physics, neutrinos produced in dark matter interactions

in the galactic disc, Earth tomography with neutrinos, oceanography and exotic particles like magnetic monopoles[135].

One particularly interesting opportunity to enhance the science potential of KM3NeT/ORCA is to detect neutrinos from a beam pointed towards it. With this controlled neutrino source at a precisely known distance, KM3NeT/ORCA is sensitive to *charge-parity* (CP) symmetry breaking in the neutrino sector. The potential of sending such a beam from the Protvino accelerator complex in Russia to KM3NeT/ORCA is currently under investigation[137].

3.2 Neutrino Detection Principle

3.2.1 Neutrino Interactions

Neutrinos interact with matter through the weak force. Based on the gauge boson exchanged, one can differentiate between *charged current* (CC) and *neutral current* (NC) interactions. A special case is given by the interaction of an electron antineutrino with an electron to a lepton plus antineutrino or hadrons through the resonance production of a W^- boson. This process is called the *Glashow resonance*. In table 3.1, the Feynman diagrams of all neutrino-matter interactions are given.

3.2.2 Neutrino Interaction Cross-Sections

(Anti)neutrino-nucleon interactions The *deep inelastic scattering* interaction cross section of (anti)neutrinos with nuclei is generally calculated using a set of functions describing the *Quantum chromodynamic* (QCD) interactions between the quarks and gluons within a nucleon. These functions are called *parton distribution functions*, and give the interaction probability as a function of two parameters related to the kinematics of the interaction (see appendix A). The parton distribution functions used in this work are obtained from a fit to experimental data[138–143], compiled by the CTEQ collaboration[144, 145].

The total (anti)neutrino-nucleon cross sections are given in figure 3.1. The cross sections are shown for charged and neutral current interactions, and for neutrinos and antineutrinos. Each of these consists of three lines with the same style; one for each neutrino flavour. For neutral current interactions, the cross sections are identical for all neutrino flavours, as in these interactions no lepton is produced.

For charged current interactions, the difference in mass of the produced lepton causes differences in the cross sections up to approximately 10 % at low energies ($E_\nu \leq 100$ GeV).

(Anti)neutrino-electron interactions To first order, the interaction cross section of (anti)neutrinos with electrons can be calculated analytically, using only fundamental physical constants as input (see for example [146]). In figure 3.1, the total cross section is given for all flavours of both neutrinos and antineutrinos. The neutrino-electron interaction cross section is about 100 times smaller than the neutrino-nucleon cross section (with exception of the Glashow resonance).

Focusing (for now) only on the interactions of (anti)neutrinos with electrons, a few things can be noted. For muon and tau antineutrinos, only the neutral current channel

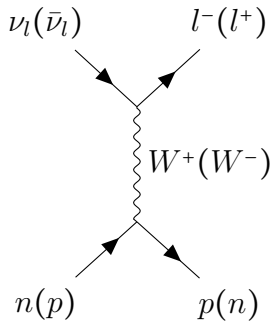
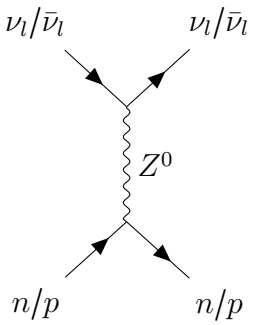
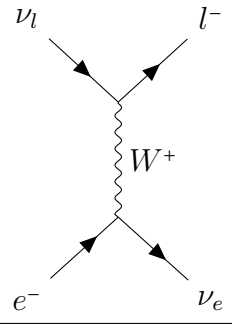
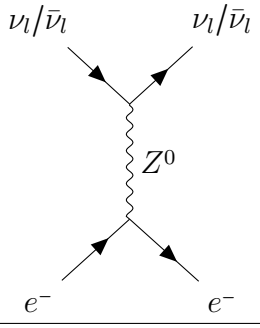
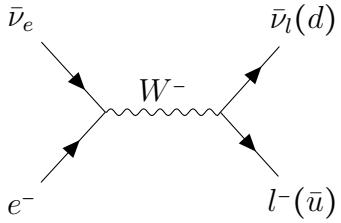
	Charged Current	Neutral Current	Glashow Resonance
Neutrino-Nucleon			
Neutrino-Electron			

Table 3.1: *Feynman diagrams of neutrino interaction channels with nuclei (top) and electrons (bottom). Named after the charge of the exchanged boson, these can be categorised into charged- (left) and neutral-current (middle) interactions. A special case is given by the interaction of an electron antineutrino with an electron via the formation of a W boson, called the Glashow resonance (bottom right). The meson produced in the Glashow resonance and (at high energies) the target nucleon in neutrino-nucleon interactions fragments into a ensemble of secondary particles (not shown in this overview).*

contributes (see table 3.1). Since this channel is kinematically identical for all neutrino flavours, the corresponding cross sections are equal. The cross section of electron antineutrinos with electrons on the other hand is significantly higher, as a result of the additional contribution from the Glashow resonance interaction channel[147]. For neutrino-electron interactions, the charged current channel increases the total cross section in case the energy of the neutrino is sufficient to produce the corresponding lepton, most prominently in the cross section of tau antineutrinos, above energies of approximately 3 TeV.

3.2.3 Expected Neutrino Interaction Flux

The expected flux of neutrino interactions in the vicinity of the detector (the number of neutrino interactions per $[s \cdot GeV \cdot sr \cdot m^3]$) $\frac{d^6\Phi_\nu^{int.}}{dE_\nu d^2\Omega d^3\vec{x}}$ for a neutrino energy E_ν, direction $\vec{d}_\nu$ and interaction point $\vec{x}_\nu$ is given by the product of the expected neutrino flux at the Earths' atmosphere $\frac{d^3\Phi_\nu^{atm.}}{dE_\nu d^2\Omega}$ (per $[s \cdot GeV \cdot sr \cdot m^2]$), the probability that the neutrino traverses the Earth $P_\oplus$ without interacting, the matter density ρ at the interaction point, Avogadro's number N_A, and the total interaction cross-section σ:

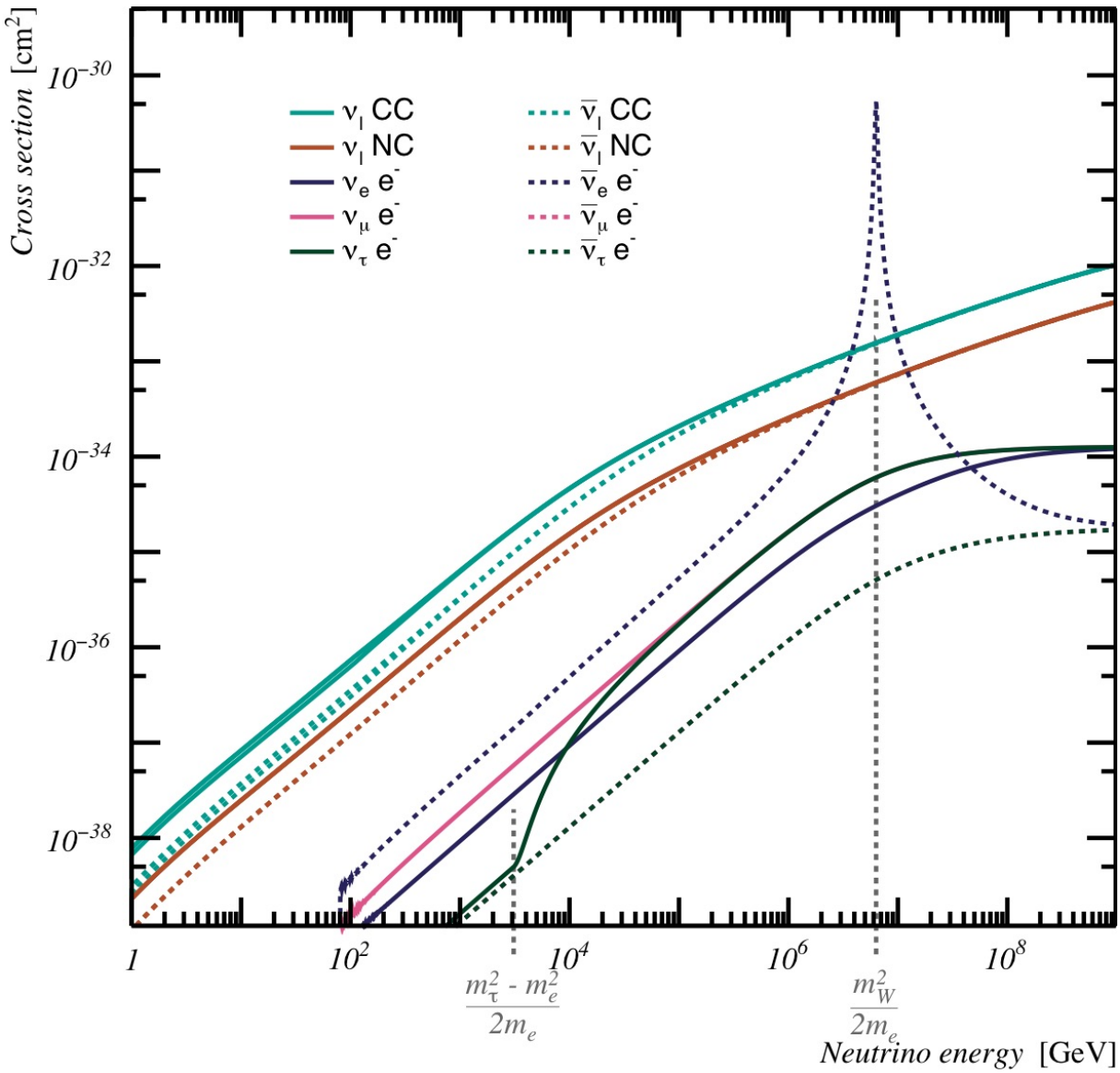


Figure 3.1: *The neutrino-nucleon and neutrino-electron interaction cross sections as function of the neutrino energy. The cross sections of antineutrinos are given by dashed lines. The neutrino-nucleon interaction cross sections (labelled CC and NC) are given by three lines, each corresponding to one neutrino flavour. In the case of neutral current interactions, these lines overlap (see text for the explanation). For neutrino-electron interactions, all contributing channels are added, to give the total cross section. The characteristic resonance from the electron antineutrino interacting with an electron (the Glashow resonance) peaks at 6.3 PeV, indicated by the dashed vertical line. The dashed line at 3 TeV corresponds to the threshold energy for the production of a tau lepton in the charged current interaction of a tau neutrino with an electron. The interaction cross sections of neutrinos with electrons are calculated using the equations given in [146]. See appendix A for the used formulae to calculate the neutrino-nucleon interaction cross sections.*

$$\frac{d^6\Phi_\nu^{int.}}{dE_\nu d^2\Omega d^3\vec{x}}(E_\nu, \hat{d}_\nu, \vec{x}_\nu) = \sigma(E_\nu) \cdot N_A \cdot \rho(\vec{x}_\nu) \cdot P_\oplus(E_\nu, \hat{d}_\nu) \cdot \frac{d^3\Phi_\nu^{atm.}}{dE_\nu d^2\Omega}(E_\nu, \hat{d}_\nu), \tag{3.1}$$

Some numbers The *mean free path* of a 1 TeV neutrino ($\sigma \approx 10^{-35}$ cm^2) through water (with a nucleon density of $N_A{\cdot}\rho \approx 6{\cdot}10^{23}$ nucleons/cm^3) is of the order of $1/\left(6 \cdot 10^{23} \cdot 10^{-35}\right) \approx 6{\cdot}10^{12}$ centimetres (or about half the distance between the Earth and the Sun). This means that only one out of approximately 60 million (1 TeV) neutrinos traversing a kilometre of water undergo an interaction. This back-of-the-envelope calculation shows 1) why neutrino detection is so challenging and 2) why the instrumented volumes of neutrino telescopes like KM3NeT/ARCA are so large.

Using a model of the Earth's density profile (see appendix F) the column density of the Earth can be estimated at approximately $7{\cdot}10^{33}$ nucleons /cm^2 for neutrinos passing right through the centre of the Earth. The corresponding probability that a 1 TeV neutrino traverses the Earth without interacting is $\exp\left(-\sigma \cdot 7 \cdot 10^{33}\right)$ ≈94%. At these (and lower) energies, the Earth is practically transparent for neutrinos.

At the Glashow resonance energy of 6.3 PeV, the electron antineutrino cross section is about 10^{-30} cm^2, which gives a significantly higher number of interacting antineutrinos per kilometre of traversed water (one in about 300). If the flux of these 'Glashow resonance' antineutrinos is significant, these should give a clearly identifiable signature in neutrino detectors. At these energies, the probability that a Glashow resonance antineutrino traversed the Earth is about zero ($\approx \exp(-6 \cdot 10^3)$), so only Glashow resonance antineutrinos from a source located above the horizon can be detected.

The expected flux of neutrino interactions will be discussed in further detail in section 11.1.

3.2.4 Cherenkov Radiation

When a charged particle traverses matter, it polarises the atoms of the medium it is propagating through. In the case that the charged particle travels at a speed $v = \beta \cdot c$ larger than the (phase) velocity of light in that medium, constructive interference of the resulting electromagnetic waves occurs. The resulting radiation is called Cherenkov radiation [148]. This process is similar to the 'sonic boom' observed in sound waves of objects moving faster than the speed of sound in a medium. The constructive interference condition results in a specific angle θ_C under which the light is emitted [149]:

$$\cos(\theta_C) = \frac{1}{\beta n} \tag{3.2}$$

where n is the phase index of refraction of the medium. The angle θ_C is called the Cherenkov angle. For a relativistic particle ($\beta \approx 1$) travelling in seawater ($n \approx 1.35$), it is approximately 42 ° .

Per metre of traversed matter dz, the number of Cherenkov photons (per unit wavelength $d\lambda$) induced by a particle of charge q is given by the Frank-Tamm equation [150]:

$$\frac{d^2N_C^\gamma}{dzd\lambda} = \frac{2\pi\alpha \cdot q^2}{\lambda^2}\left(1 - \frac{1}{(\beta \cdot n)^2}\right) \tag{3.3}$$

where α is the fine structure constant. A relativistic muon ($q = -1, \beta \approx 1$) propagating through seawater will emit approximately $3.7 \cdot 10^4$ photons per metre in the wavelength interval in which seawater is transparent (300 – 650 nm).

Cherenkov radiation is only produced in case $\beta \geq 1/n$, which gives a threshold energy for the emission of Cherenkov radiation. For a muon in seawater, the minimal energy for Cherenkov radiation is approximately 158 MeV. For electrons this *Cherenkov threshold* is significantly lower (0.76 MeV).

3.2.5 Light Emission in Seawater

Charged particles produced in neutrino interactions with sufficient energy cause Cherenkov radiation (see 3.2.4). In this section, these secondary particles and their respective distributions of emitted light will be discussed.

Electromagnetic Showers An electron moving through seawater deflects under the influence of the electric field of seawater nuclei. Consequently, a photon is emitted. This process is called *bremsstrahlung*. In general, the emitted photon will create an electron-position pair (*pair production*), which in turn results in more bremsstrahlung photons. The collection of photons, electrons and positrons is called an *electromagnetic shower*. All charged secondary particles with energies above the Cherenkov threshold induce Cherenkov radiation.

The average energy deposition of an electromagnetic shower along its principal axis is shown in figure 3.2 (left). It can be parameterised reasonably well by a Gamma distribution [151]:

$$-\frac{dE_S}{dz} = E_S \cdot \frac{z^{a-1} \cdot e^{-z/b}}{b^a \cdot \Gamma(a)}, \tag{3.4}$$

$$\begin{aligned} a &= 1.85 + 0.62 \cdot \ln\left(\frac{E_S}{[\text{GeV}]}\right), \\ b &= 0.54 \text{ [m]}, \end{aligned} \tag{3.5}$$

where E_S is the energy of the initial particle. It will be referred to as the *shower energy*. For shower energies $E_S \gtrsim 100$ MeV, the total number of emitted photons is approximately proportional to the shower energy (about $1.5 \cdot 10^5$ photons in the 300 – 650 nm wavelength range per GeV shower energy)[152]. Because of this linear relation and since the spectrum of the emitted light follows the Frank-Tamm equation (equation 3.3), it is convenient to introduce an *effective track length*[152]

$$L_S^{eff.} \equiv 4.0\left[\frac{\text{m}}{\text{GeV}}\right], \tag{3.6}$$

with which the number of emitted photons (per metre and per unit wavelength) from an electromagnetic shower with energy E_S can be parameterised as:

$$\frac{d^2N_S^\gamma}{dzd\lambda} = \frac{d^2N_C^\gamma}{dzd\lambda} \cdot L_S^{eff.} \cdot \left|\frac{dE_S}{dz}\right|. \tag{3.7}$$

As a result of the kinematics in the shower, not all shower particles propagate along the shower axis. Consequently, the distribution of the emission angle of light with respect to the shower axis is smeared around the Cherenkov angle[152]:

$$\frac{d^2P}{d\cos\theta_{\text{emit}}d\phi_{\text{emit}}} \propto \exp\left(-5.40 \times |\cos(\theta_{\text{emit}}) - \cos(\theta_C)|^{0.35}\right), \tag{3.8}$$

where θ_{emit} denotes the angle between the direction of the emitted light and the shower axis. This parameterisation is shown in figure 3.2 (right).

Hadronic Showers Similar to the way electromagnetic showers are produced, hadrons cause *hadronic showers*, consisting mainly of protons, neutrons and pions. Neutral pions are produced mainly in interactions of charged pions, and decay almost instantly to two photons, causing electromagnetic sub-showers. At high energies, the electromagnetic part of the hadronic shower dominates the total energy deposition[153]. At low energies, charged pions generally decay before interacting, mainly to an (anti)muon and muon (anti)neutrino. Since neutrons and neutrinos produced in a hadronic shower do not emit light, a smaller fraction of the total shower energy is visible.

The interaction length of a hadronic shower is larger than the interaction length of an electromagnetic shower. Consequently, a hadronic shower generally takes longer to develop. This effect is particularly present at low shower energies. A parameterisation of the longitudinal shower profile of a hadronic shower in seawater is given by equation 3.4, with (Niess[154]):

$$\begin{aligned} a &= 4.26 + 0.364 \cdot \ln\left(\frac{E_S}{[\text{GeV}]}\right), \\ b &= 0.56 \text{ [m]}. \end{aligned} \tag{3.9}$$

This parameterisation is shown in figure 3.2 (left). The angular emission spectrum of a hadronic shower is comparable to the signature of its electromagnetic counterpart [153].

Muons Muons propagating through seawater lose energy due to several processes. In addition to Cherenkov radiation, pair production and bremsstrahlung as discussed before, these processes are:

- *Ionisation*: The traversing muon knocks an electron out of a water atom.
- *Delta rays*: Similar to ionisation, except that sufficient energy is transferred to the freed electron to cause secondary ionisation of the medium.
- *Photonuclear interactions*: Inelastic scattering of the muon with a nucleus. If the transferred momentum is sufficient, the nucleus can fragment into a hadronic shower.

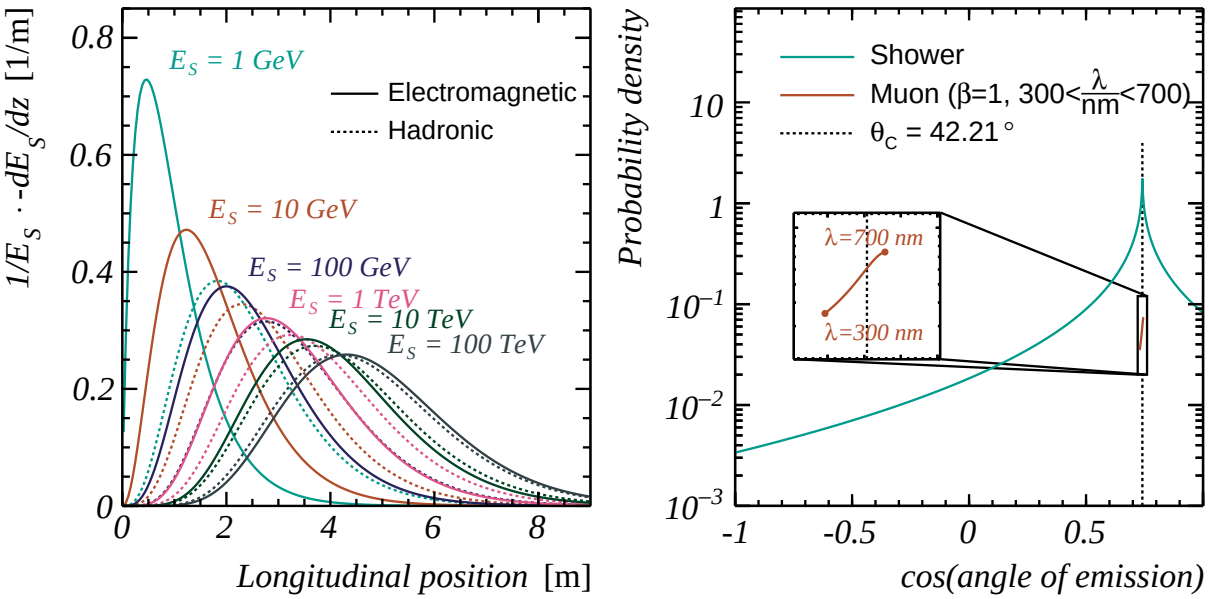


Figure 3.2: ***Left:*** *The energy loss of electromagnetic showers and hadronic showers of different energies along the propagation axis. Given are the parameterisation by Copper [151] and Niess[154].* ***Right:*** *The distribution of the angle between the light emitted by an electromagnetic shower and the shower axis, adapted from the figure by C. Kopper[151]. This distribution is approximately independent of the shower energy. For reference, the (wavelength dependent) Cherenkov emission angle of a relativistic muon ($\beta = 1$) in sea-water is shown.*

The stochastic energy losses due to pair production and bremsstrahlung and photonuclear interactions can be observed as electromagnetic showers along the muon trajectory. Light is also emitted as a result of delta rays. However, since a relatively large number of delta rays are freed per unit length, each having a relatively low energy, the angular light emission profile of delta rays can be approximated by an isotropic distribution [155]:

$$\frac{d^2P}{d\cos\theta_{\text{emit}}\, d\phi_{\text{emit}}} \approx \frac{1}{4\pi}. \tag{3.10}$$

The energy loss of a muon due to delta rays (in [GeV / m]) can be calculated using the equations given by Zyla et. al [4]. This function can be parameterised as [155]:

$$-\frac{dE_\delta}{dz} \approx \rho \cdot \begin{cases} a \cdot \log(E_\mu)^3 + b \cdot \log(E_\mu)^2 + c \cdot \log(E_\mu) + d & \text{if } E_\mu \geq 0.13078 \text{ [GeV]} \\ 0 & \text{else} \end{cases}, \tag{3.11}$$

where ρ is the density (in [gr / m^3]) of the medium the muon propagates through and the constants a, b, c and d (each in [GeV $\cdot$ m^2 / gr]) given by:

$$\begin{aligned} a &= 1.606 \cdot 10^{-10} \\ b &= -2.723 \cdot 10^{-9} \\ c &= 3.369 \cdot 10^{-8} \\ d &= 3.655 \cdot 10^{-8}. \end{aligned} \tag{3.12}$$

In figure 3.3 (left), the average energy loss due to the different processes of a muon propagating through seawater is shown. The summed average energy loss per unit length can be approximated by (Klimushin et al.[156]):

$$-\frac{dE_\mu}{dz} = A + B \cdot E_\mu(z), \tag{3.13}$$

$$\begin{aligned} A &= 2.77 \cdot 10^{-1} \text{ [GeV / m]}, \\ B &= 3.53 \cdot 10^{-8} \text{ [m}^{-1}\text{]}. \end{aligned} \tag{3.14}$$

This differential equation gives the relation between the muon energy $E_\mu(z)$ after propagating z metres and the initial muon energy $E_{\mu,0}$ (at $z = 0$):

$$E_\mu(z) = \left(\frac{A}{B} + E_{\mu,0}\right) \cdot e^{-B \cdot z} - \frac{A}{B}, \tag{3.15}$$

Solving $E_\mu(z) = 0$ gives the average range of a muon in seawater:

$$L_\mu \approx \frac{1}{B} \cdot \log\left(1 + \frac{B}{A} \cdot E_{\mu,0}\right). \tag{3.16}$$

In figure 3.3 (right), this parameterisation is shown, as well as the results of simulations including stochastic energy losses (see section 4.4.1 for details). A muon of 100 GeV will propagate about 300 m, while muons with energies above 10 TeV can propagate for tens of kilometres before losing all of their energy.

Tau leptons The mean lifetime of a tau ($2.903 \cdot 10^{-13}$ s) is about 7 orders of magnitude shorter than the lifetime of a muon. Therefore, the tau decays before its energy drops below the Cherenkov threshold. The mean lifetime of a tau results in a flight length of:

$$\begin{aligned} L_\tau &\approx c \cdot \frac{E_\tau}{m_\tau \cdot c^2} \cdot 2.903 \cdot 10^{-13}, \\ &\approx 4.9 \text{ [m]} \cdot \frac{E_\tau}{100 \text{ [TeV]}}. \end{aligned} \tag{3.17}$$

A tau lepton decays to a virtual W-boson and a tau neutrino. Subsequently, the virtual W-boson decays either to an electron or muon and corresponding antineutrino or to hadrons (mostly pions). The main branching ratios are [4]:

$$\begin{aligned} \tau^- &\to \nu_\tau + \text{hadrons} && (64.79\ \%), \\ \tau^- &\to \nu_\tau + e^- + \bar{\nu}_e && (17.83\ \%), \\ \tau^- &\to \nu_\tau + \mu^- + \bar{\nu}_e && (17.41\ \%). \end{aligned}$$

The electron or hadrons produced in the decay of a tau create a subsequent electromagnetic or hadronic shower.

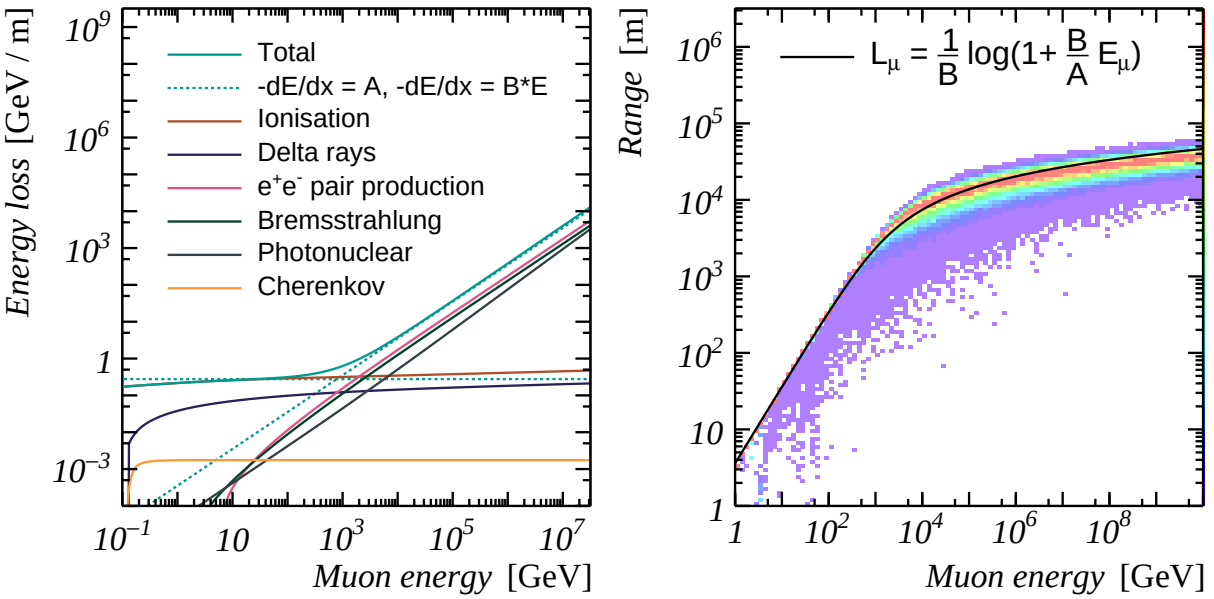


Figure 3.3: ***Left:*** *The average energy loss of a muon in seawater. The contributions of the processes discussed in the text are shown using the parameterisations given in section 3.2.4 and by Kooijman and De Jong[155]. The total muon energy loss is well approximated by an energy-independent component (related to ionisation) plus a component proportional to the muon energy (for radiative processes).* ***Right:*** *The muon range, obtained from simulations (including stochastic radiative energy losses) and calculated using the parameterisation discussed in the text. Each vertical slice of the simulated distribution is scaled to have the same maximum value.*

3.3 Neutrino Interaction Event Signatures

Distinguishing between the distribution of the emitted light, the neutrino interactions can be categorised in several event 'signatures' (see figure 3.4):

- **Shower**: A hadronic shower, either with or without associated electromagnetic shower.
- **Track**: A (high-energy) muon.
- **Double bang***: A hadronic shower with an emerging tau. The tau can propagate a few metres before decaying into a second hadronic or electromagnetic shower.
- **Muonic tau**: Same as double bang, but the tau decays to a muon and muon (anti)neutrino.
- **Lollipop**: Same as double bang, but the shower at the interaction vertex is not observed by the detector.

As a result of the kinematics of the neutrino interaction and stochastic energy losses, not every interaction type will result in the same event signature. For example, if a large

*A term introduced by Learned and Pakvasa[157]

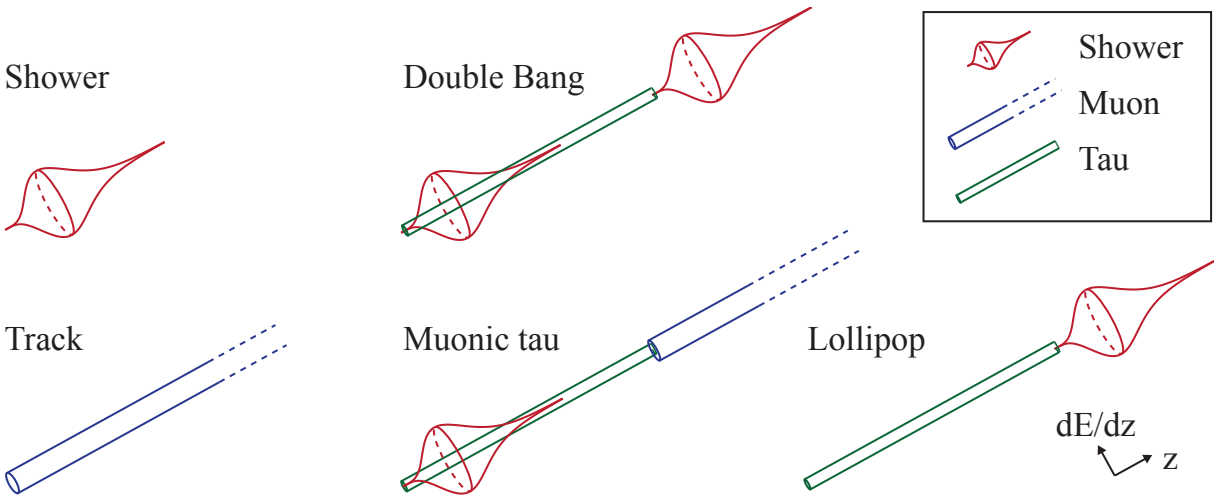


Figure 3.4: *A schematic overview of different neutrino-induced event signatures. The direction of the neutrino is indicated by the z-arrow. The radius of the different components gives an impression of the energy loss (and consequently the amount of emitted light) along the z-axis. Please see the text for a description of the presented signatures.*

fraction of the energy of a muon neutrino in a charged current interaction is transferred to the nucleus, the outgoing muon will have a low energy, and the event signature will be more similar to a shower than to a track.

3.4 Light Propagation through Seawater

While propagating through the seawater, the emitted photons scatter and get absorbed, effectively blurring and fading the detected event signatures. The (wavelength dependent) absorption and scattering lengths have been measured *in-situ* [158, 159] and are shown in figure 3.5. The model used to describe scattering (named *p00075*) consists of two components: the scattering of light on small particles such as dust (Henyey-Greenstein or Mie scattering, [160]) and on water molecules (Rayleigh scattering, [161]). The refractive index in water is given in figure 3.5. The phase index of refraction relates to the emission angle of Cherenkov radiation. The group index of refraction is inversely proportional to the group velocity of light: the speed at which photons move through the water. The distributions for the refractive indices are taken from Schuster et al.[159].

3.5 The KM3NeT Detectors

The KM3NeT/ARCA and KM3NeT/ORCA detectors each instrument a volume of seawater with a three-dimensional grid of sensitive light sensors, called *photomultiplier tubes* (PMTs). To withstand the pressure of the water overburden, 31 PMTs and electronics are housed in 42 cm diameter glass spheres called *digital optical modules* (DOMs, see figure 3.6) [162]. Eighteen DOMs are attached to two parallel vertical strings, held on the seabed by a dead weight attached to an anchor and kept upright by a buoy at the top.

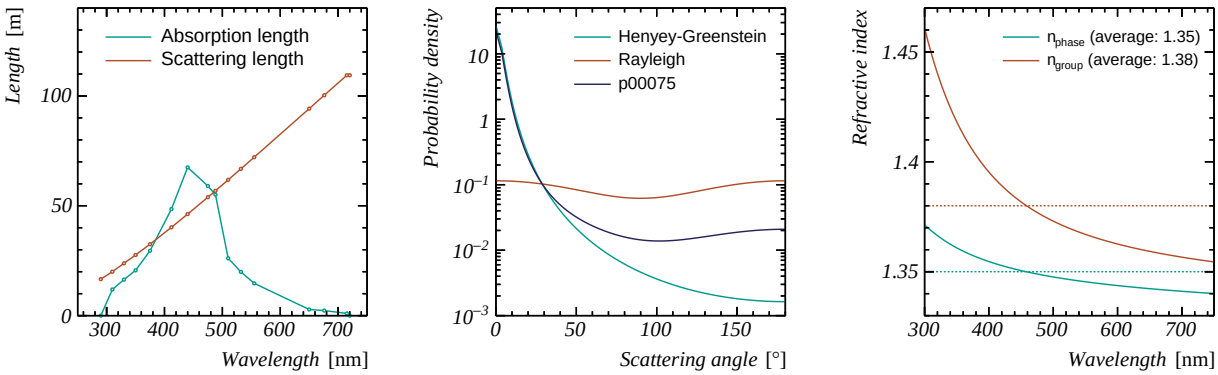


Figure 3.5: *Some optical properties of the seawater at the KM3NeT sites. Shown are the absorption and scattering length (**left**), the p00075 scattering probability model (**centre**) and group and phase refractive indices (**right**).*

Such a set is called a *detection unit* (DU). Per *building block*, 115 DUs are placed within a circular area on the seabed.

In order to differentiate between the different detection units, DOMs and PMTs, the following naming scheme will be used throughout this thesis:

- A number between 1 and 115 is assigned to each detection unit. The first deployed DUs of KM3NeT/ARCA are called *ARCA DU1* and *ARCA DU2*, the first KM3NeT/ORCA DU is called *ORCA DU1*.

- Based on the vertical positions, the DOMs are grouped in *floors*. Floor one corresponds to the lowest DOMs, floor eighteen to the top DOMs.

- The 31 PMTs within a DOM are ordered in rings and numbers. The ring is indicated by a letter, where ring F corresponds to the six uppermost PMTs in a DOM and ring A to the PMT looking straight down. The PMTs within a ring are numbered from one to six (except for ring A, with a single PMT labelled A1).

The DOM spacing of the KM3NeT detectors is optimised for the energy regime of interest. In table 3.2 the main characteristics of the KM3NeT detectors are summarised. In figure 3.6, a schematic outline of a KM3NeT/ARCA building block is given.

As the DU structure is not rigid, the DOMs can move under the influence of sea currents. To allow for the monitoring of its position and orientation, each DOM is equipped with an acoustic receiver, an accelerometer and a digital compass. A network of acoustic emitters on the seabed is foreseen to be used in the monitoring of the varying DOM positions[163].

At the moment of writing, this acoustic positioning network is not fully functional. Therefore, the DOMs will be assumed to be aligned vertically, with the DOM positions along the DU as estimated from water pressure measurements from a submersible, used for inspection of the deployed DUs. Preliminary studies indicate a maximum tilt of a KM3NeT/ARCA DU of the order of 0.6 ° [164].

		KM3NeT/ORCA	KM3NeT/ARCA
Inter-DOM vertical spacing [m]		9	36
Inter-DU horizontal spacing [m]		20	90
Instrumented cylinder (one building block)	radius [m]	106	538
	height [m]	153	612
	volume [km^3]	$5.8 \cdot 10^{-3}$	0.56
Number of building blocks		1	2
Depth	seabed [m]	2425	3450
	top DOM [m]	2232	2773
	bottom DOM [m]	2385	3385
Location	shore station	Toulon, France	Capo Passero, Italy
	shore distance [km]	40	100

Table 3.2: *Characteristics of the KM3NeT neutrino detectors. For a description of the abbreviations used, please refer to the text.*

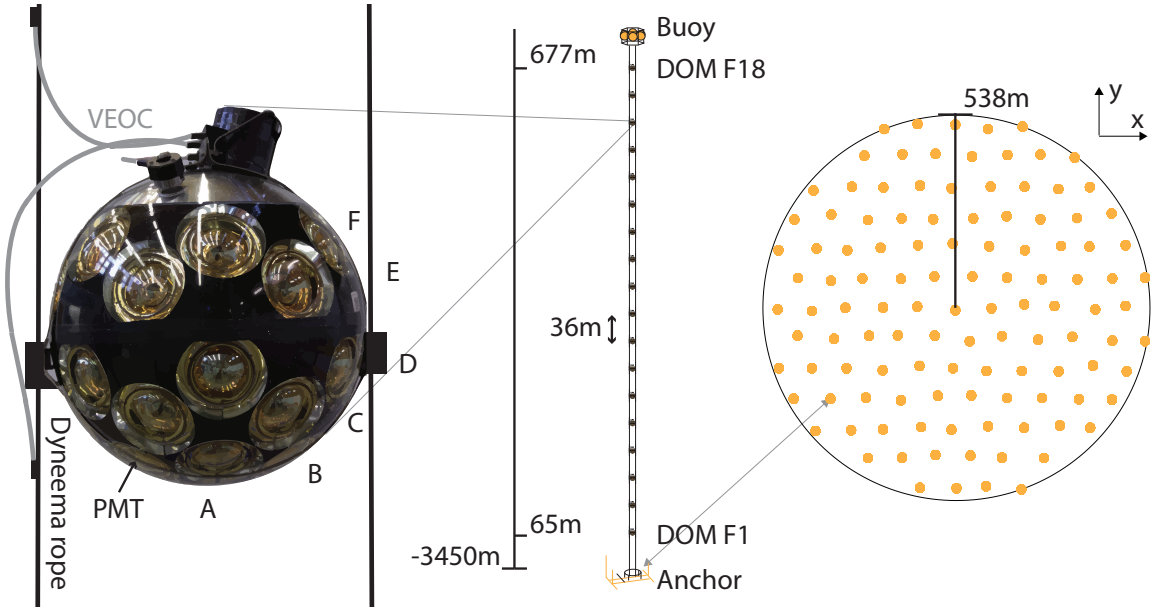


Figure 3.6: *A KM3NeT DOM (**left**), KM3NeT/ARCA detection unit (**centre**) and outline of the DU positions of one KM3NeT/ARCA building block on the seabed (**right**). The letters indicate the ordering of the PMTs in rings. An aluminium element to cool the internal electronics is located at the top. Each PMT is surrounded by a light collection ring. The grey lines marked as 'VEOC' correspond to a tube housing the communication fibres and power cables (see section 3.7).*

3.5.1 Photomultiplier Tubes

In the first construction phase, KM3NeT utilises PMTs with a diameter of 7.6 cm (Hamamatsu R12199 [165, 166]), with a photocathode area of 45.4 cm^2[†]. To enhance the *effective area*, each PMT is encircled by a reflective light collection ring. The resulting effective area of a *nominal PMT* as function of the light incident angle is shown in figure 3.7 (left). By integrating the PMT effective areas, the effective area of all 31 PMTs in a DOM for a photon arriving from a random direction is about $31 \cdot 14.8 \approx 460$ cm^2. This is 33 % of the 1384 cm^2 projected area of a DOM, implying that one third of the photons crossing a DOM will encounter a PMT.

If a photon hits the PMT photocathode, an electron (a *photoelectron*) can be emitted as a result of the photoelectric effect. The probability for this to happen is called the *quantum efficiency* (QE), see figure 3.7 (right). Using the absorption length of light (figure 3.5), the number of Cherenkov photons remaining after propagating 50 m through seawater is shown as a reference (in arbitrary units). The PMTs are sensitive in the wavelength range where most of the signal light is expected.

After emission, an electric field accelerates the photoelectron towards a chain of dynodes. The probability of the photoelectron to hit the first dynode is about 90 % (this number is called the *collection efficiency*). On impact, each electron liberates more secondary electrons that are accelerated to the next dynode. This process is repeated up to the anode. The average number of electrons reaching the PMT anode from a single emitted photoelectron is called the *gain* of the PMT, the spread in this distribution is called the *gainspread*. The ensemble of electrons hitting the PMT anode results in a measurable current.

For a KM3NeT PMT, it takes about 30 ns between a photon hitting the photocathode and the electrons reaching the anode [166, 168]. This quantity is called the *transit time* of a PMT. The *probability density function* (PDF) of the transit time of a typical KM3NeT PMT is shown in figure 3.8. The spread on the mean transit time (the *transit time spread* (TTS) is about 2 ns. This spread is the main uncertainty on the measured photon arrival time.

In addition to the main peak, the transit time distribution has a second peak between a transit time of 40 and 65 ns. These hits are attributed to *delayed pulses*, which originate from a secondary electron that back-scatters from the first dynode[165]. The back-scattered electron is re-accelerated towards the first dynode, giving a delayed pulse, typically with a lower amplitude than a 'normal' pulse. Even later hits can occur in case the photoelectron is released from a point on the photocathode where the electric field strength is relatively low. Because of this, the freed photoelectron exhibits a significantly larger drift time. These hits are called *late hits*. Photons that free a photoelectron from (one of) the first dynode(s), rather than from the photocathode, can cause a hit with a small amplitude and short transit time. These hits are called *prepulses*.

[†]Currently, the Hamamatsu R14374-20 PMT is used, with a diameter of 8.0 cm

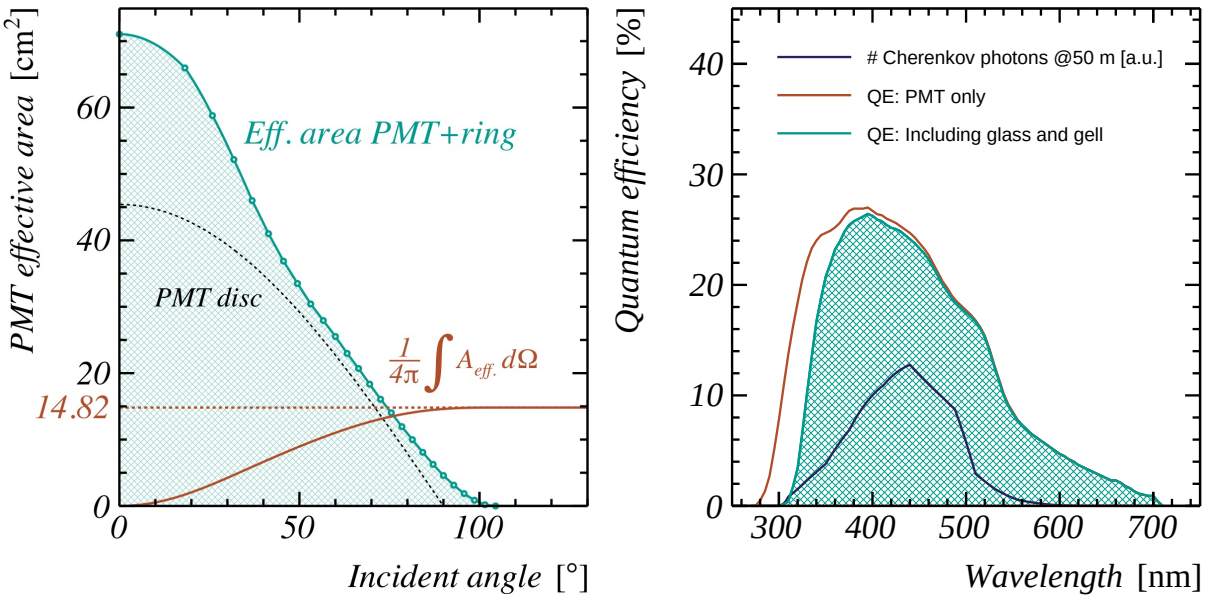


Figure 3.7: ***Left:*** *The effective area of a PMT including the light collection ring as function of the incident angle of the light [167]. The effective area of a flat disc with the area of a PMT is shown as a reference. The average effective area of the PMT plus light collection ring is* 14.82 cm^2. ***Right:*** *PMT quantum efficiency. The blue line includes the absorption of light in the glass and gel surrounding the PMT. The used data is taken from Aiello et al.[165]. The distribution of the expected number of Cherenkov photons after propagating* 50 *m through seawater is given for comparison. This distribution is given in arbitrary units.*

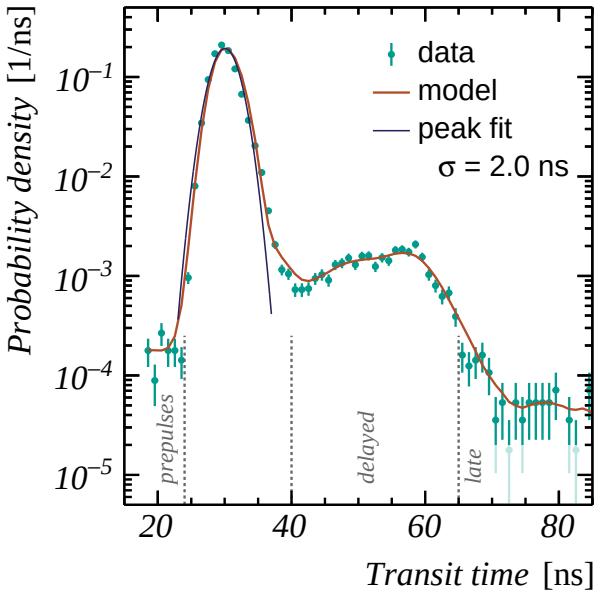


Figure 3.8: *The transit time distribution of a typical PMT used in KM3NeT (data from [165, 168]). The transit time spread (*2.1 *ns) is estimated from a fit to the peak of the distribution. Delayed pulses cause the excess at* 40 – 65 *ns. A parameterisation of the measured data is indicated by the solid line.*

3.6 Backgrounds

Neutrinos are not the only sources of detectable light in the deep sea. Three sources of optical background will be discussed in this section.

3.6.1 Potassium-40 Decays

The majority of the recorded hits in KM3NeT originate from the decays of ^{40}K isotopes, naturally abundant in sea salt. The main decay channels of ^{40}K are:

$$^{40}K \rightarrow {}^{40}Ca + e^- + \bar{\nu}_e \quad (\beta\text{-decay, 89.3 \%}),$$
$$^{40}K + e^- \rightarrow {}^{40}Ar + \gamma + \nu_e \quad (\text{electron capture, 10.7 \%}).$$

In figure 3.9, the kinetic energy distribution of the emitted electron in the β-decay channel is shown. The produced electron in the β-decay channel can carry a kinetic energy up to 1.311 MeV. In about 21 % of the β-decays, the electron has sufficient energy for the emission of Cherenkov radiation. A ≥ 0.76 MeV electron produced via the β-decay channel emits on average 38 Cherenkov photons[169].

The photon produced in the electron capture channel (with an energy of 1.461 MeV) frees electrons from the water molecules through Compton scattering, subsequently producing Cherenkov light as well. The energy of the emitted photon is larger than the energy of the electron in the β-decay channel, resulting in considerably more emitted Cherenkov photons (about 69, see [169]).

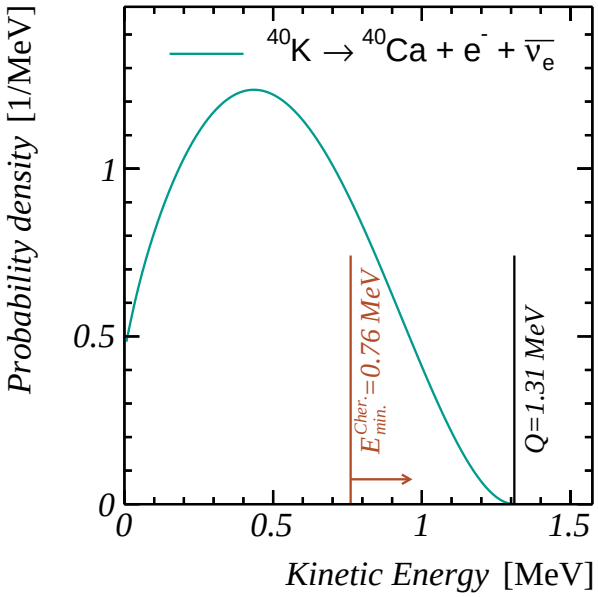


Figure 3.9: *The probability density function of the kinetic energy of the electron produced in the β-decay of ^{40}K , calculated according to [170]. The maximal kinetic energy (1.31 MeV) and threshold kinetic energy for the emission of Cherenkov radiation (0.76 MeV) are indicated by vertical lines.*

The expected hit rate from ^{40}K decays on a typical KM3NeT photomultiplier tube can be calculated using the ^{40}K decay rate, the salinity of the sea, the PMT parameters and the absorption length of light in water. This calculation is given in appendix B, yielding a singles rate of hits of about 4.7 kHz per PMT in KM3NeT/ARCA. The concentration of ^{40}K isotopes in the glass and electronic noise in the PMT are estimated to give an additional singles rate of about 1.2 kHz (on a nominal PMT) [165, 171].

Thanks to the multi-PMT design of the DOM, a large fraction of background from ^{40}K decays can be rejected from standalone DOM data by selecting hits detected on two or more PMTs in the same DOM within a time window of ≈ 10 ns (see section 7 and a recent paper by the KM3NeT Collaboration[172]).

3.6.2 Bioluminescence

Some algae and creatures in the deep sea emit light called *bioluminescence*. Bioluminescence causes an increased hit rate on a single DOM for a period lasting up to several seconds. Studies have shown that the bioluminescence activity is related to the sea current, and mainly observed close to the detector structures [173]. Therefore, it is believed that bioluminescence is stimulated by encounters with the detector and by the vortices created by the drag of the detector.

The background from bioluminescence is reduced using a *high rate veto*. In case the observed rate of a PMTs exceeds 20 kHz, the data recorded on that PMT is vetoed by the data acquisition electronics in the DOM (see next section) and will not be sent to shore.

3.6.3 Atmospheric Muons

Despite the water overburden, some muons produced in cosmic ray air showers (*atmospheric muons*) reach the KM3NeT detectors. In figure 3.10 the rate of atmospheric muons is given as function of depth below sea-level. The atmospheric muon rate is several orders of magnitude higher than the expected signal rate of (cosmic) neutrinos (see figure 2.8 (right)). Atmospheric muons can mimic track-like signatures of neutrino interactions, thus constituting a significant background for the identification of (cosmic) neutrino sources.

The probability that an atmospheric muon traverses the Earth is negligible. Using this, the atmospheric muon background can be reduced by selecting only those events with a reconstructed upward direction, effectively using the Earth as a shield. However, as the neutrino parameters are not reconstructed with absolute precision, it is inevitable that a fraction of the atmospheric muon events pass this criterion and are falsely identified as neutrino events.

3.7 Data Acquisition

The process of collecting and storing the data (the *data acquisition*) will be discussed in this section.

3.7.1 Hardware

PMT base Each PMT is soldered to an electronics board called the *PMT base*. The PMT base provides the high voltage for the PMT and amplifies the PMT output waveform. Additionally, the waveform of the PMT is transformed into a square wave pulse, by comparing the recorded waveform to a predefined threshold voltage (see figure 3.11). The threshold voltage of each base is set to a constant value for all PMT bases.

Central logic board The square wave pulses of all 31 PMT bases in a DOM are sent to a *field-programmable gate array* (FPGA) on a *central logic board* (CLB) for further digitisation. Each detected pulse is characterised by two quantities: the time at which the PMT waveform passes through the threshold (the hit time) and the period it stays above the threshold, called the *time over threshold* (ToT). The ToT gives a measure of

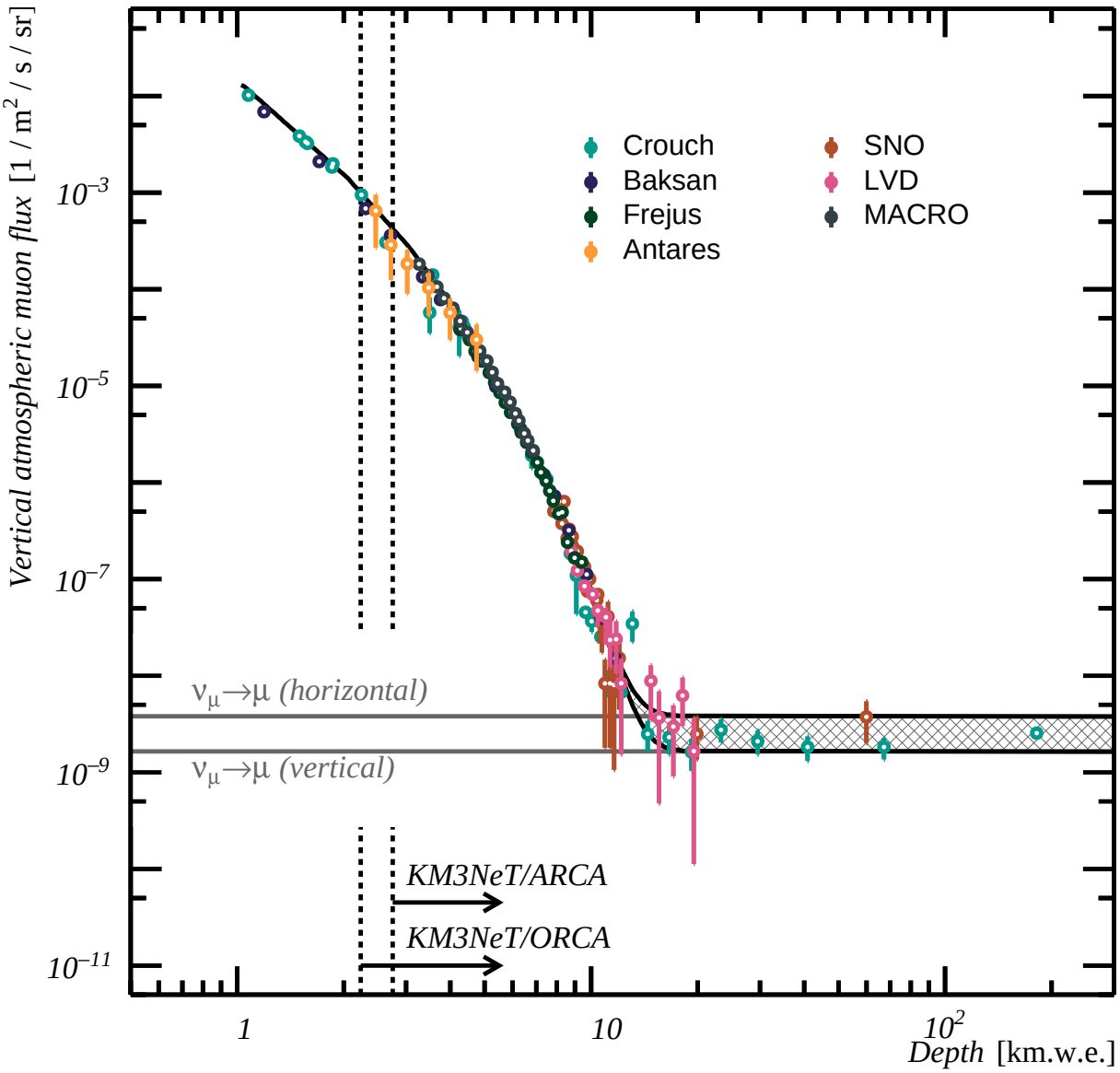


Figure 3.10: *The measured atmospheric muon flux as function of traversed distance through water. The presented data is obtained by various experiments[174–180]. At depths larger than ≈ 10 km, almost all atmospheric muons are absorbed by the water overburden. In this regime, muons produced in interactions of atmospheric muon neutrinos with the surrounding matter dominate (indicated by horizontal grey lines). The solid black curve gives a parameterisation of the horizontal and vertical flux, and is taken from the particle data group[4]. The dashed vertical lines indicate the depth of the KM3NeT detectors.*

the amplitude of each pulse. Both quantities are rounded off to an integer number of nanoseconds. The combination of a hit time and the corresponding ToT is called a *level zero hit* (L0).

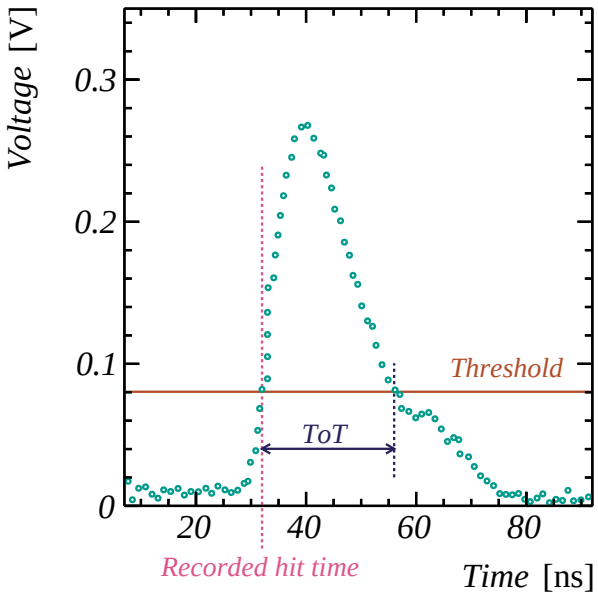


Figure 3.11: *A typical PMT pulse, as measured in a laboratory (data from Aiello et al.[165]). The horizontal line gives the threshold voltage (see text). In KM3NeT, the time the waveform passes and exceeds the threshold are stored for further analysis.*

Seafloor network A network of electrical wires and optical fibres is used to power the DOMs and to communicate with the CLBs through an Ethernet network. The 18 DOMs in a DU are connected to a *base module* on its anchor using a *vertical electro-optical cable* (VEOC). In turn, the base modules are connected to a cable running to shore using a network of junction boxes.

3.7.2 High Voltage Tuning

The high voltage setting of each PMT is set such that its gain is approximately $3 \cdot 10^6$. The unit *photoelectron equivalent* (p.e.) is introduced as a convenient quantity to describe the deposited charge; a nominal PMT has a gain of 1 p.e..

The *high voltage tuning* is done in-situ, using the distribution of the time-over-threshold values of recorded L0s. The high voltage setting of each PMT is set such that the ToT distribution corresponding to a single photoelectron peaks at a ToT value of 26.4 ns.

3.7.3 All-data-to-shore Principle

The bandwidth of the optical communication network is sufficient to allow for an *all-data-to-shore* approach. In this approach, no data reduction other than the high rate veto is applied offshore, and all recorded L0s are sent to shore. For efficient handling of the data in the shore station, the L0s are bundled in slices of 0.1 s.

3.7.4 Software and Recorded Data

On shore, a set of computers, called the *dataqueue*, sort and bundle the received data from all DOMs in *L0 timeslices*. Each L0 timeslice contains all L0s recorded in the entire detector during 0.1 seconds.

L0 runs Data taking runs in which all L0 timeslices are written to disc are called *L0 runs*. Since the vast majority of the recorded L0s originate from the ^{40}K background, L0 runs are only stored for calibration and monitoring purposes, and typically last only a few minutes.

L1 timeslices When two (or more) L0s are recorded on the same DOM within a time window of 25 ns, the corresponding cluster is called a *level one hit* (L1). The expected rate of L1s is approximately 0.75 kHz per DOM, significantly lower than the expected L0 rate ($\approx$ 200 kHz per DOM). During the data taking periods considered in this thesis, the complete data stream of L1s was stored for monitoring and calibration purposes. The corresponding timeslices are called *L1 timeslices*, taken during *L1 runs*.

Triggered events Possible neutrino induced events are identified by dedicated trigger algorithms. These algorithms search for patterns in the L0 data compatible with a shower- or track-like event signature [181, 182]. In case such a *trigger* is identified, a snapshot of all L0s recorded in the detector in a time window surrounding the trigger are written to disc. This time window equals the maximum time of flight of a muon propagating through the instrumented volume ($[-5677, 5677]$ ns for KM3NeT/ARCA). The resulting L0 snapshots are called *triggered events*. They form the main data stream used in most neutrino physics analyses.

3.8 Time Synchronisation and Calibration

Detected L0s are timestamped by the FPGA on the CLB of each DOM. These timestamps do not represent the true hit times, because of two reasons:

- Each L0 is timestamped with a PMT specific time delay, caused by the PMT transit time and electric cables between the anode and the FPGA.
- In order to assure a correct timestamp of the L0s, the clocks in the CLBs of all DOMs should be syntonised (operating on the same frequency), as well as synchronised (operating in phase). Phase differences between the FPGAs is caused by time delays in the Ethernet communication network. By design, the CLB clocks are syntonised, but not necessarily synchronised[183].

The combined effect is characterised by the total time delay of each PMT, called the *PMT time offset*. The PMT time offsets are added to the stored L0 timestamps prior to any physics analysis. In order to assure an accurate determination of the parameters of detected neutrino interactions, a goal has been set to calibrate all PMT time offsets with (sub-) nanosecond accuracy. The determination of the PMT time offsets is done in two steps:

Inter-DOM time calibration The aim of the *inter-DOM time calibration* is to determine the average of the PMT time offsets of the 31 PMTs in each DOM. This quantity is called the *DOM time offsets*.

Inter-PMT time calibration The relative PMT time offset of each PMT, defined as the difference between the PMT time offset and the DOM time offset of the corresponding DOM, is determined by the *inter-PMT calibration*.

The inter-PMT calibration is done using data taken in-situ, exploiting coincident hits originating from ^{40}K decays. The procedure is the subject of chapter 6 in this thesis.

Several signals can be exploited to determine the inter-DOM time calibration:

- *Dark room calibration*: In a laboratory environment, a calibrated laser signal is fired on a reference PMT of each DOM [184].
- *LED nanobeacon calibration*: Each DOM is equipped with an LED beacon that can be activated to illuminate the DOMs in its vicinity [31].
- *Atmospheric muons*: Based on coincident light from atmospheric muons passing by the DOMs.

Each calibration has advantages and disadvantages. The LED nanobeacon calibration is done in-situ, but depends on the position calibration of the DOMs, optical properties of the water and the LED light emission profile. Furthermore, it requires dedicated data-taking runs, effectively inducing downtime of the detector.

The laser calibration does not depend on the DOM positions, but can only be done in a laboratory environment. Data-taking conditions (such as in-situ high voltage tuning) complicate a reliable estimate of the DOM time offsets during the foreseen lifetime of the detector. Furthermore, only a single PMT per DOM is calibrated, thus giving a larger statistical uncertainty on complete time calibration.

Of the three methods, the in-situ calibration exploiting the light emitted by atmospheric muons has the biggest advantages. It does depend on the DOM position calibration, but no interruption of data taking is necessary, providing a large statistical set of signal over the full lifetime of the detector. Two methods for inter-DOM time calibration based on detected muons are presented and discussed in chapter 8.

Inter-DU time calibration The largest time delays in the detector are caused by the seafloor network of optical fibres and junction boxes connecting the DU base modules to the shore station. To determine these time offsets, each base modules is equipped with a clock, which is both syntonised and synchronised with the 'master' clock on shore. The corresponding *inter-DU time offsets* are known with sub-nanosecond accuracy by the use of a white-rabbit protocol[183].

3.9 Effective Area

The effective area of neutrino detectors is generally parameterised as a function of the neutrino energy (E_ν) and zenith angle (θ_ν). Using the effective area, it is straightforward to calculate the expected rate of detected events (Φ_ν, in hertz) with neutrino energies in a range $E_A \leq E_\nu \leq E_B$ and directions $\hat{d}_n$ in a fraction $\hat{d}_n \in \omega$ of the full solid angle for a given neutrino flux $\frac{d^3\Phi_\nu^{atm.}}{dEd^2\Omega}(E_\nu, \hat{d}_\nu)$:

$$\Phi_{\nu}^{det.}(E_A \leq E_\nu \leq E_B, \hat{d}_n \in \omega) = \iiint_{E_A,\omega}^{E_B} A_{\nu}^{eff.}(E_\nu, \theta_\nu) \cdot \frac{d^3\Phi_{\nu}^{atm.}}{dEd^2\Omega}(E_\nu, \hat{d}_\nu) d^2\Omega dE. \tag{3.18}$$

In figure 3.12, the effective area of one KM3NeT/ARCA building block detector is shown (obtained from simulations, see chapter 4). It includes the probability that a neutrino propagated through the Earth without being absorbed, interacts in the vicinity of the detector and causes an triggered event. As such, this effective area has to be combined with a flux defined at the Earth's atmosphere to give the expected rate of triggered events.

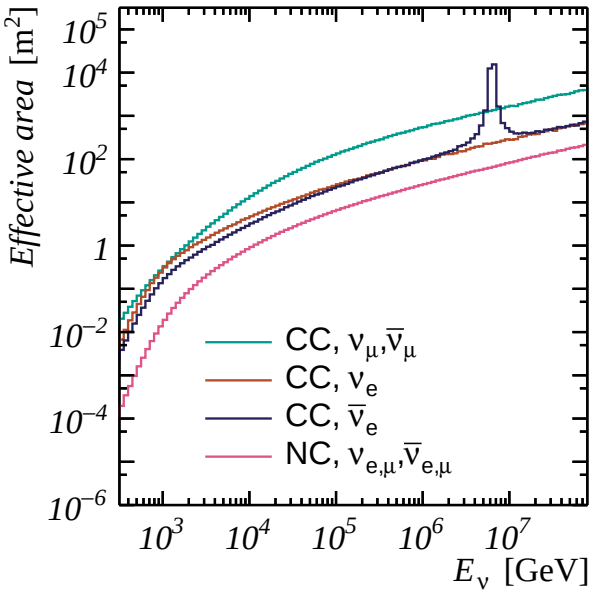


Figure 3.12: *The direction-integrated effective area of a charged-current interaction of a muon (anti)neutrino (detected mainly as tracks), charged-current ν_e interactions (giving a hadronic and EM shower), charged-current $\bar{\nu}_\mu$ interactions (same as ν_e, but with an enhanced cross section at 6.3 PeV from electron-$\bar{\nu}_e$ interactions) and a neutral current interaction of a neutrino or antineutrino of any type (hadronic shower only). These effective areas are given for triggered events in a single KM3NeT/ARCA building block, as obtained from simulations (see chapter 4).*

The enhanced electron-$\bar{\nu}_e$ cross section at the Glashow resonance results in a significantly larger corresponding effective area at 6.3 PeV. The effective area of muon (anti)neutrinos is larger than the effective area of electron neutrinos at neutrino energies above ≈ 1 TeV. The difference is explained by the range of the muon produced at these energies, resulting in a higher detection efficiency. The effective area of neutral current interactions, on the other hand, is about five times smaller than the charged-current ν_e effective area, a difference caused by the lower neutral current cross-section and the fact that (on average) less light is emitted in a neutral current interaction.

3.10 Detector Status

KM3NeT/ARCA The construction of the KM3NeT/ARCA detector started in December 2015, with the successful deployment of the first KM3NeT/ARCA detection unit. A second KM3NeT/ARCA DU was deployed in May 2016. Both detection units took data until the 20th of April 2017, at which time a technical issue with one of the junction boxes in the sea-floor network prevented further data taking. During a sea campaign in

January 2019, this junction box was replaced, solving this issue. However, when the DUs were powered on again, one of the DUs showed symptoms of an electrical short-circuit, and communication with this detection unit was lost. This detection unit was recovered for inspection, which revealed a water leakage in the connection of one DOM with the VEOC. It is likely that this leakage caused an unfortunate shortcut with the failing junction box. Fortunately, the other DU was not affected by this, and has taken data ever since, interrupted only by a power-off on November the 11th, 2019, to investigate a pending issue in the shore station. Since April 2021, following the successful deployed of five more detection units, the KM3NeT/ARCA detector is operating with six detection units.

The parameters of the first detected atmospheric muons events were reconstructed already within the first 15 minutes of operation of the first KM3NeT/ARCA detection unit. Using the algorithm developed for the work presented in this thesis (see section 5.3.4), the parameters of the corresponding muon tracks were estimated. In figure 3.13, the arrival time distribution of the corresponding hits on the DOMs of the KM3NeT/ARCA detection unit are shown, as well as the expected arrival times of the reconstructed muon track. The detected event is compatible with an atmospheric muon passing by the DU at a *zenith angle* of approximately 174 ° .

KM3NeT/ORCA The KM3NeT/ORCA detector started operation with one detection unit in September 2017. A second KM3NeT/ORCA detection unit is operational since May 2019. Four additional DUs were deployed at the first of July, 2019 (two DUs) and on January the 27th, 2020, respectively.

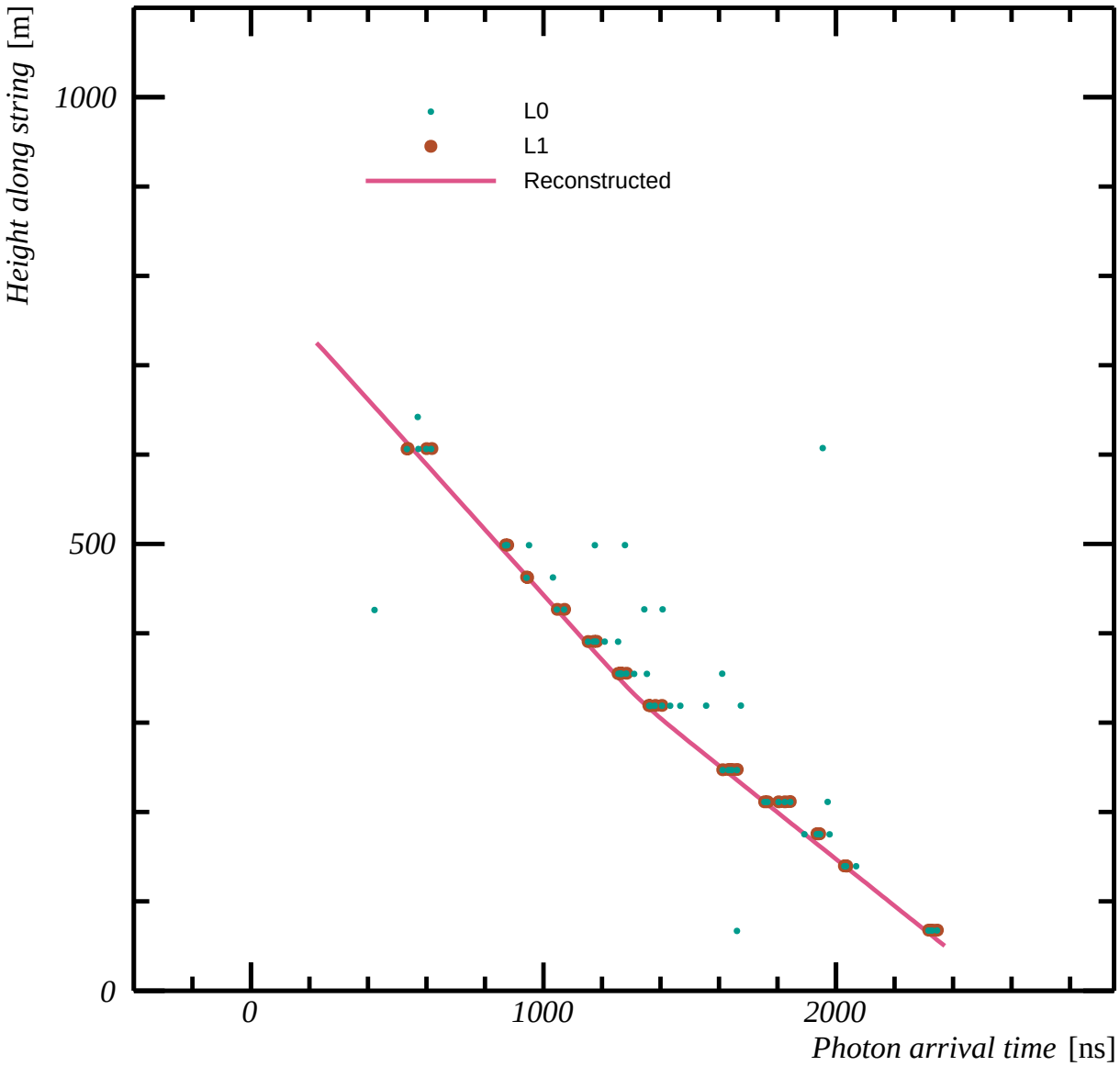


Figure 3.13: *One of the first events of KM3NeT (likely an atmospheric muon), reconstructed within 15 minutes after the collection of the first KM3NeT/ARCA data on December the 4th, 2015. The dots give the arrival times of L0s on the DOMs. The corresponding muon track has been reconstructed using the algorithm described in section 5.3.4. The solid line shows the expected arrival time distribution of the reconstructed muon track. The most likely track is moving downwards, with a zenith angle of* 174 ° .

Chapter 4

Simulations

Simulations are widely used in KM3NeT to develop and characterise the performance of (algorithms used in) physics analyses. For the studies presented in this thesis, events originating from neutrino interactions, as well as the backgrounds from ^{40}K decays and atmospheric muons, are simulated. The used techniques and simulation algorithms will be discussed in this chapter.

4.1 Simulation Outline

The simulation generally follows three stages:

1. In the first stage, interactions of the primary particles of interest are simulated in the vicinity of the detector, resulting in a list of the secondary particles produced in these interactions.

2. The propagation and interactions of the secondary particles, as well as the light induced by these particles and subsequent light propagation through the seawater is simulated in the second stage. This stage includes the optical properties of the DOM, a model of the PMT angular acceptance and the wavelength-dependent quantum efficiency distribution. The output of this stage is equivalent to a list of emitted photoelectrons in each PMT.

3. The remaining detector response is simulated in a third stage. This stage accounts for the PMT-to-PMT efficiency, PMT transit time and incorporates a model for the simulation of L0s.

The format of the simulated events after the third stage is (by design) identical to that of observed data. All preceding algorithms (such as the trigger and reconstruction algorithms, see sections 3.7.4 and 5) are identical for both data and simulations.

Various software packages are available for each stage, optimised for the physics processes of interest (see table 4.1). In the following sections, these will be discussed.

Stage	^{40}K decays	Atm. muons	Neutrinos
1. Primary interactions	OMGsim	MUPAGE	Genhen
2. Propagation and light		JSirene	KM3
3. Detector response	JPP		

Table 4.1: *Overview of different software packages used in KM3NeT simulations.*

4.2 Generation Volume and the 'Can'

Interactions of the primary particles are generated in a volume called the *generation volume*. The dimensions of the generation volume depend on the interaction of interest, and are chosen such that particles originating outside of it have a negligible probability to traverse a second (smaller) volume called the *can*.

The *can* defines a volume in which light emission by particles is simulated (stage two). All (light emitting) particles outside of the can are propagated to the can taking into account energy losses and interactions, but without the actual simulation of the light emission. The dimensions of the can are chosen such that the probability that a particle not traversing it causes a triggered event is negligible. The combination of the can and generation volume result in minimal CPU usage while guaranteeing a negligible probability that an interaction happening outside of the generation volume causes a triggered event. A schematic sketch of the used generation volumes is given in figure 4.1.

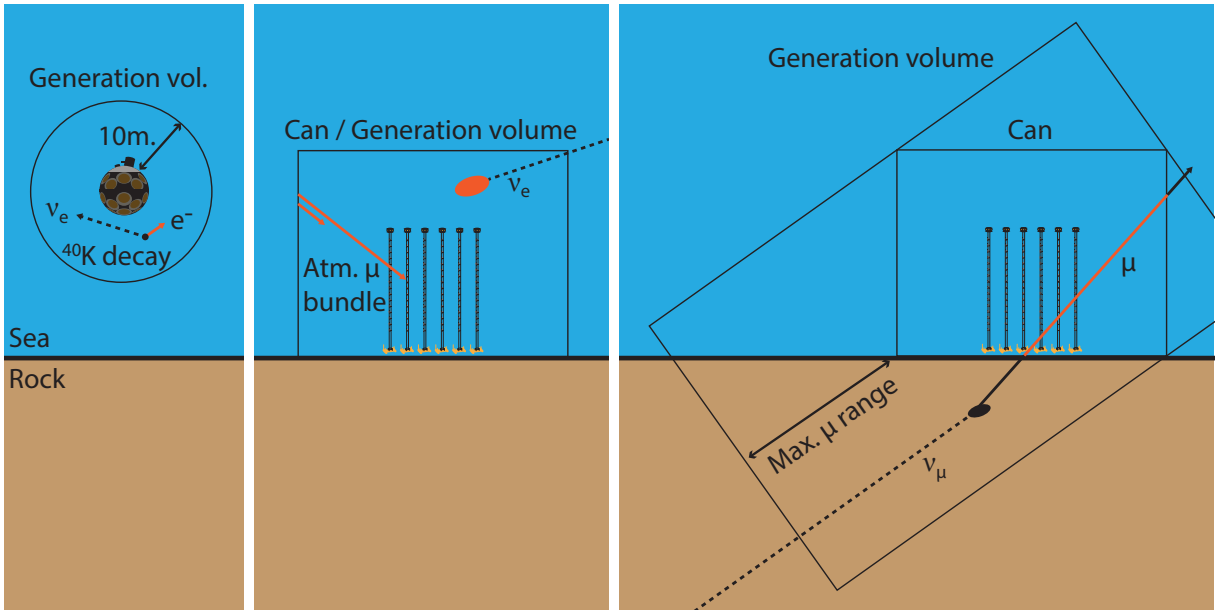


Figure 4.1: *A sketch of the detector, light simulation "can" (see text) and generation volume used in the simulation of ^{40}K decays (**left**), atmospheric muon and shower-like events (**middle**) and track-like events (**right**). Particle trajectories indicated by a dashed line do not induce the emission of light, trajectories with a solid line do. Light emission is simulated only for particle trajectory segments with an orange colour.*

^{40}K decays On average, each ^{40}K decay leads to about 41 emitted photons in the wavelength range to which KM3NeT is sensitive[169]. Because of this relatively low intensity, ^{40}K decays are generally observed by a single PMT only. For the work presented in this thesis only the light from ^{40}K decays detected in coincidence by two or more PMTs within the same DOM are of interest. The probability to observe such a coincidence from a ^{40}K decay occurring more than 10 m away from a DOM is negligible. Consequently, the generation volume for ^{40}K isotopes consists of a spherical volume with a radius of 10 m surrounding a single DOM.

Atmospheric muons Atmospheric muons are simulated starting from the point where they enter the can. The dimensions of the can are chosen such that a muon not traversing it has a negligible probability to cause L1(s) in the detector. Generally, a volume extending the instrumented volume by 200 m (which is about four light absorption lengths) is chosen, which is a compromise between the number of simulated undetected (low-energy) muons and the relatively large can needed for high-energy muons. The dimensions of the can do not depend on the energy of the muon.

Neutrino interactions The generation volume and can used in the simulation of neutrino interactions depends on the neutrino flavour and interaction channel. For neutral current neutrino interactions and electron neutrino charged-current interactions, the generation volume is taken to be identical to the can, as the resulting showers have a range of a few metres. The muon and tau leptons produced in charged current muon neutrino and tau neutrino events can propagate significantly further (see section 3.2.5). The generation volume used for these events is a cylinder with its' axis parallel to the neutrino direction, encompassing the can, with an extension of the maximum muon and tau flight length in the neutrino direction respectively (see figure 4.1).

4.3 Primary Interactions

4.3.1 The OMGsim Software Package

Particles produced in the decays of ^{40}K isotopes are simulated with the OMGsim software package [185]. The used simulation chain differs from the simulation of atmospheric muons and neutrino interactions in the sense that the same routine is used for both the simulation of the interactions (stage 1) and subsequent light emission, propagation and detection (stage 2). OMGsim is based on the GEANT4 software package [186] for the simulation of ^{40}K decay products, light emission and light propagation through the seawater and DOM. Subsequent simulation of the PMT photocathode is performed using a detailed model of the PMT photocathode material [187]. The used model of the KM3NeT DOM includes the optical properties of the glass, gel, light collection ring and aluminium cooling element (see section 3.5 and figure 3.6), but (for practical reasons) does not include the DOM support structure and VEOC.

4.3.2 Simulation of Atmospheric Muons: MUPAGE

Events caused by atmospheric muons are simulated using the MUPAGE software[188]. Individual muons, as well as multiple muons originating from the same cosmic ray air shower (*muon bundles*) are simulated.

A (CPU-intensive) full simulation of the cosmic ray air shower and subsequent propagation of the muons is avoided by the use of parameterisations of the expected flux of atmospheric muons, as a function of the bundle multiplicity, zenith angle and depth below the sea surface[188] (see figure 4.2). In this parameterisation, all muons in a bundle are assumed to be moving parallel to one common axis. The bundle multiplicity, the shower axis direction and the position where this axis intersects with the can are generated accordingly. In case the simulated multiplicity is one (i.e. a single muon), the energy of the muon is picked from a second parameterisation, depending on the slant depth of the muon. For muon bundles (multiplicity larger than one), the distance of each muon to the shower axis is sampled from a third parameterisation. The energy of each muon is then sampled from the parameterisation used for single muon events.

The parameterisations used in MUPAGE are tuned to match a set of simulated cosmic ray air showers using the HEMAS [189] software package. The main uncertainty in this approach comes from the hadronic interaction models used in the air shower simulation. These uncertainties are generally taken into account by assuming an uncertainty of 10 % of the total atmospheric muon flux. A more detailed description of MUPAGE is given by Carminati et al.[188].

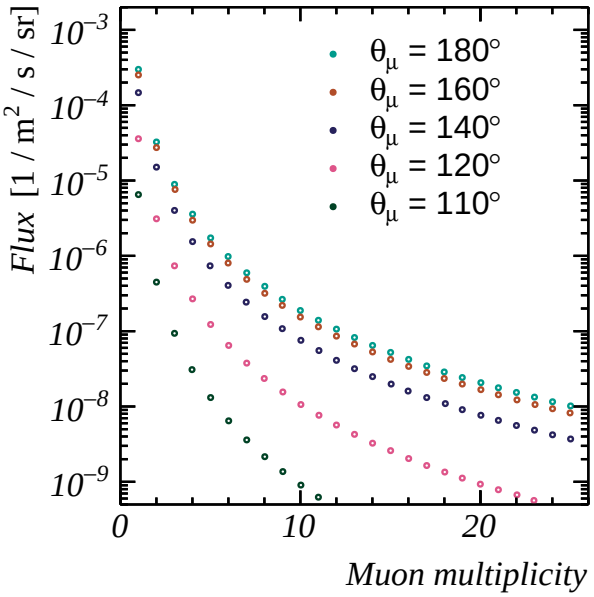


Figure 4.2: *The flux parameterisation used in MUPAGE at 2.5 km water equivalent vertical depth as function of the muon multiplicity of the bundle. The different sets correspond to various zenith angles (θ_μ =180 ° corresponds to vertically down going) of the muon (bundle). Data from Carminati et al.[188].*

4.3.3 Neutrino Interactions: Genhen

Neutrino interaction induced events are simulated using the Genhen software package [190]. The flavour of the neutrino and the interaction channel (charged/neutral current) are set by the user. Each simulated neutrino is forced to interact anywhere in the gen-

eration volume. The neutrino direction is taken isotropically from the (4π solid angle) hemisphere, the neutrino energy is sampled from an energy spectrum $dN/dE \propto E^{-1.4}$.

The momentum transfer onto the target nucleon and scattering angle of the neutrino are randomly sampled from the double differential neutrino interaction cross-section, described in sections 3.3 and appendix A. In case the interaction of a charged-current muon or tau neutrino is simulated outside of the can, the resulting lepton is propagated to the can using the MUSIC software package [191]. MUSIC takes into account the energy losses described in 3.2.5, as well as the scattering of the muon direction.

The energy spectrum used to simulate the neutrino energy ($dN/dE \propto E^{-1.4}$) is motivated by statistics rather than by physics (the diffuse astrophysical neutrino flux has a spectral index of approximately 2.5, see section 2.3). The signal from cosmic neutrino sources is expected to be prominent at (very) high energies (see section 2.3). If neutrinos would be simulated according to a $dN/dE \propto E^{-2.5}$ energy spectrum, a (statistically speaking) sparse sample of high-energy events would be obtained, while wasting computing resources on the relatively uninteresting low-energy events. Any (physical) neutrino flux can be inferred from the generated events using *event weights*. This procedure is described in appendix C.

4.4 Propagation of Secondaries and Light Simulation

Two software packages, named *JSirene*[192] and *KM3*, are used to simulate the propagation of muons and other (secondary) particles through seawater, according to the energy losses discussed in section 3.2.5. The emission of light in (some of) these processes, consecutive light propagation through seawater and the detection of photons by the photocathode of KM3NeT PMTs are simulated as well. In this, models of the optical properties of the DOM, the PMT effective area and the quantum efficiency of a nominal PMT are used. The output of both applications consists of a list of the arrival times of all detected (i.e. causing the emission of a photoelectron) photons on each PMT.

A charged particle induces thousands of Cherenkov photons per metre of traversed seawater (see section 3.2.4). As a consequence, it is computationally demanding to simulate the absorption and scattering of all these photons propagating through the medium. The assumptions and simplifications made to overcome this problem will be discussed in the following paragraphs.

4.4.1 JSirene

In JSirene, each muon is assumed to follow a straight path. Photons caused by Cherenkov radiation along its track are simulated until the muon energy drops below the Cherenkov energy threshold. The stochastic energy losses due to bremsstrahlung, pair production and photonuclear interactions are parameterised as electromagnetic showers at random positions (distributed according to their respective interaction lengths) along the muon track. Light from delta rays is emitted isotropically and continuously along the muon track. The energy loss due to ionisation is taken into account as well, but does not lead to the emission of light.

For each of the light emitting processes, the number of expected photoelectrons is evaluated at a set of predefined abscissae covering the phase-space of the arrival time and position and direction of the PMT relative to the track. The calculated points are stored in tables. Evaluation of the expected number of photons on points between the stored abscissae is done by means of multi-dimensional interpolation [193]. In appendix D, these functions are described in detail.

4.4.2 KM3

A software package, complementary to the JSirene software (see section 4.4.1), used to simulate the emission propagation and detection of photons is named KM3. Even though the time required to simulate an event is significantly longer for KM3, it is still maintained as a cross-check of the JSirene package and, for historical reasons, used in the simulation of neutrino events.

Similar to the JSirene software package, KM3 makes use of pre-computed tables of the number of expected photoelectrons as a function of the arrival time of the photon, the position and orientation of the PMT. These tables are evaluated for muon tracks and electromagnetic showers, using a set of simulated events for each. In the simulation, the muon energy loss is simulated using the MUSIC software package [191]. For the simulation of showers, the GEANT4 software package is used [186]. The used model of the water properties and PMTs are discussed in sections 3.5.1 and 3.4.

4.5 Detector Response Simulation

A model of the PMT hardware is used to generate L0s from the simulated photoelectron emission times on each PMT. This model is part of a general software framework used in KM3NeT, called JPP[194]. Random light from ^{40}K decays is added prior to the detector simulation, according to the expected ^{40}K singles rate (about 6.7 kHz per PMT, see section 3.6.1). The detection time of the given list of photoelectron emission times are sampled from a parameterisation of the measured transit time distribution (shown in figure 3.8).

Differences between the PMT properties, which can vary with time, are taken into account by a model in which the relevant parameters can be set for each individual PMT. Among the tune-able parameters are the relative PMT gain g and gainspread σ_g, which are used to describe the probability density to observe a certain deposited charge $n_{p.e.}$ for a given number of detected photons n_γ:

$$\frac{\partial P}{\partial n_{p.e.}}(n_{p.e.}|n_\gamma) = \frac{1}{\sqrt{2\pi}\sigma_g} \exp\left(-\frac{(n_{p.e.} - g \cdot n_\gamma)^2}{2\sigma_g^2}\right), \tag{4.1}$$

and a set of parameters (the 'offset' ToT_0, 'slope' $\frac{d\mathrm{ToT}}{dn_{p.e.}}$, 'curvature' Γ, 'threshold' $n_{p.e.}^{min}$ and 'saturation' ToT^{max}) used to parameterise the relation between the deposited charge

$n_{p.e.}$ and the ToT of an L0:

$$\mathrm{ToT}(n_{\mathrm{p.e.}}) = f(n_{\mathrm{p.e.}}) \cdot \frac{\mathrm{ToT}^{max}}{\sqrt{f(n_{\mathrm{p.e.}})^2 + (\mathrm{ToT}^{max})^2}}, \tag{4.2}$$

$$f(n_{\mathrm{p.e.}}) = \left(\mathrm{ToT}_0 + n_{\mathrm{p.e.}} \cdot \frac{d\mathrm{ToT}}{dn_{p.e.}}\right) \cdot \left(1.0 - e^{-\Gamma \cdot (n_{\mathrm{p.e.}} - n_{p.e.}^{min})}\right). \tag{4.3}$$

This model is shown in figure 4.3. The used model constants are obtained by a fit to the observed time over threshold distribution of one PMT obtained with in-situ measurements. The default settings correspond to a model of a PMT with nominal characteristics (a *nominal PMT*). This will be referred to as a nominal simulation. The analyses based on ^{40}K decays and neutrinos presented in this thesis are based on a nominal PMT, with a relative efficiency of one for all PMTs. For atmospheric muons, realistic PMT parameters are used, obtained using the procedures discussed in chapter 6.

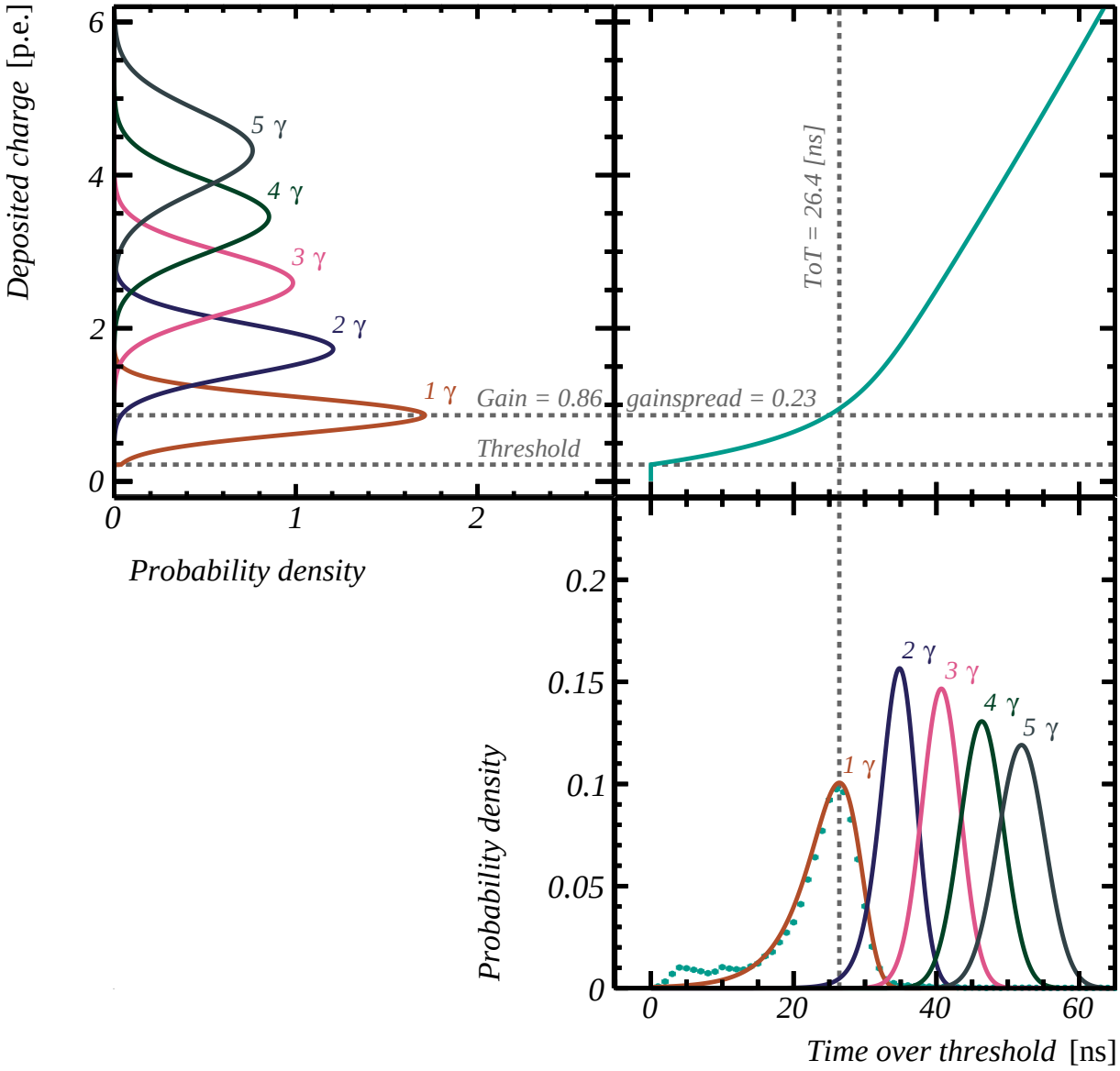


Figure 4.3: *The expected distributions of the deposited charge (top left panel) and the time over threshold (bottom right panel) for one or more detected photons. The relation between the deposited charge and the time over threshold used in the detector simulation (see text) is shown in the top right panel[194]. The gain and gainspread of the PMT are obtained from the single photon ToT distribution of data taken by an arbitrarily chosen PMT (labelled 'data'). The observed hits with time over threshold values below* 10 *ns are presumably caused by prepulses and electronic noise.*

Chapter 5

KM3NeT/ARCA Event Reconstruction

K. Melis, A. Heijboer, and M. de Jong
(on behalf of the KM3NeT Collaboration),
"KM3NeT/ARCA Event Reconstruction Algorithms"
Proceedings of Science (ICRC2017), 950, Aug. 2017.

Event reconstruction algorithms are used in KM3NeT to estimate the properties (and associated errors) of detected neutrinos. The available information for this estimate consists of the distribution of L0s in each triggered event. The procedure to obtain a parameter estimate from the data is called an *estimator*. Several estimators are used in KM3NeT/ARCA, each optimised for one of the event signatures discussed in section 3.3. The concepts of parameter estimation and the event reconstruction algorithms used in KM3NeT/ARCA will be discussed in this chapter.

5.1 Parameter Estimators

An parameter estimator $\vec{y} \rightarrow \tilde{\boldsymbol{x}}$ is a procedure, or rule, to obtain an estimate $\tilde{\boldsymbol{x}}$ of some (unknown) parameters $\boldsymbol{x}_{true}$ from a given set of N data points $\vec{y} = \{y_0, y_1, ..., y_N\}$. Parameter estimators are often designed to have the following desirable properties[195]:

- *Unbiased*: Suppose the data $\vec{y}$ are (randomly) sampled under any true parameters $\boldsymbol{x}_{true}$, then the expectation value of the corresponding estimated parameter values $\tilde{\boldsymbol{x}}$ equals $\boldsymbol{x}_{true}$: $E[\tilde{\boldsymbol{x}}|\boldsymbol{x}_{true}] \equiv \int \tilde{\boldsymbol{x}} \cdot \frac{\partial P(\vec{y})}{\partial \vec{y}}(\boldsymbol{x}_{true}) d\vec{y} = \boldsymbol{x}_{true}$, where $\frac{\partial P(\vec{y})}{\partial \vec{y}}(\boldsymbol{x}_{true})$ is the probability density to obtain a data set $\vec{y}$ under $\boldsymbol{x}_{true}$. If $E[\tilde{\boldsymbol{x}}|\boldsymbol{x}_{true}] \neq \boldsymbol{x}_{true}$, the estimator is called *biased*. It is relatively straightforward to construct an unbiased estimator from a biased estimator if the bias is known.
- *Efficient*: The variance of $\tilde{\boldsymbol{x}}$, given by $\mathrm{Var}[\tilde{\boldsymbol{x}}|\boldsymbol{x}_{true}] \equiv \int \tilde{\boldsymbol{x}}^2 \cdot \frac{\partial P(\vec{y})}{\partial \vec{y}}(\boldsymbol{x}_{true}) d\vec{y} - E[\tilde{\boldsymbol{x}}|\boldsymbol{x}_{\boldsymbol{true}}]^2$, is as small as possible. The efficiency of an estimator gives a measure of its accuracy.

The lower bound of the variance of an unbiased estimator* is given by the Cramér-Rao bound[196, 197]. The corresponding estimator is called the *maximum likelihood*

*A biased estimator, for example 'always return the same value' can be (much) more efficient than the Cramér-Rao bound.

estimate[195].

5.1.1 Maximum Likelihood Estimate

The maximum likelihood estimate corresponds to the value(s) of $\boldsymbol{x}$ at which the likelihood[†] to obtain the data $\vec{y}$, given the set of parameters $\mathcal{L}(\vec{y}|\boldsymbol{x})$ is maximum:

$$\left.\frac{\partial \mathcal{L}(\vec{y}|\boldsymbol{x})}{\partial x}\right|_{\tilde{\boldsymbol{x}}} = 0 \quad \forall \quad x \in \boldsymbol{x}. \tag{5.1}$$

The better the likelihood function $\mathcal{L}(\vec{y}|\boldsymbol{x})$ describes all underlying stochastic processes, the more efficient and less biased the corresponding maximum likelihood estimate is. For complex problems, in which the likelihood function is not known or too complicated to calculate, an approximation of the likelihood function is generally used.

In case the data consists of a set of uncorrelated data points, the likelihood function is given by the product of the likelihood of each individual data point:

$$\mathcal{L}(\vec{y}|\boldsymbol{x}) = \prod_{i=1}^{N} \left[\mathcal{L}(y_i|\boldsymbol{x})\right], \tag{5.2}$$

Log-likelihood For numerical purposes, it is often more convenient to use an alternative, but equivalent, definition of the maximum likelihood estimate, in which the maximum of the logarithm of the likelihood (the log-likelihood) is used:

$$\left.\frac{\partial l(\vec{y}|\boldsymbol{x})}{\partial x}\right|_{\tilde{\boldsymbol{x}}} = 0 \quad \forall \quad x \in \boldsymbol{x}, \tag{5.3}$$

with the log-likelihood function defined as

$$l(\vec{y}|\boldsymbol{x}) \equiv \log\left(\mathcal{L}(\vec{y}|\boldsymbol{x})\right). \tag{5.4}$$

In case of uncorrelated data points, the log-likelihood can be written as:

$$\begin{aligned} l(\vec{y}|\boldsymbol{x}) &= \sum_{i=1}^{N} \left[l(y_i|\boldsymbol{x})\right], \\ l(y_i|\boldsymbol{x}) &\equiv \log\left(\mathcal{L}(y_i|\boldsymbol{x})\right). \end{aligned} \tag{5.5}$$

Start value problem Even with the use of computers, the maximum of the likelihood function can be difficult to find. A few examples are 1). Each evaluation of the (log-)likelihood function requires a large (computational) effort. 2). The (log-)likelihood function is not a continuous function, or 3). It has (many) local minima. In these cases, the problem of finding the maximum likelihood is often split in two steps. First, a first estimate of $\tilde{\boldsymbol{x}}$ is obtained using a simplified model, or by incorporating less information from the data. This *seed* is used in a second stage, in which the maximum likelihood estimate is determined using an iterative scan. A commonly used strategy for the first stage is based on a subclass of maximum likelihood estimates, called χ^2 minimisation.

[†]The *likelihood* function $\mathcal{L}(\vec{y}|\boldsymbol{x})$ (instead of the *probability density function* $P(\vec{y}|\boldsymbol{x})$) is used to indicate that the model parameters are varied, rather than the data. The likelihood function is proportional to $P(\vec{y}|\boldsymbol{x})$, but is not necessarily normalised.

5.1.2 Chi^2-Fits

In case the data consists of N uncorrelated data points, of which each value y_i is Gaussian distributed with a known mean μ_i and standard deviation σ_i, and where each value μ_i is a linear combination of the model parameters $\boldsymbol{x}$

$$\vec{\mu} \equiv \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{bmatrix} = \boldsymbol{A} \cdot \boldsymbol{x}, \tag{5.6}$$

the likelihood function $\mathcal{L}(\vec{y}|\boldsymbol{x})$ is given by a multivariate Gaussian distribution:

$$\begin{aligned} \mathcal{L}(\vec{y}|\boldsymbol{x}) &= \frac{1}{\sqrt{2\pi}^N \cdot |\boldsymbol{\Sigma}|} \cdot \exp\left(-\frac{1}{2} \cdot \chi^2\right), \\ \chi^2 &\equiv (\vec{y} - \vec{\mu})^T \cdot \boldsymbol{\Sigma}^{-1} \cdot (\vec{y} - \vec{\mu}), \end{aligned} \tag{5.7}$$

with

$$\vec{y} \equiv \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \qquad \boldsymbol{\Sigma} \equiv \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 & \cdots & \sigma_1\sigma_N \\ \sigma_1\sigma_2 & \sigma_2^2 & \cdots & \sigma_2\sigma_N \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1\sigma_N & \sigma_2\sigma_N & \cdots & \sigma_N^2 \end{bmatrix}. \tag{5.8}$$

This (log-)likelihood function has a single maximum, given by a minimal χ^2, at coordinates:

$$\tilde{\boldsymbol{x}} = \left(\boldsymbol{A}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{A}\right)^{-1} \cdot \left(\boldsymbol{A}^T \boldsymbol{\Sigma}^{-1} \vec{y}\right). \tag{5.9}$$

A fit of this kind is known as a χ^2-fit. Because of the analytic solution, χ^2-fits are widely used. However, χ^2-fits are particularly sensitive to *outliers*, i.e. data points y_i several standard deviations away from the expected mean μ_i. One way to deal with these outliers is to remove potential outliers from the data used in the fit. A second way is to limit their impact on the fit, for example by the use of an M-estimator.

5.1.3 M-estimators

M-estimators form an alternative to χ^2 fits. They lack an analytic solution, but are more robust, and are generally used as a compromise between a χ^2-fit and a full maximum likelihood estimate[198, 199]. A M-estimator determines the coordinates $\tilde{\boldsymbol{x}}$, at which the 'score'

$$M(\vec{y}|\boldsymbol{x}) \equiv \sum_{i=1}^{N} \left[\rho\left(\frac{y_i - \mu_i(\boldsymbol{x})}{\sigma_i}\right)\right], \tag{5.10}$$

is minimal. The function $\rho(z)$ is called the score function. Various score functions can be used. In general, these functions have a minimum at $z = 0$ (for it to be an unbiased estimator) and flattens towards higher values of z to suppress the influence of outliers. A few examples of commonly used M-estimator score functions are given in table 5.1.

Name	$\rho(z)$
Lorentzian	$\log\left(1+\frac{1}{2}\cdot z^2\right)$
$L_1 - L_2$	$\sqrt{1+\frac{1}{2}\cdot z^2}-1$
Fair	$c^2\cdot\left(\frac{\lvert z\rvert}{c}-\log\left(1+\frac{\lvert z\rvert}{c}\right)\right)$
χ^2	$\frac{1}{2}z^2$

Table 5.1: *A few typical M-estimator score functions[200, 201]. The 'score function' of a χ^2-fit is given for completeness.*

5.2 Shower Reconstruction

An algorithm, called *aashowerfit*[202], is used to reconstruct shower-like events. The estimated parameters are

- The position of the shower. In this, the light emitted by a shower is generally assumed to originate from a single point. The corresponding shower position ($\vec{x}_S$) and time (t_S) are called the shower *vertex*.
- The *visible shower energy* (E_S). It gives an estimate of the total energy of all light emitting particles in an electromagnetic or hadronic shower (see section 3.2.5).
- The direction of the shower ($\hat{d}_S$). The shower direction is related to the average direction of all light-emitting particles, each weighted with the total number of photons they emit.

The reconstruction follows two stages. First, the shower vertex is estimated, which serves as a pivot point for the determination of the shower direction and energy.

5.2.1 Vertex Fit

The first step in the shower reconstruction aims at estimating the shower position $\vec{x}_S$ and interaction time t_S. Background hits from ^{40}K decays can result in outliers. Their contribution is reduced by selecting coincidences of L0s on the same DOM within 20 ns. The time of the first hit of this coincidence $t_i^{\text{1st hit}}$ and the position $\vec{x}_i$ of the corresponding DOM are used in the fit.

A first estimate of the shower vertex (position and time) is obtained by calculating the centre of gravity of the selected coincidences:

$$\begin{bmatrix}\vec{x}_{\text{C.O.G.}}\\ t_{\text{C.O.G.}}\end{bmatrix} = \frac{1}{N}\cdot\sum_{i=1}^{N}\begin{bmatrix}\vec{x}_i\\ t_i^{\text{1st hit}}\end{bmatrix}, \tag{5.11}$$

where N is the number of coincidences. This estimate is used as start value for a (slightly modified) "$L_1 - L_2$" M-estimator (see table 5.1), with score function:

$$M = \sum_{\text{hits}}\left[\sqrt{1+(t_i^{\text{1st hit}} - \hat{t}_{S,i}(\vec{x}_S, t_S))^2}\right], \tag{5.12}$$

where $\hat{t}_{S,i}(\vec{x}_S, t_S)$ is the expected time of unscattered light on a DOM located at $\vec{x}_i$. Using the approximation that all light is emitted from a point source at $\vec{x}_S$ at time t_S, and travels unscattered through the water with the velocity of light in water, $\hat{t}_{S,i}$ is given by:

$$\hat{t}_{S,i}(\vec{x}_S, t_S) = t_S + \frac{n_g}{c} \cdot r_i, \tag{5.13}$$

where r_i is the distance from the DOM to the shower position (see figure 5.1)

$$r_i \equiv |\vec{x}_i - \vec{x}_S| \,. \tag{5.14}$$

The minimum of the M-estimator score function (equation 5.12) is found using the MINUIT minimization algorithm[203].

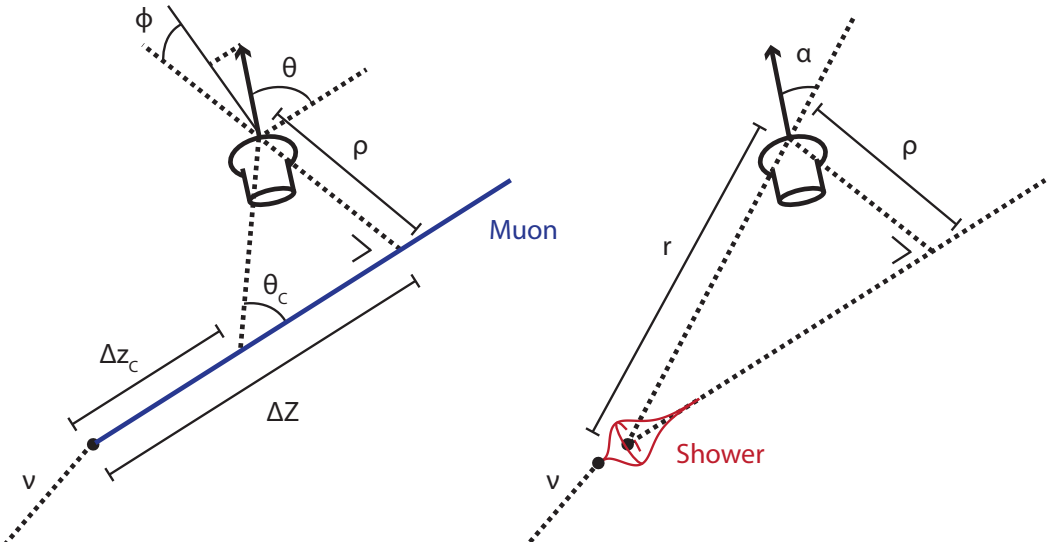


Figure 5.1: *Coordinates used in reconstruction algorithms to describe the position and orientation of a PMT with respect to a muon (**left**) and shower (**right**).*

5.2.2 Direction Fit

The primary source of information used in the direction fit of a shower is given by its angular light emission profile (equation 3.8), which is strongly peaked at the Cherenkov angle. Even though dispersion and scattering blurs this distribution, a similar distribution should still be observed in the distribution of the hit and non-hit PMTs. Since the total amount of emitted light is proportional to the energy of the shower, both the direction and shower energy are reconstructed simultaneously.

In the shower direction reconstruction, all level zero hits are used. A PMT is 'hit' in case it records a L0 within a time window of $[t_{min.}, t_{max.}] \equiv \left[\hat{t}_{S,i} - 100, \hat{t}_{S,i} + 900\right]$ ns, where $\hat{t}_{S,i}$ is the expected arrival time of unscattered light emitted from the estimated shower

vertex. Starting from twelve isotropically distributed directions, the shower direction $\hat{d}_S$ and energy E_S are found (using the MINUIT algorithm) that maximise the likelihood function:

$$\mathcal{L}(\boldsymbol{x}) \equiv \prod_{hitPMTs} \left[P^{\text{hit}}\left(N_i^{\text{p.e.}}\right)\right] \times \prod_{nothitPMTs} \left[P^{\text{not hit}}\left(N_i^{\text{p.e.}}\right)\right], \tag{5.15}$$

where $P^{\text{not hit}}(N)$ is the Poisson probability that the PMT did not record a hit:

$$P^{\text{not hit}}(N) = \frac{\exp(-N) \cdot N^0}{0!} = \exp(-N), \tag{5.16}$$

and

$$P^{\text{hit}}(N) = 1 - P^{\text{not hit}}(N) = 1 - \exp(-N). \tag{5.17}$$

The time-integrated expected number of photoelectrons $N_i^{\text{p.e.}}$ on PMT i is given by the sum of the light emitted by the shower (which is assumed to scale linearly with the shower energy, see section 3.2.5) and the expected background rate from ^{40}K decays:

$$N_i^{\text{p.e.}} \equiv E_S \cdot N_{S,1\text{ GeV}}^{\text{p.e.}}(r_i, \rho_i, \alpha_i) + f_{40K}(1) \cdot \Delta t, \tag{5.18}$$

where $N_{S,1\text{ GeV}}^{\text{p.e.}}$ gives the expected number of photo-electrons from a 1 GeV shower. This function is parameterised by means of a three-dimensional histogram, obtained from simulations. It depends on the distance from the shower to the PMT (r_i), its perpendicular distance to the shower axis (ρ_i) and the orientation of the PMT with respect to the line from the shower to the PMT (given by a single angle α_i), see figure 5.1. The expected background rate from ^{40}K decays is given by the measured singles rate $f_{40K}(1)$ (about 6 kHz per PMT) and the time window (1000 ns) used to define whether a PMT is hit or not.

5.3 Track Reconstruction

For a track-like events, the neutrino interaction position can be located very far away from the instrumented volume. In the KM3NeT track reconstruction, this problem is avoided by reconstructing the following parameters:

- The position $\vec{x}_\mu$ and corresponding time t_μ at which the first detected light was emitted. This position is called the *start position* of the muon.
- The muon energy (E_μ) at the start position $\vec{x}_\mu$.
- The muon direction $\hat{d}_\mu$.

The reconstruction of track-like events consists of a chain of algorithms, part of the JPP software framework[204]. First, a prefit is used to provide a seed, after which the remaining properties are estimated in two stages, both using a maximum likelihood approach.

5.3.1 Prefit

The aim of the prefit is to provide a (set of) first estimate(s) of the muons position $\vec{x}_\mu$, direction $\hat{d}_\mu$ and time t_μ at which the muon passes through $\vec{x}_\mu$. To this extent, the prefit loops over a grid of about 20000 (approximately isotropically distributed) muon directions, sampling the full solid angle with steps of about one degree. For each of the sampled directions, the most likely $\vec{x}_\mu$ and t_μ are determined. The twelve fits with the best quality (the definition of the fit quality will be discussed in this section) are stored for use in a subsequent stage (see section 5.3.2) to estimate the muon parameters with a maximum likelihood fit.

For each of the 20000 muon direction $\hat{d}_\mu$, the most likely muon position $\vec{x}_\mu$ and time t_μ is found using the difference between the hit times of L0s and the expected arrival time $\hat{t}_{\mu,i}$ of unscattered Cherenkov light. This arrival time depends on the position $\vec{x}_i = \{x_i, y_i, z_i\}$ of the PMT (i) and the parameters of the muon via

$$\begin{aligned}
\hat{t}_{\mu,i}(\vec{x}_\mu, t_\mu) &= t_\mu + \frac{1}{c}\cdot\Delta z_{C,i} + \frac{n_g}{c}\cdot\left(\frac{\rho_i}{\sin(\theta_C)}\right), \\
&= t_\mu + \frac{1}{c}\cdot\left(\Delta Z_i - \frac{\rho_i}{\tan(\theta_C)}\right) + \frac{n_g}{c}\cdot\left(\frac{\rho_i}{\sin(\theta_C)}\right), \\
&= t_\mu + \frac{1}{c}\cdot\left(\Delta Z_i + \rho_i\cdot\kappa\right),
\end{aligned} \tag{5.19}$$

where c is the speed of light in vacuum, n_g is the (group) refractive index of seawater, θ_C is the Cherenkov angle and

$$\begin{aligned}
\Delta Z_i &\equiv (\vec{x}_i - \vec{x}_\mu)\cdot\hat{d}_\mu, \\
\Delta z_{i,C} &\equiv \Delta Z_i - \frac{\rho_i}{\tan(\theta_C)}, \\
\rho_i &\equiv \sqrt{(\vec{x}_i - \vec{x}_\mu)^2 - \Delta Z_i^2},
\end{aligned} \tag{5.20}$$

The constant κ is defined as

$$\begin{aligned}
\kappa &\equiv \frac{n_g}{\sin(\theta_C)} - \frac{1}{\tan(\theta_C)}, \\
&= \frac{n_g - 1/n}{\sqrt{1 - 1/n^2}},
\end{aligned} \tag{5.21}$$

Since the phase (n) and group (n_g) refractive indices are approximately equal

$$\kappa \approx \tan(\theta_C) \approx 0.95. \tag{5.22}$$

For each sampled $\hat{d}_{\mu,j}$, a convenient coordinate system is defined in which the z-axis aligns with the (assumed) muon direction. In this rotated coordinate system (with coordinates x', y' and z'), the muon position can be described by three uncorrelated coordinates[‡];

[‡]The origin of this system is not relevant for this problem (generally, it is set to the centre of gravity of the detector).

the time t_μ at which the muon crosses the $z' = 0$ plane and the corresponding position x'_μ, y'_μ. Using these coordinates, equations 5.19 and 5.20 can be written as:

$$\hat{t}_{\mu,i}(x'_\mu, y'_\mu, t_\mu) = t_\mu + \frac{1}{c} \cdot (z'_i + \rho_i \cdot \kappa), \tag{5.23}$$

$$\rho_i \equiv \sqrt{\left(x'_i - x'_\mu\right)^2 + \left(y'_i - y'_\mu\right)^2}, \tag{5.24}$$

where x'_i, y'_i, z'_i give the PMT position in the rotated coordinate system and κ is defined in equation 5.19.

In this coordinate system, the problem can be cast in the form of equation 5.9 by introducing the following quantity[204]:

$$\tilde{\rho}_i(t_\mu) \equiv \frac{c \cdot t_i - z'_i}{\kappa} - \frac{c \cdot t_\mu}{\kappa}, \tag{5.25}$$

which can be seen as an estimate of ρ_i (using the measured hit time) for a given t_μ. By taking the difference of the square of this parameter between two hits i and j:

$$\tilde{\rho}_i^2(t_\mu) - \tilde{\rho}_j^2(t_\mu) \approx \rho_i^2 - \rho_j^2, \tag{5.26}$$

with

$$\kappa^2 \cdot \left(\tilde{\rho}_i^2(t_\mu) - \tilde{\rho}_j^2(t_\mu)\right) = \left(c \cdot (t_i - t_\mu) - z'_i\right)^2 - \left(c \cdot (t_j - t_\mu) - z'_j\right)^2, \tag{5.27}$$

$$= \left(c \cdot t_i - z'_i\right)^2 - \left(c \cdot t_j - z'_j\right)^2 - 2 \cdot \left(c \cdot (t_i - t_j) - \left(z'_i - z'_j\right)\right) \cdot c \cdot t_\mu, \tag{5.28}$$

$$\rho_i^2 - \rho_j^2 = {x'_i}^2 - {x'_j}^2 + {y'_i}^2 - {y'_j}^2 - 2 \cdot \left(x'_i - x'_j\right) \cdot x'_\mu - 2 \cdot \left(y'_i - y'_j\right) \cdot y'_\mu, \tag{5.29}$$

an alternative set of data points is constructed:

$$\vec{y} \equiv \begin{bmatrix} \left(c \cdot t_1 - z'_1\right)^2 / \kappa^2 - \left(c \cdot t_2 - z'_2\right)^2 / \kappa^2 + {x'_1}^2 - {x'_2}^2 + {y'_1}^2 - {y'_2}^2 \\ \left(c \cdot t_1 - z'_1\right)^2 / \kappa^2 - \left(c \cdot t_3 - z'_3\right)^2 / \kappa^2 + {x'_1}^2 - {x'_3}^2 + {y'_1}^2 - {y'_3}^2 \\ \vdots \\ \left(c \cdot t_{N-1} - z'_{N-1}\right)^2 / \kappa^2 - \left(c \cdot t_N - z'_N\right)^2 / \kappa^2 + {x'_{N-1}}^2 - {x'_N}^2 + {y'_{N-1}}^2 - {y'_N}^2 \end{bmatrix}, \tag{5.30}$$

each of which has an expectation value that depends linearly on the muon parameters

$$\vec{\mu} = \boldsymbol{A} \cdot \tilde{\boldsymbol{x}}_\mu = \begin{bmatrix} 2 \cdot (x'_1 - x'_2) & 2 \cdot (y'_1 - y'_2) & -2 \cdot (c \cdot (t_i - t_j) - (z'_1 - z'_2)) \\ 2 \cdot (x'_1 - x'_3) & 2 \cdot (y'_1 - y'_3) & -2 \cdot (c \cdot (t_i - t_j) - (z'_1 - z'_3)) \\ \vdots & \vdots & \vdots \\ 2 \cdot \left(x'_i - x'_j\right) & 2 \cdot \left(y'_i - y'_j\right) & -2 \cdot \left(c \cdot (t_i - t_j) - \left(z'_i - z'_j\right)\right) \end{bmatrix} \cdot \begin{bmatrix} x'_\mu \\ y'_\mu \\ c \cdot t_\mu \end{bmatrix}. \tag{5.31}$$

Consequently, the problem can be solved using a χ^2-fit, in which the best fit muon coordinates x'_μ, y'_μ, t_μ for the given muon direction $\hat{d}_\mu$ are found using the matrix identity (see equation 5.9):

$$\tilde{\boldsymbol{x}}_\mu = \left(\boldsymbol{A}^T \boldsymbol{A}\right)^{-1} \cdot \left(\boldsymbol{A}^T \cdot \vec{y}\right). \tag{5.32}$$

The χ^2-fit outlined above is sensitive to outliers, caused by hits from ^{40}K -decays, bioluminescence and by photons following a strongly scattered path. In order to reject

these, a series of χ^2-fits is done for each sampled direction, each with a selection of n out of the N total number of hits in the event. The fit with the highest fit quality q, defined as

$$\begin{aligned} q &= -\chi^2 + N_{d.o.f.} \cdot \sigma^2, \\ \chi^2 &\equiv \sum_{i=1}^{n} \left[\left(t_i - t_{\mu,i}(\vec{i}) \right)^2 \right], \\ N_{d.o.f.} &\equiv n - 3, \end{aligned} \tag{5.33}$$

is selected as best estimate of x'_μ, y'_μ, t_μ for the given $\hat{d}_\mu$. The constant σ (typically set to a value of three) is included to favour fits with more data points. In case $N < 9$, all possible permutations of $n \in \{N, N-1, ...4\}$ hits are tested, where $n = 4$ is the minimum number of required hits for a constrained fit. In case $N \geq 9$, an iterative procedure is followed in which the hit with the largest difference between measured and fitted hit time is removed from the sample, until $n = 4$ hits remain or the quality parameter q does not increase by the rejection of the next hit.

The twelve prefit solutions with the highest quality parameter (with χ^2 and $N_{d.o.f.}$ as defined in equation 5.33):

$$Q \equiv N_{d.o.f.} - \frac{1}{4} \cdot \left(\frac{\chi^2}{N_{d.o.f.}} \right), \tag{5.34}$$

are used as initial values for a maximum likelihood estimate of the track parameters.

5.3.2 Likelihood Trajectory Fit

A more accurate estimate of the geometrical properties (time, position and direction) of the muon is obtained by a maximum likelihood approach. Using the twelve prefit seeds, the maximum likelihood is found using the Levenberg-Marquandt algorithm [205, 206].

The data $\vec{y}$ used in the likelihood consists of the first L0 on each PMT recorded within a time window of $[t_{min.}, t_{max.}] \equiv \left[\hat{t}_{\mu,i} - 50, \hat{t}_{\mu,i} + 450 \right]$ nanoseconds, where $\hat{t}_{\mu,i}$ is the expected arrival time of direct Cherenkov light (see equation 5.19) from the prefit muon. The likelihood function is based on the time-dependent probability to observe the hits at their respective times:

$$\mathcal{L}(\vec{y}|\boldsymbol{x}) \equiv \prod_{hit PMTs} \left[\frac{P^{\text{hit}}\left(\dot{n}_{\mu,i}(t_i|\boldsymbol{x}_\mu)\right) \cdot P^{\text{not hit}}\left(n_{\mu,i}(t_i|\boldsymbol{x}_\mu)\right)}{P^{\text{hit}}\left(N_{\mu,i}(\boldsymbol{x}_\mu)\right)} \right], \tag{5.35}$$

in which the terms in the numerator correspond to the probability that the first L0 is detected in a time window of 1 ns around t_i (e.g. equation 5.17), and the probability that no L0 is detected on the same PMT before, respectively. The denominator corresponds to the total probability that the PMT is hit in the time window. The variables $\dot{n}_{\mu,i}(t_i|\boldsymbol{x}_\mu)$, $n_{\mu,i}(t_i|\boldsymbol{x}_\mu)$ and $N_{\mu,i}(\boldsymbol{x}_\mu)$ relate to the (time-dependent) number of expected photoelectrons on PMT i for the given muon parameters $\boldsymbol{x}_\mu$. These quantities are evaluated using (five-dimensional) interpolation of precalculated tabulated values. It accounts for the light induced by the muon (see section 3.2.5) and from ^{40}K decays, the propagation of the light

through water, as well as the PMT efficiency and transit time distribution. The evaluation of $\dot{n}_{\mu,i}(t_i|\boldsymbol{x}_\mu)$, $n_{\mu,i}(t_i|\boldsymbol{x}_\mu)$ and $N_{\mu,i}(\boldsymbol{x}_\mu)$ is discussed in appendix D.

At this stage of the reconstruction, the muon is assumed to have an infinite range, and a constant energy of 1 TeV along its track. Under this assumption, the parameter space of the fitted muon properties reduces from seven to five dimensions (two for the direction and three for an arbitrary position along its track, E_μ and the start position are not reconstructed). This assumption makes the fit faster to evaluate and more robust.

5.3.3 Interaction Vertex and Energy

The muon energy E_μ and the start position of the muon are estimated in the final stage of the reconstruction chain. This position and the corresponding muon energy are defined at the first point along its trajectory where detected light is emitted. Wen the muon is produced outside of the instrumented volume, this position may differ significantly from the neutrino interaction vertex and corresponding muon energy.

The procedure to estimate the muon position is relatively straightforward: For each DOM, the position $\Delta z_{i,C}$ (defined in equation 5.20) along the fitted muon corresponding to the point where unscattered Cherenkov light must have been emitted in order to be detected by the DOM is determined. Subsequently, the number of L0s within $[t_{min.}, t_{max.}] \equiv \left[\hat{t}_{\mu,i} - 50, \hat{t}_{\mu,i} + 450\right]$ nanoseconds on each DOM is counted. The start position of the muon is taken to be either the lowest value of $\Delta z_{i,C}$ of a DOM with at least two recorded L0s in this time window, or the point where two DOMs with adjacent values of $\Delta z_{i,C}$ both recorded one L0 in this time window. The probability that one of these criteria is met by the background rate from ^{40}K decays is less than approximately 10^{-3}.

The energy of the muon at the estimated start position is estimated from the distribution of all hit and non-hit PMTs with $\Delta z_{i,C}$ larger than the fitted start position. In this case, a PMT is 'hit' in case it detects one or more L0s in a time window of $[t_{min.}, t_{max.}] \equiv \left[\hat{t}_{\mu,i} - 50, \hat{t}_{\mu,i} + 450\right]$ nanoseconds. For a given muon energy E_μ, the time-integrated expected number of photoelectrons on each PMT is evaluated using the functions discussed in appendix D. The muon energy corresponding to the maximum of the likelihood function, defined as:

$$\mathcal{L}(\vec{y}|\boldsymbol{x}) \equiv \prod_{\text{hit PMTs}} \left[P^{\text{hit}}\left(N_i(\boldsymbol{x}_\mu)\right)\right] \times \prod_{\text{not hit PMTs}} \left[P^{\text{not hit}}\left(N_i(\boldsymbol{x}_\mu)\right)\right], \tag{5.36}$$

is taken as an estimate of the muon energy.

5.3.4 Single Line Track Reconstruction

For a period of about half a year, until the deployment of a second line, the KM3NeT/ARCA detector consisted of a single DU (see section 3.10). The non-operational acoustic positioning system (see section 3.5), motivates the use of a detector geometry in which all DOMs are assumed to be aligned vertically. Because of the resulting rotational symmetry, the fit is insensitive to the azimuth angle ϕ_μ of the track.

The prefit routine used in the track reconstruction chain (see section 5.3.1) was adapted

to reconstruct tracks using the data taken during this period.[§]. In this fit, the azimuth angle ϕ_μ is set to zero and the coordinate system is shifted such that the DU is located at $x = y = 0$. The zenith angle θ_μ is sampled with steps of 0.1 ° . The prefit with the highest fit quality (equation 5.34) is taken as best estimate of the track parameters.

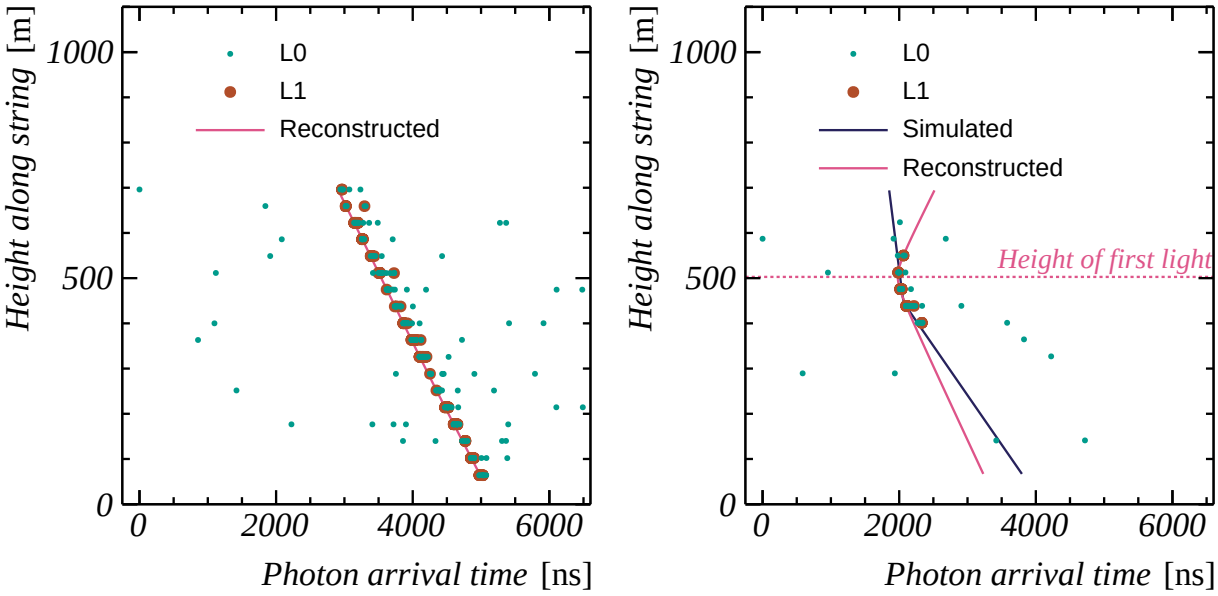


Figure 5.2: *Two examples of events recorded by a single detection unit. The event in the* ***left*** *figure is obtained from the experiment, the event in the* ***right*** *figure from simulations. Shown are the L0s and L1s, as well as the expected arrival time of the best-fit muon hypothesis. For the event from simulations, the expected arrival time of the simulated muon is shown as well. The horizontal axis corresponds to the time elapsed since the first L0 of the triggered event. The reconstructed muon shown in the left figure is moving almost vertically down, in the right figure it moves approximately horizontally.*

With $x = y = 0$, the data used in each fit consists of two observables; the height z_i along the DU and the time t_i of the first L1 in the event detected by DOM i. In figure 5.2, the relationship between these two observables is illustrated. The expected arrival time of unscattered photons $\hat{t}_\mu(z)$ emitted by the best fit track is shown as well, which depends on the height z along the DU and the estimated track parameters (x_μ, y_μ, z_μ, t_μ and θ_μ) via the following relations:

$$\begin{aligned} \hat{t}_\mu(z) &= t_\mu + \frac{1}{c} \cdot (\Delta Z + \rho \cdot \kappa)\,, \\ \Delta Z &= (z - z_\mu) \cdot \cos(\theta_\mu) - x_\mu \cdot \sin(\theta_\mu), \\ \rho^2 &= (\vec{x} - \vec{x}_\mu)^2 - (\Delta Z)^2 \\ &= ((z - z_\mu) \cdot \sin(\theta_\mu) + x_\mu \cdot \cos(\theta_\mu))^2 + y_\mu^2. \end{aligned} \tag{5.37}$$

The estimated muon position $\vec{x}_\mu = [x_\mu, y_\mu, z_\mu]$ has no particular physical meaning. In

[§]A second, alternative routine was developed in parallel, but turned out to be inferior in terms of speed and performance, and will therefore not be discussed in further detail.

the following, some derived quantities describing the track are discussed.

Distance of closest approach Since ϕ_μ is set to zero, the reconstructed track travels through the plane given by the reconstructed y_μ. The closest distance between the reconstructed track and the axis through the DU (at $x = y = 0$):

$$\rho_{\text{clos.}} = \begin{cases} |y_\mu| & \text{if } 0 < \theta_\mu < 180° \\ \sqrt{x_\mu^2 + y_\mu^2} & \text{else} \end{cases}, \tag{5.38}$$

is called the *distance of closest approach*.

Height of closest approach The height $z_{\text{clos.}}$ where the track is closest to the axis at $x = y = 0$ is called the *height of closest approach*. It is defined only if $0 < \theta_\mu < 180°$, in which case it is calculated by solving $\rho^2 = \rho_{\text{clos.}}^2$ for z:

$$\begin{aligned} y_\mu^2 &= \left((z_{\text{clos.}} - z_\mu) \cdot \sin(\theta_\mu) + x_\mu \cdot \cos(\theta_\mu)\right)^2 + y_\mu^2, \\ 0 &= (z_{\text{clos.}} - z_\mu) \cdot \sin(\theta_\mu) + x_\mu \cdot \cos(\theta_\mu), \\ z_{\text{clos.}} &= \begin{cases} z_\mu + \frac{x_\mu}{\tan(\theta_\mu)} & \text{if } 0 < \theta_\mu < 180° \\ \text{undefined} & \text{else} \end{cases}. \end{aligned} \tag{5.39}$$

Height of first light The height $z_{1\text{st}}$ along $x = y = 0$ where Cherenkov light from the muon can be first observed is called the *height of first light*. It is given by solving:

$$\begin{aligned} 0 &= \left.\frac{\partial \hat{t}_\mu(\vec{x})}{\partial z}\right|_{z_{1\text{st}}} \quad \Rightarrow \quad 0 = \cos(\theta_\mu) + \kappa \cdot \frac{\partial \rho}{\partial \rho^2} \cdot \frac{\partial \rho^2}{\partial z}, \\ 0 &= \cos(\theta_\mu) + \frac{\kappa}{2\rho} \cdot 2 \cdot \left((z_{1\text{st}} - z_\mu) \cdot \sin(\theta_\mu) + x_\mu \cdot \cos(\theta_\mu)\right) \cdot \sin(\theta_\mu), \\ \rho^2 \cdot \cos^2(\theta_\mu) &= \kappa^2 \cdot \sin^2(\theta_\mu) \cdot \left((z_{1\text{st}} - z_\mu) \cdot \sin(\theta_\mu) + x_\mu \cdot \cos(\theta_\mu)\right)^2, \\ &= \kappa^2 \cdot \sin^2(\theta_\mu) \cdot \left(\rho^2 - y_\mu^2\right), \\ \rho^2 &= y_\mu^2 \cdot \left(1 - \frac{\cos^2(\theta_\mu)}{\kappa^2 \cdot \sin^2(\theta_\mu)}\right), \\ \left((z_{1\text{st}} - z_\mu) \cdot \sin(\theta_\mu) + x_\mu \cdot \cos(\theta_\mu)\right)^2 &= \frac{-y_\mu^2}{\kappa^2 \cdot \tan^2(\theta_\mu)}. \end{aligned} \tag{5.40}$$

This quadratic formula is relatively straightforward to solve.

5.4 Flavour Identification and Event Selection

In KM3NeT, both the shower and track reconstruction algorithms are applied to each event. The output of these algorithms (the most likely event signature parameters, error estimates of these parameters and fit quality parameters) is used in a third estimator which aims at the identification of the flavour and interaction channel of the neutrino.

This is called *flavour identification*. In KM3NeT, flavour identification is done primarily by means of machine learning algorithms[207].

The output of the event reconstruction algorithms can also be used to reject misreconstructed events and events originating from the atmospheric muon background. More stringent criteria can be applied to select a subset of the events with a very good angular and energy resolution. Such a selection may be beneficial for studies in which the (angular) resolution of the subsample of selected events is the leading factor in the sensitivity. Such event selection criteria are generally based on the quality of the fit, the estimated angular accuracy and the distance from the reconstructed interaction position to the detector.

5.5 Performance of the Reconstruction

The performance of the reconstruction algorithms has been quantified by means of simulations. The shower reconstruction has been applied to shower-like events (charged-current $\nu_e + \bar{\nu}_e$ and neutral current neutrino interactions), and the track reconstruction to events from charged-current $\nu_\mu + \bar{\nu}_\mu$ interactions.

Event Selection As discussed in the previous section, only a fraction of all reconstructed events are selected for preceding analyses. The criteria of this selection are generally optimised to maximise the sensitivity of the analysis. In this section, selection criteria discussed in the letter of intent[135] will be used.

For tracks, the maximum likelihood $\mathcal{L}$ is required to be $\mathcal{L} \geq e^{60} \cdot \mathcal{L}_{40K}$, where $\mathcal{L}_{40K}$ is the likelihood that the event is caused by light from ^{40}K decays only (i.e. $E_\mu = 0$). Additionally, the estimated accuracy (defined at a 1σ level) of the fitted direction has to be less than 0.1 ° . Shower-like events are selected if at least 2200 L0s are used in the direction and energy fit, and the reconstructed position is contained in a cylindrical volume extending from the seabed to a height of 600 m above the seabed, and of radius 500 m around the centre of the detector. The fraction of triggered events that pass these selection criteria is shown in figure 5.3.

Resolution The selected fraction of reconstructed charged current $\nu_e + \bar{\nu}_e$ events increases up to a neutrino energy of about 10^4 GeV. Above this energy, the number of events that pass the first selection criterion (at least 2200 L0s used in the fit) still increases with energy, but the number of triggered events outside of the containment volume increases even more. A similar effect is seen for shower-like events from neutral current interactions, but at a higher neutrino energy. This shift is explained by the fact that hadronic showers typically emit less light. For track-like events (charged-current $\nu_\mu + \bar{\nu}_\mu$ interactions), the fraction of selected events increases with energy, up to a neutrino energy of about 10^6 GeV.

The accuracy of the reconstruction algorithms is characterised using the following quantities:

- The angular resolution, parameterised by means of three angles: The angle between the directions of the neutrino and the reconstructed muon or shower ($\Delta\theta_{\text{reco.},\nu}$), between the neutrino and the true muon/shower ($\Delta\theta_{\mu,\nu}$ and $\Delta\theta_{S,\nu}$) and the angle

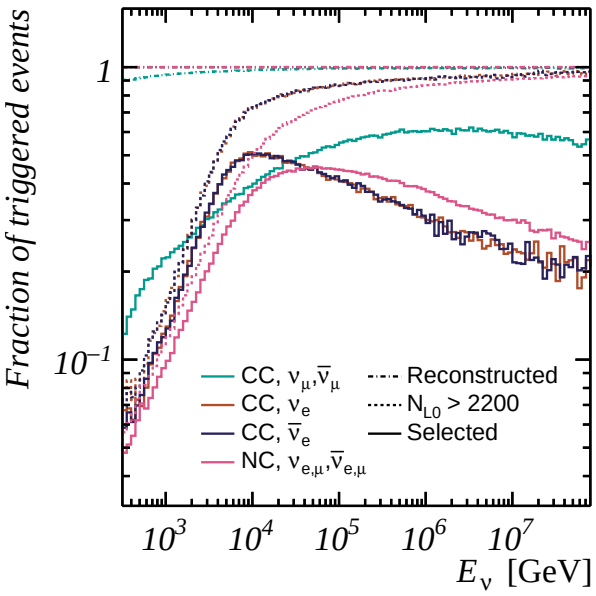


Figure 5.3: *The reconstructed and selected fraction of the total number of triggered events, for simulated charged-current $\nu_\mu + \bar{\nu}_\mu$, charged-current ν_e and $\bar{\nu}_e$ and neutral current $\nu_e + \bar{\nu}_e + \nu_\mu + \bar{\nu}_\mu$ events. An event is reconstructed in case at least one solution is found by the reconstruction algorithms. It is selected only in case the reconstructed quantities pass the quality conditions discussed in the text.*

between the true muon and the reconstructed muon direction ($\Delta\theta_{\text{reco.},\mu}$). The direction of a shower is defined as the average direction of all of its particles, weighted by the energy of each.

- The energy resolution, defined as the ratio between the reconstructed energy ($E_{\text{reco.}}$) and the true neutrino energy (E_ν), true muon energy (E_μ) and/or the true shower energy (E_S). Additionally, a corrected energy $E^{\text{corr.}}_{\text{reco.}}$ is introduced to account for biases in the reconstructed energy $E_{\text{reco.}}$. The corrected energy is defined such that the median of $E^{\text{corr.}}_{\text{reco.}}$ equals the (true) neutrino energy E_ν.

- The longitudinal distance from the reconstructed shower or muon start position $\vec{x}_{\text{reco.}}$ to the true neutrino interaction position, defined as $\Delta_{\text{long.}} \equiv (\vec{x}_{\text{reco.}} - \vec{x}_\nu) \cdot \hat{d}_\nu$, where $\vec{x}_\nu$ is the true neutrino interaction position and $\hat{d}_\nu$ is the true neutrino direction.

- The perpendicular distance, defined as $\Delta_{\text{perp.}} \equiv \sqrt{(\vec{x}_{\text{reco.}} - \vec{x}_\nu)^2 - (\Delta_{\text{long.}})^2}$.

The distributions of the reconstructed quantities, obtained from the selected set of events, are shown in figures 5.4 and 5.5.

Discussion A few features of the reconstruction performance distributions are noticeable:

- All distributions of reconstructed neutral-current events are very similar to those of reconstructed charged-current electron (anti)neutrino events, but the features appear shifted along the energy axis. This shift can be explained by the fraction of the neutrino energy carried away by the neutrino in neutral-current events. This contrasts with charged-current ν_e and $\bar{\nu}_e$ events, where the outgoing electron/positron causes an additional electromagnetic shower. This effect is particularly observable in the energy resolution of the reconstructed neutrino energy; the true shower energy is reconstructed with an accuracy of a few tens of percents in both cases, but the reconstructed neutrino energy differs much more in the case of neutral current

events. The same effect is observed in the distribution of the angle between the neutrino direction and the direction of the shower.

- At energies below approximately 10 TeV, the number of recorded coincident hits in shower-like events is too low for an accurate position fit. From about 10 TeV up to neutrino energies of approximately 10^7 GeV, the shower position is estimated with an accuracy of about one metre in the perpendicular plane. In the longitudinal distance, a systematic bias of the reconstructed position is present. One of the contributions to this bias comes from the elongation of the shower emission profile with energy (section 3.2.5). However, this does not explain the full effect. Another contribution comes from the fact that the hits used in the position fit (the first hit of a coincidence on each DOM) are biased: a DOM positioned close to the shower is likely to record a large number of hits and thus assumes an early arrival time of the light, while DOMs positioned farther away will mainly record scattered (and thus late) hits.[208]. This effect may also be the cause of the worsened performance of the shower reconstruction at energies above 10^7 GeV.

- From about 10 TeV up to neutrino energies of approximately 10^7 GeV, the shower angular accuracy is of the order of a few degrees, and the accuracy of the reconstructed neutrino energy (in charged-current ν_e and $\bar{\nu}_e$ neutrino events) is about a few tens of percents. At neutrino energies below 10 TeV and above 10^7 GeV, the (relatively) inaccurate position fit affects the accuracy of the reconstruction of the direction and energy.

- The angular accuracy of reconstructed muons is about 0.1 ° . This resolution is about ten times better than the angular resolution of shower events, which can be explained by the range of the muon (on average more than a kilometre for muon energies above 1 TeV), providing these events a long 'lever arm'. For neutrino energies below approximately 10 TeV, the intrinsic muon-neutrino angle dominates angular resolution of the reconstructed neutrino direction.

- For muons with a long range, the probability of the hadronic shower being observed in the detector decreases. This causes a larger uncertainty of the determined neutrino energy, since the muon loses energy before reaching the reconstructed start position.

- In case the (true) muon start position is within the instrumented volume, it is generally reconstructed with an accuracy of a few metres. However, if the muon is produced in an interaction vertex (far) outside of the instrumented volume, the estimated start position of the muon is strongly biased. These two cases are reflected in the distribution of the longitudinal accuracy (figure 5.5, top left); the 5% and 16% percentiles are relatively close to zero, while the median and higher percentiles reflect the distribution of the range of the muon, and are therefore at significant larger longitudinal distances.

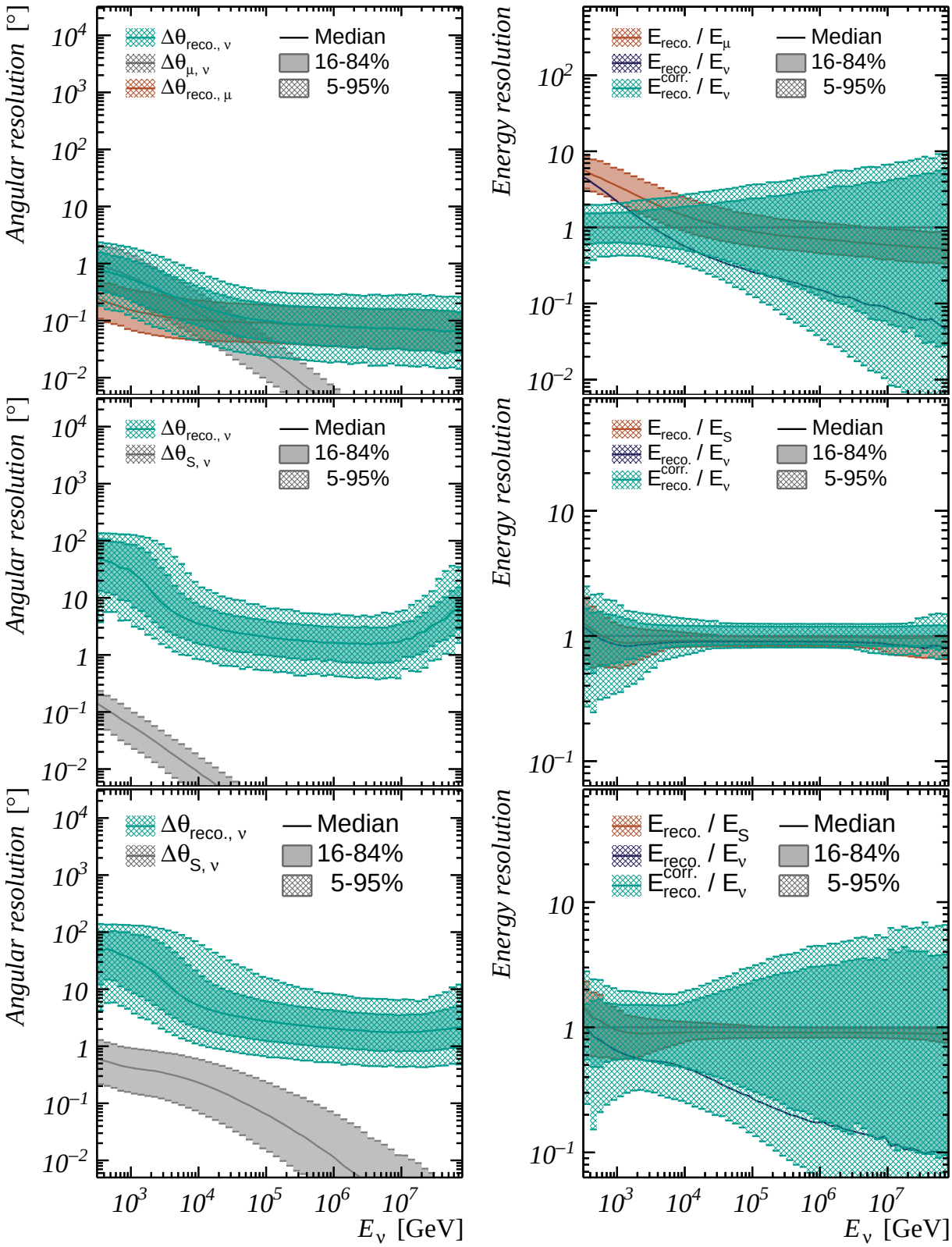


Figure 5.4: *The angular and energy resolution for charged-current $\bar{\nu}_\mu + \nu_\mu$ (**top**), charged-current $\bar{\nu}_e + \nu_e$ (**middle**) and neutral current $\nu_e + \nu_\mu$ (**bottom**) interactions. Event selection criteria as discussed in the text are applied. Please note the different vertical energy scale for charged-current $\bar{\nu}_\mu + \nu_\mu$ events.*

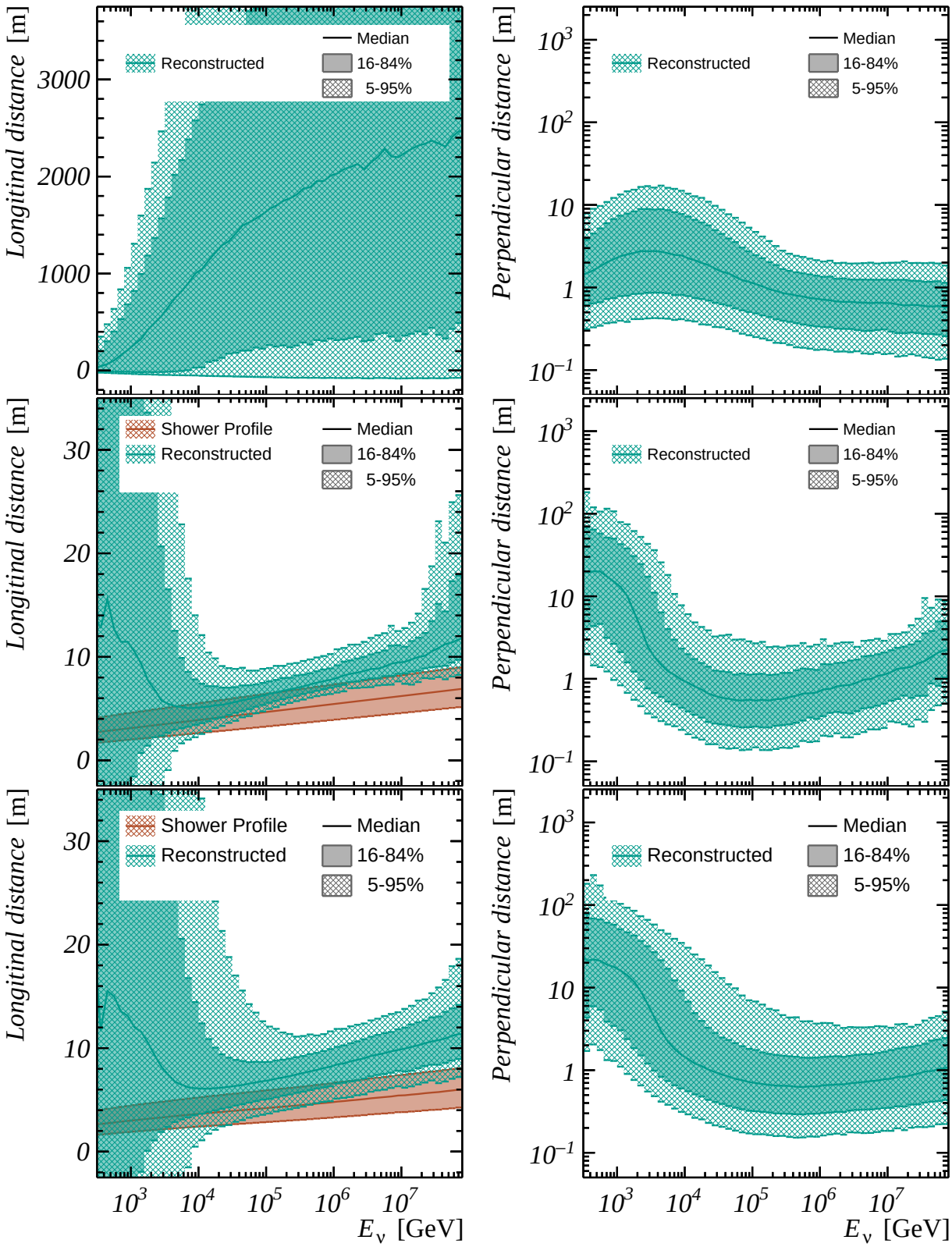


Figure 5.5: *Accuracy of the reconstructed position, obtained from simulations. The longitudinal and perpendicular component (see text) with respect to the true neutrino properties are given for charged-current $\bar{\nu}_\mu + \nu_\mu$ (**top**), charged-current $\bar{\nu}_e + \nu_e$ (**middle**) and neutral current $\nu_e + \nu_\mu$ (**bottom**) interactions. For shower-like events, the median (and 16-84% percentiles) of the light emission profile is given, which includes the distribution of the shower energy for a given neutrino energy (see section 3.2.5 and appendix A).*

Part II

First Data

Abstract

Data-taking with the KM3NeT detector(s) started on December 2015, shortly after the deployment of the first detection unit. Three calibration techniques (the 'in-situ' calibration) have been developed to calibrate the detector. These techniques exploit the light induced by the ^{40}K -decay and atmospheric muon backgrounds. The main advantages of these techniques is that the detector can be calibrated without the need of dedicated calibration 'runs', high statistics are available, and the performance of the detector can be monitored over long periods of time.

The in-situ calibration is discussed in this part. Sub-nanosecond timing accuracy is validated by comparing the results with other calibration techniques. The PMT efficiencies (determined with an accuracy of a few percent) are used in realistic simulations of the detector. These simulations are cross-checked with data, leading to the first scientific measurement with KM3NeT; the depth dependence of the atmospheric muon flux. Having shown that the detector works as expected, candidate neutrino events have been identified in the first data.

Chapter 6

Inter-PMT Calibration

K. Melis (on behalf of the KM3NeT Collaboration),
"In-Situ Calibration of KM3NeT"
Proceedings of Science (ICRC2017), no. 1059, Aug. 2017.

One KM3NeT/ARCA building block has 64170 PMTs, each having different characteristics. A laboratory-based calibration of all PMTs can thus be a tedious job, requiring a large effort from (presumably) fellow PhD students. Fortunately, the detector can be calibrated in-situ, exploiting the presence of light from ^{40}K decays in the seawater. As the rate of ^{40}K decays is approximately constant over time, an advantage of this calibration procedure is that the detector performance can be monitored over long periods of time, unaffected by changes in the detector setup (such as the PMT high voltage settings), and without the need of dedicated calibration runs. The algorithms developed to calibrate the PMTs in-situ, and some results obtained with the first DUs KM3NeT/ARCA, will be discussed in this chapter.

6.1 Principle

In this section, a procedure will be discussed to calibrate the relative PMT time offset $t_{0,i}^{PMT}$, transit time spread TTS_i and relative efficiency ϵ_i of each PMT in the KM3NeT detectors, exploiting the light from the background of ^{40}K decays. The definitions of the relative PMT time offset and transit time spread are given in sections 3.5.1 and 3.8, the relative PMT efficiency is defined as the ratio between the efficiency of a PMT and the efficiency of a PMT with nominal characteristics, as described in sections 3.5.1 and 4.5. The PMT quantum efficiency, collection efficiency and threshold, as well as the effect of shadowing are described by the fitted relative PMT efficiency. The determined relative PMT time offsets are used in the data acquisition system to calibrate the data (section 3.7.4). The fitted relative efficiencies and transit time spreads are used as inputs for the detector simulations (section 4.5).

Model Simulations (see section 4.3) have been used to obtain the distributions of the time difference in the arrival times of two (or more) photons and the resulting L0s from

the same ^{40}K decay, detected by any combination of two or more nominal PMTs in a DOM. These distributions are given in figures 6.1 (left) and 6.1 (right).

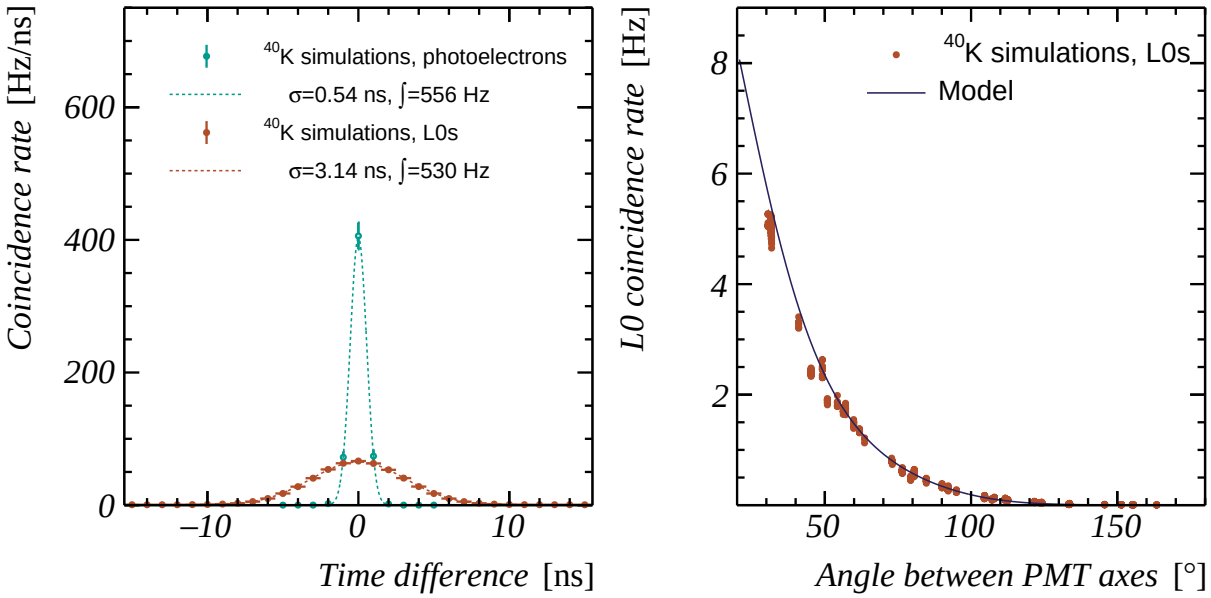


Figure 6.1: ***Left:*** *The expected distribution of the time difference between two photons caused by a single simulated ^{40}K decay, detected on two PMTs in a DOM. The distribution of the time difference between the resulting L0s (i.e. after simulation of the PMT quantum efficiency and transit time) is shown as well. The simulations do not take random coincidences from multiple ^{40}K decays into account.* ***Right:*** *The simulated total coincidence rate as a function of the angle between the PMT axes. The points correspond to the total coincidence rate of pairs of nominal PMTs in simulations, the line gives a fit to simulations (see text). Coincidences from by multiple ^{40}K decays are not included in the simulations.*

As a result of the (isotropic) distribution of ^{40}K decays around the DOM, the time difference is on average zero, with a spread of 0.54 ns (3.14 ns) before (after) simulation of the detector response. The total coincidence rate of L0s (i.e. after detector simulation) R on a pair of PMTs i, j depends on the *opening angle* between the PMTs $\theta_{i,j}$. This relation is parameterised by

$$R(\theta) = \exp\left[1.69 \cdot \cos^3(\theta) \; -1.36 \cdot \cos^2(\theta) + 3.17 \cdot \cos(\theta) - 1.07\right]. \tag{6.1}$$

With these observations in mind, the expected distribution of L0 coincidences $R_{i,j}(\Delta t)$ for a pair of PMTs i and j as function of the time difference Δt between the L0s can be modelled as a Gaussian distribution. The mean of this distribution depends on the difference in relative PMT time offsets ($t_{0,i}$ and $t_{0,j}$), the width is related to the squared sum of the (mean) transit time spreads (TTS_i and TTS_j) and the total coincidence rate

is proportional to the product of the PMT efficiencies (ϵ_i and ϵ_j):

$$\begin{aligned} R_{i,j}(\Delta t) &= \frac{\epsilon_i \cdot \epsilon_j \cdot R(\theta_{i,j})}{\sqrt{2\pi}\sigma_{i,j}} \cdot \exp\left[-\frac{(\Delta t - \mu_{i,j})^2}{2\sigma_{i,j}^2}\right], \\ \mu_{i,j} &= t_{0,i} - t_{0,j}, \\ \sigma_{i,j}^2 &= TTS_i^2 + TTS_j^2 + 0.54^2, \end{aligned} \tag{6.2}$$

where Δt is the time difference between the detected L0s.

Fit procedure In order to estimate the parameters ($t0$, TTS and ϵ) of all 31 PMTs in a DOM, the model given by equation 6.2 is fitted to the L0 time difference distributions of all $31 \cdot 30/2 = 465$ pairs of PMTs in a DOM. As each parameter affects the model of 30 coincidence rate distributions, all distributions are fitted simultaneously, with 92 free parameters in the fit; 31 PMT relative efficiencies, 31 transit time spreads and 30 relative PMT time offsets. The remaining relative PMT time offset is constrained by the fact that the average of the 31 relative PMT time offsets is defined to equal zero.

The model (equations 6.2 and 6.1) does not account for coincidences caused by uncorrelated ^{40}K decays. This random coincidence rate is estimated from the tails of the distributions ($19 \leq |\Delta t| \leq 25$ ns) and subtracted from the measured distributions prior to the fit.

6.2 Results

The inter-PMT calibration procedure has been applied to KM3NeT/ARCA L1 timeslice data, taken between December 23, 2016 12:00:48 PM and January 13, 2017 6:00:18 AM (one DU operational). This period starts after an in-situ tuning of the PMT high voltage settings (see section 3.7.2). During this three-week period, the KM3NeT/ARCA detector was switched off for approximately 20.5 hours, to protect the electronics at the shore-station against a thunderstorm. After exclusion of calibration runs (9 minutes), a total of 470 hours 48 minutes and 9 seconds of data was collected with 99.9 % data-collection efficiency. The results and some cross-checks will be discussed in this section. Conclusions will be drawn in section 6.4.

Time corrected coincidence rate distribution In figure 6.2 (left), the L0 time difference distributions are shown for a typical pair of adjacent PMTs on an arbitrarily chosen DOM, before and after the inter-PMT time calibration is applied to the data (by subtracting the PMT time offsets from the raw L0 times). The fitted random coincidence rate is indicated by the shaded area. The coincidence rate distribution is, as expected, symmetric around zero.

Cross-check of time calibration with LED nanobeacons For all downward looking PMTs in a DOM (rings A-D), an alternative and independent time calibration can be obtained from the arrival time distribution of light emitted by the nanobeacon flasher on the DOM below (for details, please refer to [30, 31]). In figure 6.2 (right), the determined PMT

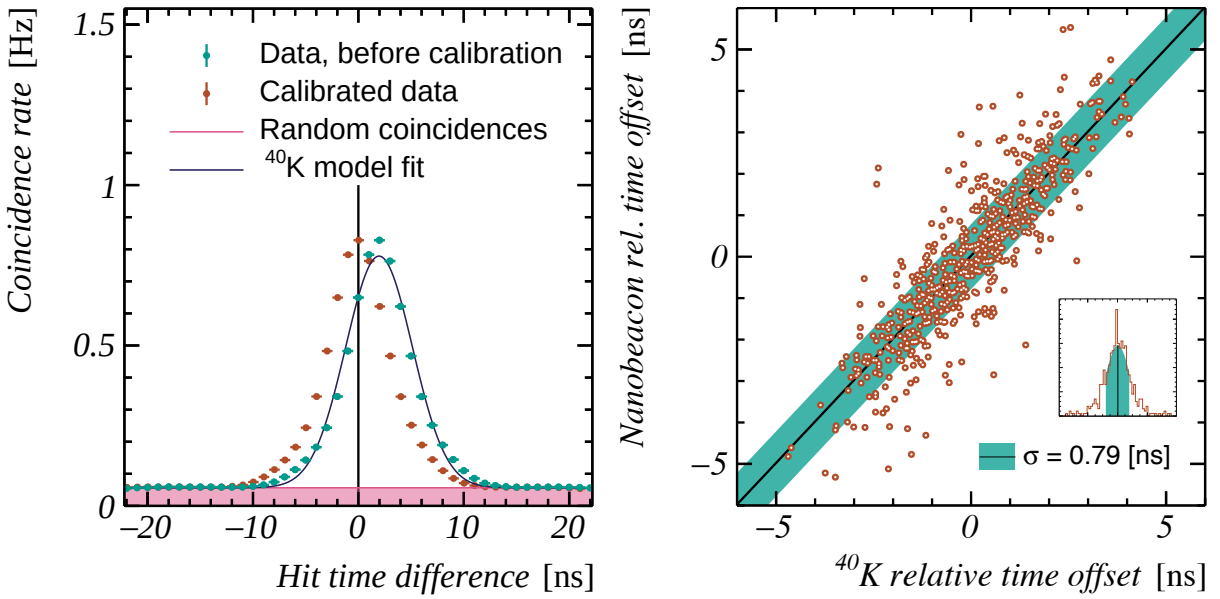


Figure 6.2: ***Left:*** *The time difference distributions on two arbitrarily chosen neighbouring PMTs, obtained from (un)calibrated data and simulations. The uncalibrated data is based on the raw L0 times. The calibrated data show the hit time difference after subtracting the PMT time offsets, determined using the inter-PMT calibration procedure described in the text. The time-uncorrelated random coincidence background is fitted and added to the fitted model.* ***Right:*** *Comparison between the best fit PMT time offsets obtained from the calibration routine using ^{40}K decays and those obtained from the arrival time distributions of light emitted by a LED nanobeacon one DOM below [30, 31]. Only downward looking PMTs (rings A-D) are shown, as upward looking PMTs record only scattered light from the LED nanobeacon. The panel in the top left corner shows the distribution of the difference between the relative PMT time offsets. This distribution is well fitted by a Gaussian, with mean* 0.024 *ns and standard deviation* 0.79 *ns.*

time offsets of both methods are compared with each other. The difference between the relative PMT time offsets determined with both methods follows a Gaussian distribution, with a standard deviation of 0.79 ns.

Fitted relative PMT efficiencies The distribution of the fitted relative PMT efficiencies is shown in figure 6.3 (left). A set of PMTs with systematically lower PMT efficiencies can be identified below $\epsilon \leq 1$. All of these PMTs are situated close to the structure connecting the DOM to the VEOC, which effectively shadows the surrounding PMTs. As the DOM support structure is not included in the simulations of ^{40}K decays (see section 4.3.1), this effect is neither included in the model used to fit the PMT efficiencies (e.g. equation 6.1). The relative efficiencies of all PMTs with an unobstructed view (referred to as *unobstructed PMTs*) follow a Gaussian distribution with a mean of 1.18 and a spread of 7 %.

Efficiency corrected L0 coincidence rate In figure 6.3 (right), the total measured coincidence rate for all PMT pairs in an arbitrarily chosen DOM is shown. The contribution

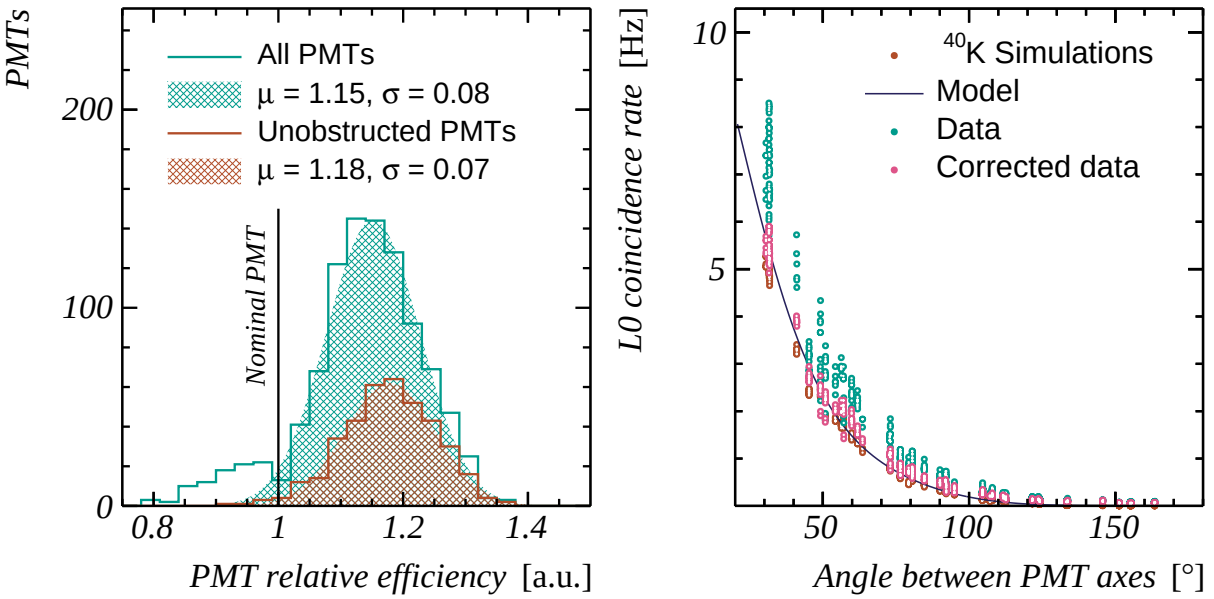


Figure 6.3: ***Left:*** *The fitted PMT relative efficiencies of all PMTs in KM3NeT/ARCA DU1+2. A nominal PMT as used in detector simulations has a relative PMT efficiency of one. The distribution called unobstructed PMTs includes all PMTs with an unobstructed view.* ***Right:*** *Expected and observed total ^{40}K coincidence rate as function of the angle between the PMTs in an arbitrarily chosen DOM. The background from random coincidences has been subtracted from each data point. The parameterisation used to fit the PMT efficiencies is shown as well (see text).*

from random coincidences is subtracted using the routine explained in section 6.1. Differences in the PMT photon detection efficiencies cause a scatter of the points. Nevertheless, it is clear that the observed coincidence rate is significantly higher than the simulated coincidence rate. The points labelled 'corrected data' are obtained by dividing the observed coincidence rate of each PMT pair by the fitted relative efficiencies of the corresponding PMTs. The differences between the calibrated total coincidence rate of PMT pairs sharing the same opening angle is significantly smaller, indicating that differences in the observed L0 coincidence rate can be mainly attributed to the efficiencies of the PMTs. The corrected data points lie on the parameterised model and share (approximately) the same scatter as the simulations.

PMT efficiency comparison with L0 singles rates Most ^{40}K decays (and other light-sources) are detected only by a single L0. This *singles rate* should be (approximately) proportional with the efficiency of the PMT.

In appendix B, the relation between the PMT relative efficiency and the expected rate of L0s from ^{40}K decays in the surrounding seawater is derived. The L0 rate is proportional to the PMT efficiency, with a rate of 4.7 kHz on a nominal PMT (relative efficiency of one). In figure 6.4, this relation is compared to the measured PMT L0 rates and the fitted relative PMT efficiencies. The following contributions to the total L0 rate are included as well:

- Decays of ^{40}K in the DOM glass sphere: 7.5 % of the contribution from ^{40}K in the

seawater [171].

- Dark counts: (Thermo-)electric noise in the PMT, most significantly due to thermal emission of electrons from the photocathode. The contribution (approximately 250 Hz [165]) does not dependent on the PMT efficiency.

- Late hits and delayed pulses (see section 3.5.1): 7.1 % and 3.2 % of all hits respectively [165].

The measured singles rate is indeed approximately proportional to the PMT efficiency. The measure L0 rates agrees well with the expected total L0 rate.

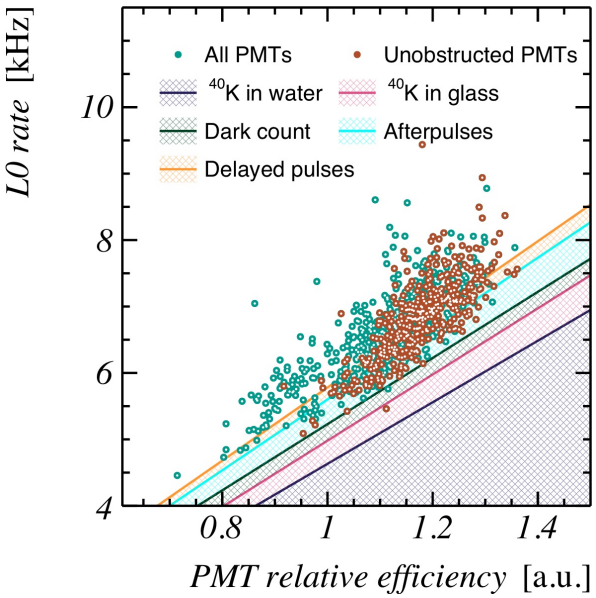


Figure 6.4: *The mean PMT L0 rate versus the fitted relative efficiencies using the ^{40}K calibration technique. The contribution of light from ^{40}K decays in the seawater as well as several additional sources of light are given by the shaded areas.*

6.3 Long-term Monitoring of PMT Parameters

The constant rate of ^{40}K decays provides for a convenient way to monitor the PMT time offsets and efficiencies over time. This is of particular interest since the PMT parameters may vary over time due to ageing of the PMTs or sedimentation on the DOMs. Additionally, the in-situ calibration can account for changes in the detector setup, such as the PMT high-voltage settings.

The PMT efficiencies and time offsets of an arbitrarily chosen DOM are shown in figure 6.5. The considered data taking period is considerably larger than the period used to obtain the other results in this chapter (indicated by the grey band in the figure). The vertical black lines indicate the moments when the high voltage settings of the PMTs were tuned (see section 3.5). A human error was made in the first high voltage tuning, causing a significant decrease in the efficiencies of the PMTs. This error was rectified in the second tuning. Since the transit time of a PMT is related to the applied high voltage, both changes are apparent in the PMT time offsets.

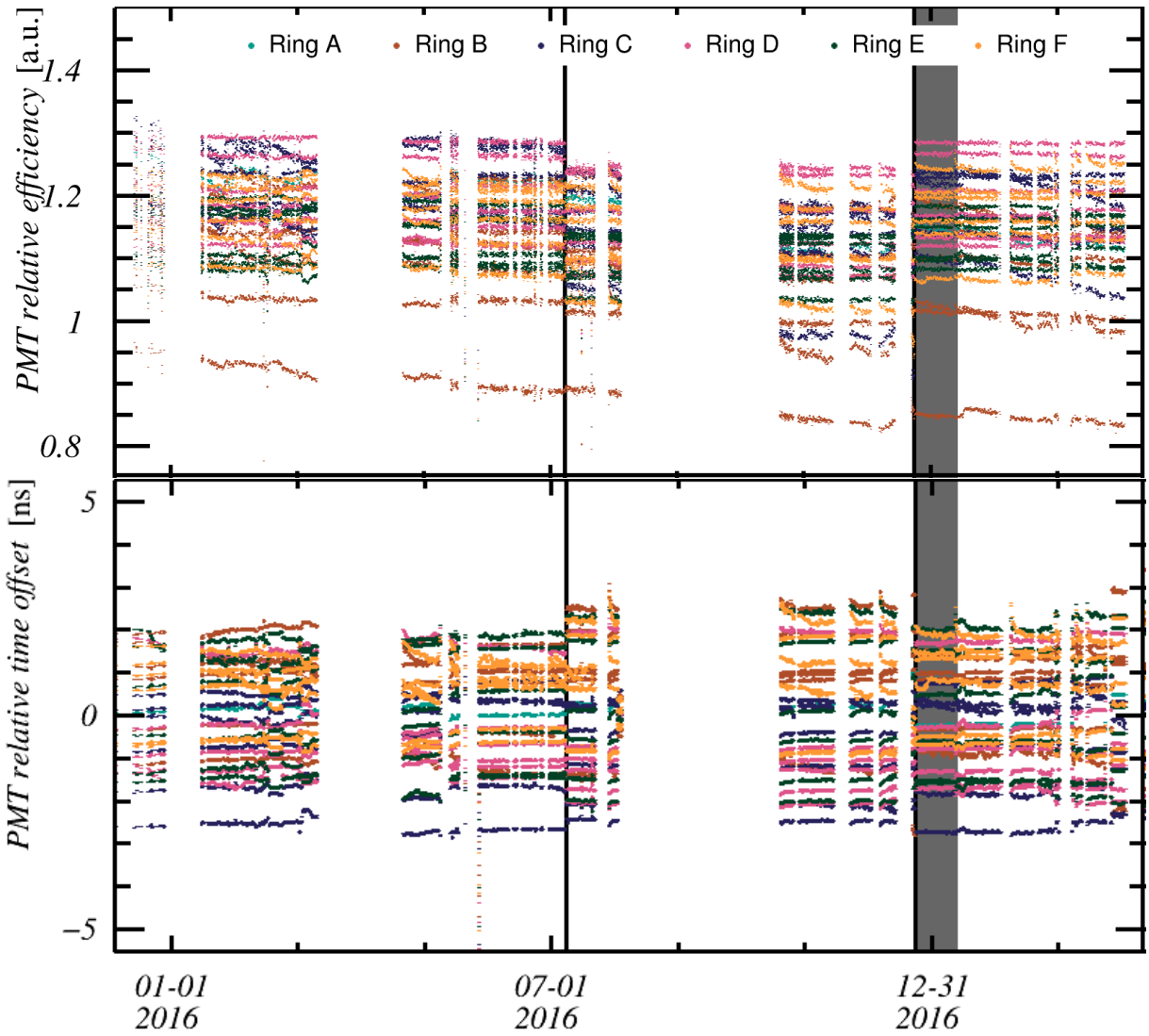


Figure 6.5: *The stability over time of the fitted PMT efficiencies (**top**) and PMT time offsets (**bottom**) of an arbitrarily chosen DOM. The black lines correspond to tuning of the PMT high voltage settings (see text). The grey shaded area corresponds to the data taking period considered for the results presented in this chapter. Periods without data taking are left blank. The colour of each marker corresponds to the ring the PMTs is located in.*

In the figure, periods without data taking are left blank. During these periods, the detector was turned off to investigate and solve various complications. After such a period (when the detector is powered on again) it takes up to a few days before all PMTs are in a stable data taking condition. Even though this effect is small, it can still be observed in the time stability plots.

A few PMTs suffer from more significant variations, over longer periods of time. An example is given by PMT B2, which has the lowest overall efficiency. Over the considered time period of about 16 months, its efficiency degrades by about 13 %. In order to

investigate this decrease, the PMT relative efficiencies obtained from one of the first runs, taken at the fifth of December 2015, are compared to those obtained from data taken at the 12th of April 2017. The results are shown in figure 6.6, where the PMTs in different rings (see section 3.5) are represented by different colours.

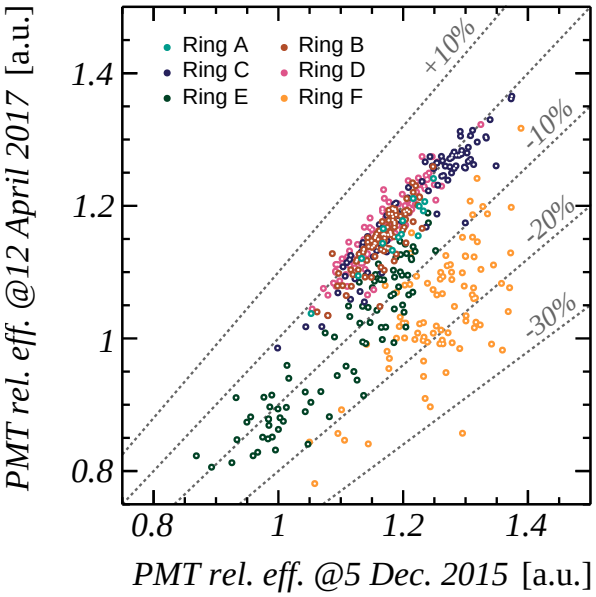


Figure 6.6: *The PMT relative efficiencies as obtained from data taken in-situ using the inter-PMT calibration method based on ^{40}K decays. The fitted efficiencies of an early run (taken on the fifth of December, 2015) are compared to those of a later run (taken on the 12th of April, 2017). The different colours represent the PMTs located in different rings (see 3.5).*

6.4 Conclusions

Despite its simplicity, the model used in the inter-PMT calibration describes the observed L0s coincidence rate distributions reasonably well (see, for example, figure 6.2 (left)). However, the assumed Gaussian distribution is not a perfect fit to the measured L0 time difference distributions. Investigations of the time over threshold values indicate that the deviations in the tails of the distribution (large Δt) are likely to be caused by prepulses (see section 3.5). One possible improvement of the inter-PMT calibration consists of an independent fit of the 31 relative PMT efficiencies, minimising the difference between the expected and the efficiency-corrected time-integrated L0s coincidence rate (figure 6.3 (right)). This option has been investigated, giving comparable PMT efficiencies within a few percent.

The fitted relative PMT time offsets agree well with the time offsets as estimated from nanobeacon data (figure 6.2 (right)). Both methods agree within a standard deviation of 0.79 ns, well within the goal of sub-nanosecond accuracy of the time calibration. The spread between the results is believed to be dominated by the inaccuracy of the nanobeacon calibration method, implying an even better accuracy of the inter-PMT calibration using ^{40}K decays.

After correcting for the differences in the (fitted) relative PMT efficiencies, the observed L0 coincidence rates agree well with the expected rates from simulations, both in the dependence on the opening angle between the PMTs as well as the scatter in the points (e.g. figure 6.3 (right)). The fact that the relative PMT efficiencies are, on average,

significantly higher than one could mean that light sources other than ^{40}K decays in the seawater contribute to the L0s coincidence rate, but are not taken into account in the model. Decays of ^{40}K in the glass of the DOM could be such a source. It is estimated that these decays contribute about 7.5% to the singles rates, but their impact on the coincidence rate could be considerably higher, since the light is emitted close to the PMTs. Another explanation could be that the PMTs have better photon detection efficiencies than expected.If this would be the case, the PMTs are also more sensitive to the light emitted in neutrino interactions, implying a better effective area and resolution of the detector than expected (chapter 5). More detailed simulations are currently in development to investigate the expected L0s coincidence rate.

From figure 6.6, it is clear that some PMTs loose efficiency over time. The most significant loss is observed for the PMTs in rings E and F, i.e. the PMTs located at the top of the DOM. Therefore, it is believed that the loss in efficiency is caused by sedimentation. A much less significant loss in efficiency (of the order of a few per cent over the considered time period of 16 months) is observed for all other PMTs.

Chapter 7

Depth Dependence of the Atmospheric Muon Rate

M. Ageron et al. (KM3NeT Collaboration),*
"Dependence of atmospheric muon flux on seawater depth measured with the first KM3NeT detection units"
Eur. Phys. J., vol. C80, no. 2, p. 99, 2020.

Since muons lose energy while propagating through seawater, fewer atmospheric muons are expected at larger depths (see 3.6.3). In this chapter, a measurement of the depth dependence of the atmospheric muon flux will be discussed, using data from KM3NeT/ARCA (two DUs) and KM3NeT/ORCA (one DU). Each DOM is used as a standalone detector. Using DOMs from both KM3NeT/ARCA and KM3NeT/ORCA, the depth dependence relation is measured over a range of about 1200 meters. The presented results have been published in a paper by the KM3NeT Collaboration[172].

7.1 Data Sample

The results presented in this section are obtained from L1 timeslice data. An L1 is defined as a coincidence of at least 2 L0s within a time window of 25 ns. For KM3NeT/ARCA, this data was taken with two DUs, between the 23rd of December 2016 and the 13th of January 2017 (about 471 hours). For KM3NeT/ORCA, about 33 hours of data are used, taken between the 30th of November and the 2^{nd} of December 2017 with a single DU.

7.2 Background Rejection

In order to obtain an almost background-free measurement of the atmospheric muon rate by each DOM, selection criteria are implemented to reject the bioluminescence and ^{40}K backgrounds.

*Corresponding authors: S. Biagi, M. Lincetto, K. Melis

Bioluminescence Bioluminescence is characterised by an increased hit rate on one or more PMTs on the same DOM for a period of several seconds (see 3.6.2). Bioluminescence is expected to be more prominent in KM3NeT/ORCA than in KM3NeT/ARCA, due to its location and shallower depth. In order to reject light from bioluminescence, a rather stringent cut is introduced. All timeslices in which any of the PMTs within a DOM records an L0 rate above 10 kHz are discarded from the analysis. In KM3NeT/ORCA, only about 5 % of the timeslices pass this criterion (95 % of the data is discarded). For KM3NeT/ARCA, a negligible fraction of the timeslices are discarded by this criterion.

^{40}K Decays In order to distinguish atmospheric muons from the background of ^{40}K decays, a cut on the *multiplicity* of the L1 is used. The multiplicity is defined as the number of PMTs contributing to the L1. To illustrate this cut, the measured and expected multiplicity distributions of two DOMs (floor 1 of KM3NeT/ARCA DU1, at a depth of 2285 m and floor 18 of KM3NeT/ARCA DU2, at a depth of 3385 m) are shown in figure 7.1. The simulated multiplicity distributions of both atmospheric muons and ^{40}K decays agree very well with the distribution obtained from the data.

In this figure, the expected rate from single ^{40}K decays and atmospheric muons is obtained from simulations (see chapter 4). In the simulation of atmospheric muons, all PMT efficiencies are set to the fitted efficiencies using the inter-PMT calibration technique presented in section 6. In the simulations of ^{40}K decays, all PMT efficiencies of the DOM are set to the average fitted efficiency of 1.15 . The contribution from random coincidences is calculated using the measured average singles rate of a PMT ($R_S = 6.7$ kHz, see section 6.2). The random coincidence rate for multiplicity n is given by:

$$\begin{aligned}
R_{RC}(n) &= \frac{10^9\ [\text{ns}]}{[s]} \cdot P_{\text{one PMT hit}} \cdot P_{\text{no PMT hit before}} \cdot P_{\text{n-1 in coincidence}}, \\
P_{\text{one PMT hit}} &\equiv 1 - \exp\big(-31 \cdot R_S[\text{Hz}] \cdot 10^{-9}[\text{s}]\big), \\
P_{\text{no PMT hit before}} &\equiv \exp\big(-31 \cdot R_S[\text{Hz}] \cdot 25 \cdot 10^{-9}[\text{s}]\big), \\
P_{\text{n-1 in coincidence}} &\equiv \binom{30}{n-1} \cdot (1 - P_{RC})^{n-1} \cdot (P_{RC})^{31-n}, \\
P_{RC} &\equiv 1 - \exp\big(-R_S[\text{Hz}] \cdot 25 \cdot 10^{-9}[\text{s}]\big),
\end{aligned} \tag{7.1}$$

with $P_{\text{one PMT hit}}$ the Poisson probability that any of the 31 PMTs recorded an L0 in a time window of 1 ns, $P_{\text{no PMT hit before}}$ the Poisson probability that no PMT recorded an L0 in the 25 ns before this L0 and $P_{\text{n-1 in coincidence}}$ the Binomial probability that $n-1$ out of the remaining 30 other PMTs recorded an L0 within 25 ns after the L0.

Two regimes can be identified in the multiplicity distribution (figure 7.1); for multiplicities lower than 8 PMTs, the contribution from ^{40}K decays dominates, while for higher multiplicities only atmospheric muons contribute. This separation is caused by the fact that the atmospheric muon rate is lower than the rate of ^{40}K decays, but per muon more light is emitted. The fact that the atmospheric muon rate decreases with depth causes the total rate of L1s with multiplicity ≥ 8 to be lower on the lower DOM. In the following section, this feature will be used to measure the depth dependence of the atmospheric muon rate.

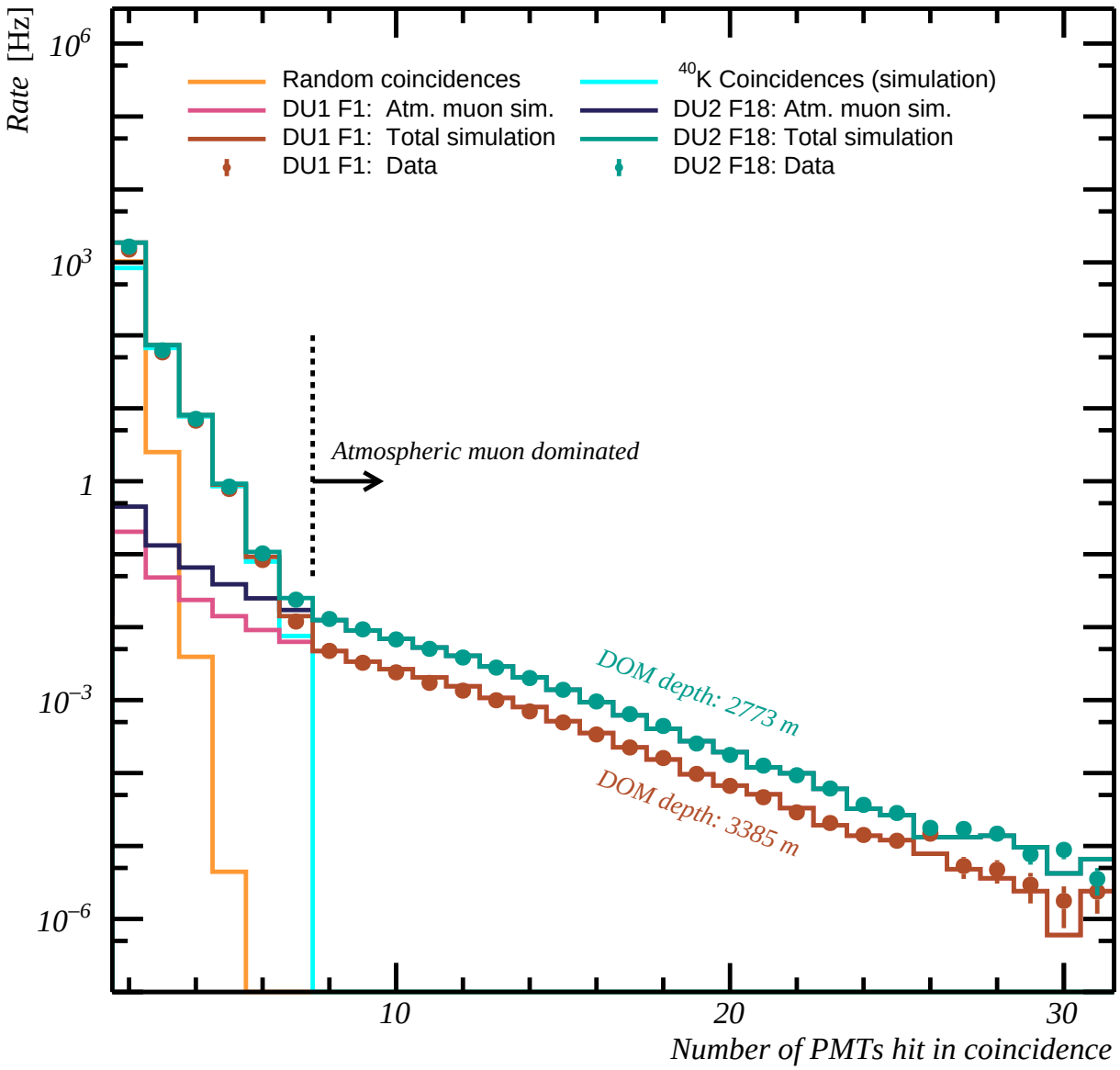


Figure 7.1: *The distribution of the number of PMTs in a DOM hit in coincidence (the multiplicity of an L1) as observed in data, obtained from atmospheric muon and ^{40}K simulations and as expected from random coincidences. The distributions of the highest (labelled L2F18, at a depth of 2773 m) and lowest positioned DOM (L1F1, at a depth of 3385 m) along KM3NeT/ARCA DU1 and DU2 are shown.*

7.3 Depth Dependence of the Atmospheric Muon Rate

The rate of L1s with multiplicity ≥ 8 gives a measure of the flux of atmospheric muons passing by a DOM. Due to the absorption of muons in water (see section 3.6.3), the ≥ 8-fold coincidence rate decreases approximately exponentially with depth.

Expected atmospheric muon rate A model of the direction-integrated atmospheric muon flux R^μ as function of the depth d below the surface is given by Bugaev et al.[32, 209], with an (estimated) 8% uncertainty on the flux normalisation $R^\mu(0)$ as the dominating factor in the systematic error[33]:

$$R^\mu(d) = R^\mu(0) \times \frac{B_1 \cdot \exp(A_1 \cdot d) + B_2 \cdot \exp(A_2 \cdot d)}{C_1 + C_2 \cdot d}, \tag{7.2}$$

$$R^\mu(0) = 1^{0.08}_{-0.08}, \tag{7.3}$$

where the depth d is measured in units of kilometre water-equivalent (which takes the density of the water at these depths into account), and

$$\begin{aligned} A_1 &= -2.91 \cdot 10^{-7}, \quad & A_2 &= -1.17 \cdot 10^{-3} && [1 \text{ / m}], \\ B_1 &= 1.31 \cdot 10^{-5}, \quad & B_2 &= 7.31 \cdot 10^{-7} && [1 \text{ / cm}^2 \text{ / s / sr}], \\ C_1 &= 4.16 \cdot 10^{-1}, \quad & C_2 &= 1.07 \cdot 10^{-4} && [1 \text{ / m / sr}]. \end{aligned}$$

PMT efficiencies Differences in the efficiencies of the PMTs in each DOM can have a significant impact on the measured ≥ 8-fold coincidence rate (especially if one of the PMTs has a very low efficiency). In order to correct for this effect, an efficiency corrected ≥ 8-fold coincidence rate $R_i^{8+,\epsilon\text{-corr.}}$ is determined for each DOM:

$$R_i^{8+,\epsilon\text{-corr.}} \equiv \frac{R_i^{8+}}{\epsilon_i^{8+}}, \tag{7.4}$$

where R_i^{8+} is the measured ≥ 8-fold coincidence rate of DOM i and ϵ_i^{8+} is the relative efficiency of a DOM to detect a ≥ 8-fold coincidence. The value of ϵ_i^{8+} for each DOM is estimated using two sets of simulations:

- A simulation using the nominal PMT model for all PMTs. These simulations will be referred to as the 'nominal simulation'.
- A more realistic simulation, using the fitted PMT efficiencies from the inter-PMT calibration procedure outlined in chapter 6.

For each DOM, the ratio between the ≥ 8-fold coincidence rate from the realistic $R_i^{8+,\text{real. sim}}$ and the nominal simulations $R_i^{8+,\text{nom. sim}}$ is used as an estimate of ϵ_i^{8+}:

$$\epsilon_i^{8+} = \frac{R_i^{8+,\text{real. sim}}}{R_i^{8+,\text{nom. sim}}}. \tag{7.5}$$

The measured and simulated ≥ 8-fold coincidence rates are shown as a function of the depth of each DOM in figure 7.2 (left).

DOM effective area In order to calculate the direction-integrated atmospheric muon rate R_i^μ from the efficiency corrected ≥ 8-fold coincidence rate ($R_i^{8+,\epsilon\text{-corr.}}$), an effective area $A_{\text{eff.}}^{8+}$ is introduced:

$$R_i^\mu = \frac{R_i^{8+,\epsilon\text{-corr.}}}{A_{\text{eff.}}^{8+}}. \tag{7.6}$$

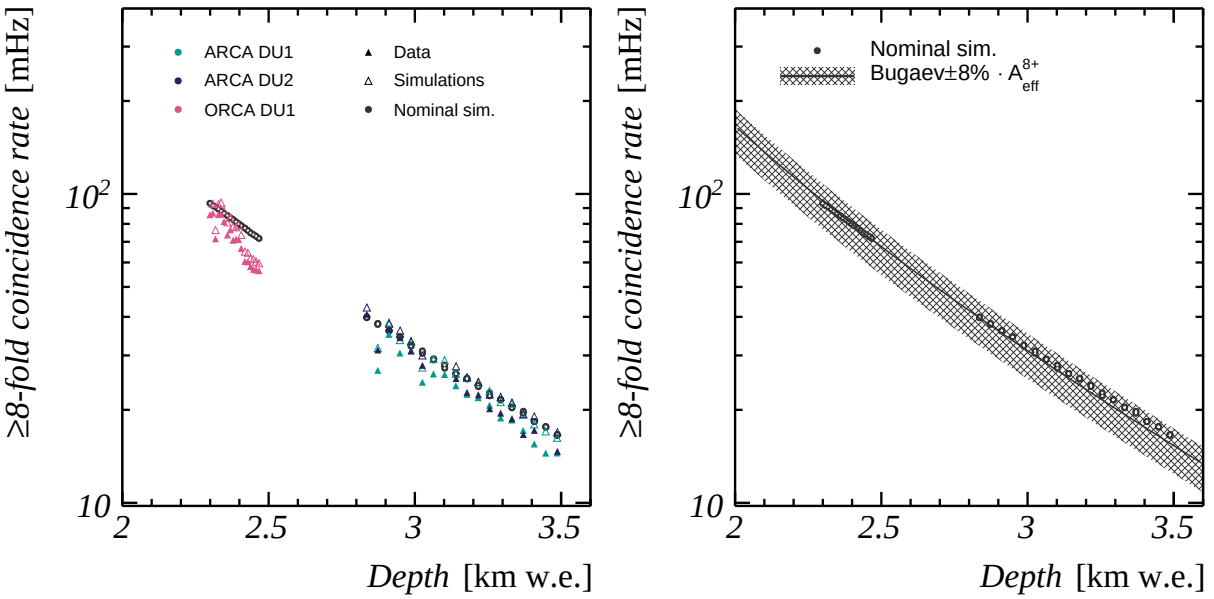


Figure 7.2: ***Left:*** *The rate of ≥ 8 multiplicity coincidences as a function of DOM depth below the sea surface (given in units of kilometre water equivalent). The solid triangles correspond to data, the open triangles to simulations. The grey dots correspond to the nominal simulation (see text).* ***Right:*** *The result of the fit used in the determination of the DOM effective area. The band gives the eight per cent uncertainty on the flux normalisation of the model, as well as the systematic error on the fitted effective area (see text). The dots give the ≥ 8-fold coincidence rate obtained from the nominal simulation.*

All DOMs are assumed to have the same effective area. The value of $A^{8+}_{\text{eff.}}$ is estimated from a fit of the expected direction-integrated muon flux ($R^{\mu}(d)$, see equation 7.3) to the nominal ≥ 8-fold coincidence rates $R_i^{8+,\text{nom. sim}}$. In figure 7.2 (right), this fit is shown. The fitted value is:

$$A^{8+}_{\text{eff.}} = 96^{+5}_{-12}\ [m^2], \tag{7.7}$$

where the error gives a measure of the systematic error caused by the (unknown) distribution of muons in a atmospheric muon bundles. This error is estimated using a set of dedicated simulations in which these distributions are varied. The most significant contribution to the total error is given by the number of muons in each bundle.

Corrected data The effective area $A^{8+}_{\text{eff.}}$ and the values of ϵ_i^{8+} are used to calculate an estimate of the direction-integrated muon flux at the depth d_i of each DOM:

$$R_i^{\mu} = \frac{R_i^{8+}}{A^{8+}_{\text{eff.}} \cdot \epsilon_i^{8+}}. \tag{7.8}$$

In figure 7.3 these measurements are compared to the model from Bugaev et al. (equation 7.3). The errors on the data points correspond to the systematic error in the effective area $A^{8+}_{eff.}$, the error band around the model corresponds to the eight per cent uncertainty on the flux normalisation.

7.4 Conclusions

Since the effective area of the DOMs is fitted to the model, the presented results represent a measurement of the dependence of the atmospheric muon flux with depth, rather than a direct measurement of the absolute atmospheric muon flux. At the probed depths by the KM3NeT/ARCA and KM3NeT/ORCA DOMs, ranging over more than one kilometre, the measurements and the model agree within the estimated errors on both. These results underline the fact that both detectors can be calibrated using the inter-PMT calibration procedures described in chapter 6, even though the detectors are deployed at locations more than 1000 kilometres apart.

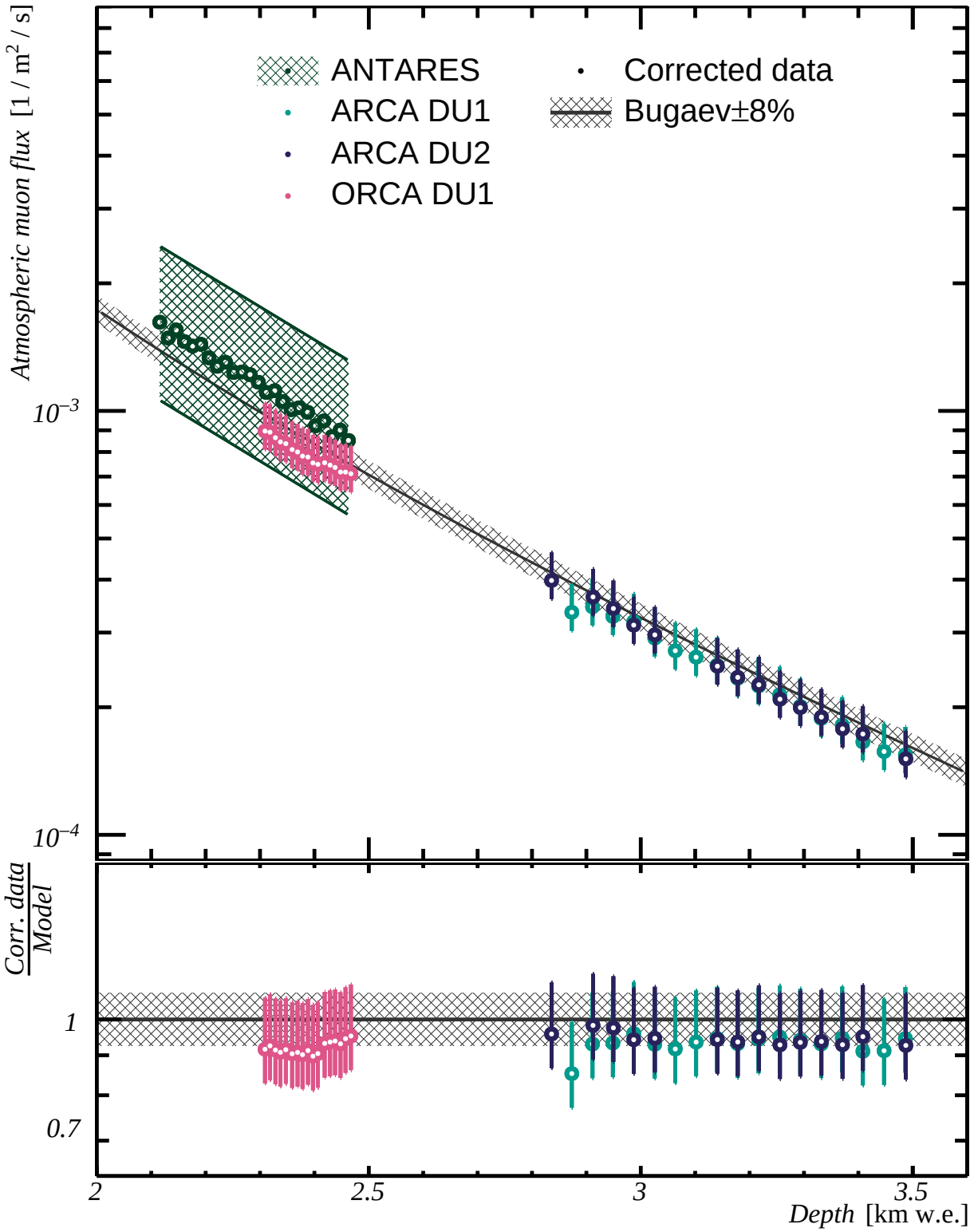


Figure 7.3: *The direction-integrated atmospheric muon rate as a function of depth (in water equivalent kilometres), measured by the DOMs of KM3NeT/ARCA (two DUs) and KM3NeT/ORCA (one DU). The error bands correspond to the systematic error in the effective area of the DOMs (see text). The model given by Bugaev et al.[32], with the estimated error on the total flux[33], is indicated by the grey line. A measurement (with corresponding error band) of the atmospheric muon flux carried out by Aguilar et al. (the ANTARES collaboration)[33] is shown for comparison.*

Chapter 8

Inter-DOM Calibration

In chapter 6, the calibration of the PMT time offsets is discussed. These account for PMT specific time delays (such as differences in the PMT transit times) relative to each other. For an absolute time calibration, an additional time offset is needed, called the DOM time offset (see section 3.8). In this chapter, two routines to calibrate the DOM time offsets will be discussed.

Both methods are based on comparing the arrival times of the Cherenkov light emitted by atmospheric muons passing by a detection unit with a model based on simulations. This model depends on the positions of the DOMs. However, the acoustic positioning system was, at the moment of writing, not completely functional. Therefore, the DOMs are assumed to have no horizontal displacement, thus following a vertical straight DU. This is a reasonable assumption, since the maximum tilt of a detection unit is of the order of $0.6\,^{\circ}$. The heights of the DOMs along the line are estimated from pressure measurements. Please see section 3.5 for details on the acoustic positioning system, maximum tilt and estimated DOM positions.

8.1 Coincident Light between DOMs

The light emitted by atmospheric muons can be detected by two or more DOMs in coincidence. Similar to the inter-PMT calibration procedure using ^{40}K decays (where coincidences of L0s on two or more PMTs within the same DOM are used, see 6) coincidences between L1s observed on two or more DOMs are used to calibrate the DOM time offsets. This is done by comparing the observed distributions of the time differences between L1s on each DOM pair with a model obtained from simulations.

Model The expected distribution of the time difference between two coincident L1s observed on each pair of DOMs is obtained from simulations (see chapter 4). In figure 8.1, these distributions are shown for all DOMs with respect to DOM 9 in KM3NeT/ARCA DU1. The labels (F1, F2, ...) correspond to the floor of the DOM (see section 3.5). In this simulation, hits from ^{40}K decays are only added in a time window of a few microseconds surrounding each simulated atmospheric muon. Hence, the contribution of random coincidences between the uncorrelated ^{40}K decays and simulated atmospheric

muons is negligible.

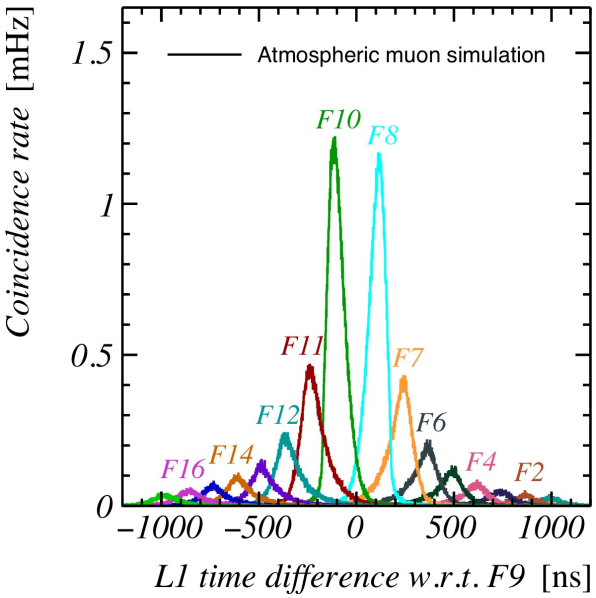


Figure 8.1: *The time difference distribution of L1s between DOMs of KM3NeT/ARCA DU1 with respect to the DOM positioned in DU1 on floor 9. The labels (F1, F2, ...) correspond to the floor of the DOM (see 3.5), where the F1 is the lowest DOM. The uppermost DOM (floor eighteen) was not functional during the considered period, and therefore not included in the simulations.*

As a result of the zenith angle distribution of the atmospheric muons and absorption of the muons in the water, fewer coincidences of L1s are expected with increasing distance between the DOMs. Each L1 time difference distribution peaks at approximately $\Delta d/C$, where Δd denotes the distance between the DOMs and $C = 0.3$ m/ns is the speed of light. This is expected, as most atmospheric muons passing by multiple DOMs will be travelling approximately vertically downwards. The spread in each distribution is caused mainly by the zenith angle distribution of the simulated atmospheric muons. This, in combination with the characteristic light emission signature of Cherenkov light, causes the asymmetric shape of each distribution.

Random coincidence background By design, each timeslice in the atmospheric muon simulation contains at least one atmospheric muon. Therefore, the contribution of random coincidences of L1s caused by ^{40}K decays is almost completely missing in the model (figure 8.1). To account for this contribution, the random coincidence rate is estimated from data, by taking the average rate of the tails (± 1.5 μs to ± 2.5 μs away from the maximum) of the L1 time difference distribution. In figure 8.2 (left), the L1 time difference distribution is shown for a pair of two adjacent DOMs. To illustrate the fit procedure, a time offset of 500 ns with respect to the dark room calibration (see section 3.8) is applied to the L1 data of one DOM. The estimated random L1 coincidence rate from uncorrelated ^{40}K decays and atmospheric muons is given by the shaded area.

Fit procedure The calibration procedure of the DOM time offsets $\vec{t0} = \{t0_0, t0_1, \cdots\}$ is based on maximising the agreement between the observed coincidence rate distributions and the expected distributions obtained from simulations for each DOM pair i, j. The 'agreement' between these distributions is quantified by a quality parameter $Q_{i,j}$, related to the likelihood to observe the detected coincidence rate given the expected rate from

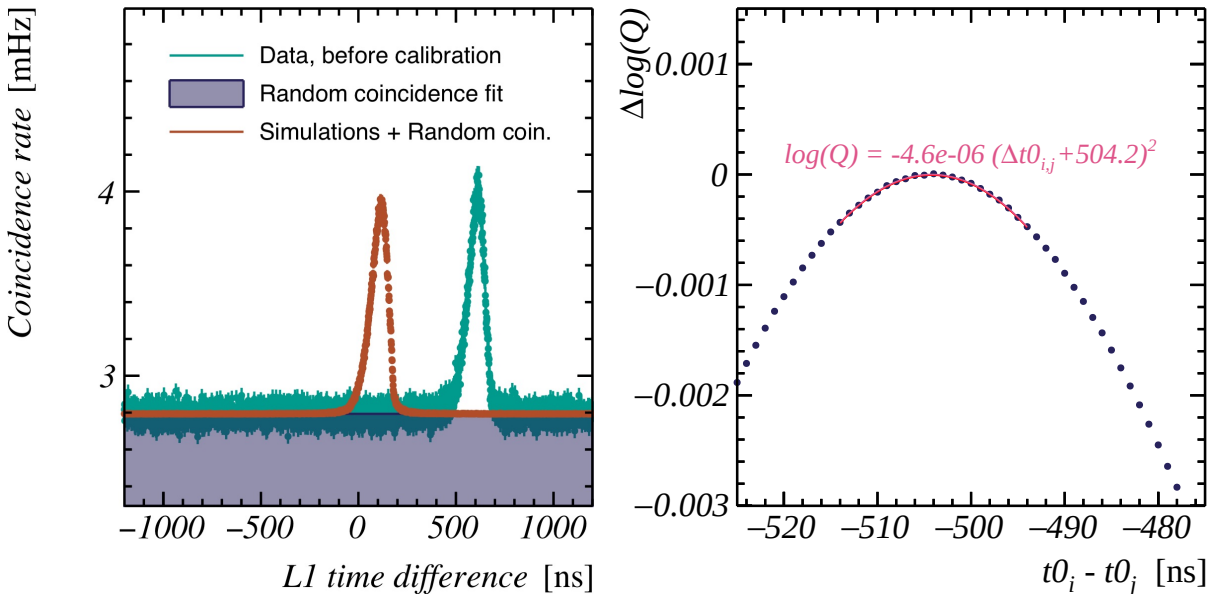


Figure 8.2: ***Left:*** *The hit time difference distribution of L1s on a typical pair of adjacent DOMs, as obtained from data (arbitrarily shifted to illustrate the calibration procedure) and simulated atmospheric muon events. From the tails of the data distribution, the random L1 coincidence rate is estimated (shaded area) and added to the distribution obtained from simulations.* ***Right:*** *The agreement of the L1 time difference distributions between data and simulations, quantified using the quality parameter Q (see text). $\Delta t0$ denotes the difference in DOM time offsets. The agreement between data and simulations is optimal at a shift (in the left figure) of* 504 *ns.*

simulations in each bin of the histograms:

$$Q_{i,j}\left(t0_i - t0_j\right) \equiv \prod_{\text{bins } k}\left[\frac{e^{-m_{i,j,k}} \cdot \left(m_{i,j,k}\right)^{d_{i,j,k+l}}}{d_{i,j,k+l}!}\right], \tag{8.1}$$

$$l \equiv \frac{t0_i - t0_j}{\text{bin width}}, \tag{8.2}$$

where $m_{i,j,k}$ is the expected coincidence rate for DOM pair i,j in bin k (obtained from simulations) plus the estimated random coincidence rate for this DOM pair, and $d_{i,j,k+l}$ is the observed coincidence rate for this DOM pair in bin $k+l$. The integer l depends on the difference between the DOM time offsets of the two DOMs. Typically, the bin width is chosen to be 1 ns.

In figure 8.2 (right), the value of Q is plotted as function of the difference in the DOM time offsets $t0_i - t0_j$. The horizontal spacing between the points corresponds to the chosen bin width of the histograms (in this case 1 ns). The point with the best agreement between the data and simulations lies at a DOM time offset difference of approximately −504 ns. Since the accuracy of this number is limited to the bin width of 1 ns in the coincidence rate distributions, a more precise estimate of the best agreement is obtained by a fit of a quadratic function to the values of $\log(Q_{i,j})$ of the points within ±10 ns of this maximum:

$$q_{i,j}(t0_i - t0_j) \equiv A_{i,j} \cdot ((t0_i - t0_j) - B_{i,j})^2 + C_{i,j}. \tag{8.3}$$

In the example distribution (figure 8.2 (right)), the maximum of the fit lies at $B_{i,j}$ =−504.2 ns. This value is different from the artificially applied −500 ns shift. The difference of 4.2 ns can be explained by a systematic inaccuracy of the vertical DOM positions, as will be discussed in section 8.3.

The aim of the calibration is to find the set of DOM time offsets $\vec{t0}$ that simultaneously maximises the agreement between the hit time difference distributions from data and simulations of (a selection of) all DOM pairs. The corresponding global quality parameter is defined as:

$$Q(\vec{t0}) = \prod_{j \neq i} \left[Q_{i,j}(t0_i - t0_j) \right], \tag{8.4}$$

or, equivalently, the logarithm of this parameter:

$$q(\vec{t0}) \equiv \log\Big(Q(\vec{t0})\Big) = \sum_{j \neq i} \left[q_{i,j}(t0_i - t0_j) \right], \tag{8.5}$$

where $Q_{i,j}$ is the quality function of a single DOM pair (equation 8.2) and $q_{i,j}$ is the fitted quadratic function (equation 8.3). From the set of fit parameters $A_{i,j}$, $B_{i,j}$ and $C_{i,j}$ describing $q_{i,j}$ of each DOM pair, the set of DOM time offsets maximising $Q(\vec{t0})$ can be derived:

$$\begin{aligned} 0 &= \frac{\partial}{\partial t0_k} log\left(Q(\vec{t0})\right) \qquad\qquad \forall \quad k, \\ 0 &= \sum_{j \neq k} \left[\frac{\partial}{\partial t0_k} q_{k,j}(t0_k - t0_j) \right], \\ &= \sum_{j \neq k} \left[\frac{\partial}{\partial t0_k} \left(A_{k,j} \cdot (t0_k - t0_j - B_{k,j})^2 + C_{k,j} \right) \right], \\ &= \sum_{j \neq k} \left[A_{k,j} \cdot (t0_k - t0_j - B_{k,j}) \right], \end{aligned} \tag{8.6}$$

where in the last step the identities $A_{i,j} = A_{j,i}$ and $B_{i,j} = -B_{j,i}$ are used, following from the fact that $Q_{i,j}(t0_i - t0_j) = Q_{j,i}(t0_j - t0_i)$. The set of DOM time offsets can be determined using basic linear algebra:

$$\tilde{\vec{t0}} = \begin{bmatrix} t0_0 \\ t0_1 \\ t0_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \sum_{i \neq 0} [A_{0,i}] & -A_{0,1} & -A_{0,2} & \cdots \\ -A_{1,0} & \sum_{i \neq 1} [A_{1,i}] & -A_{1,2} & \cdots \\ -A_{2,0} & -A_{2,1} & \sum_{i \neq 2} [A_{2,i}] & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum_{i \neq 0} [A_{0,i} \cdot B_{0,i}] \\ \sum_{i \neq 1} [A_{1,i} \cdot B_{1,i}] \\ \sum_{i \neq 2} [A_{2,i} \cdot B_{2,i}] \\ \vdots \end{bmatrix}. \tag{8.7}$$

In a similar way as the inter-PMT time calibration using coincident light from ^{40}K decays, the procedure outlined for the inter-DOM time calibration is not sensitive to the absolute values of DOM time offsets; only to the DOM time offsets relative to each other.

As a result of the matrix inversion method, the average of the fitted DOM time offsets will always be zero.

The results of the inter-DOM time calibration will be discussed in section 8.3, by comparing the determined DOM time offsets with the results of the nanobeacon calibration (see section 3.8) and a second time calibration method based on reconstructed atmospheric muon tracks, which will be discussed first.

8.2 Hit Time Residuals

A complementary method to calibrate the DOM time offsets is based on *hit time residuals* of reconstructed events. The time residual of an L0 is defined as the time difference between an observed L0 on a DOM with respect to the expected arrival time of light from a reconstructed muon track.

For each DOM, the distribution of the hit time residuals of a large number of simulated atmospheric muon events serves as a reference for the inter-DOM calibration. Each simulated (triggered) event is reconstructed using the algorithm outlined in section 5.3.4, including the hits on all DOMs on the same DU, except the DOM under consideration. A triggered event with, for example, 9 DOMs hit will thus be used in 9 independent fits, using the hits on 8 out these 9 DOMs each. Excluding this DOM makes the reconstruction more computationally intensive, but is believed to make the inter-DOM time calibration less sensitive to systematic uncertainties. Each reconstructed muon track is characterised by five independent parameters (see section 5.3.4). If the arrival times of five DOMs (i.e. five data points) are used in the fit, the number of degrees of freedom is zero, and the fitted parameters will have a relatively large error compared to reconstructed tracks with six or more DOMs used in the fit. In order to select reconstructed tracks of relatively good quality, only triggered events with at least 7 DOMs on the same DU contributing to the fit are used.

The same procedure and selection criteria are used to obtain the hit time residual distributions of data taken in-situ. For each DOM i, the expected hit time residual distribution from simulations and the observed hit time residual distribution are used to determine the optimal DOM time offset $t0_i$, maximising the agreement between both distributions using a quality parameter Q very similar to the one given in equation 8.2:

$$Q_i(t0_i) \equiv \prod_{\text{bins } k} \left[\frac{e^{-m_{i,k}} \cdot (m_{i,k})^{d_{i,k+l}}}{d_{i,k+l}!} \right], \tag{8.8}$$

$$l \equiv \frac{t0_i}{\text{bin width}}, \tag{8.9}$$

where $m_{i,k}$ and $d_{i,k+l}$ are the expected and observed number of reconstructed events with a hit time residual in bin k and $k+l$, respectively. The used bin width in the hit time residual distributions is 1 ns. The DOM time offset maximising Q_i is found by fitting a quadratic function to $\log(Q)$:

$$q_i(t0_i) \equiv A_i \cdot (t0_i)^2 + C_i, \tag{8.10}$$

A major difference between this method (using hit time residuals) and the method using L1 time differences lies in the fact that the fitted DOM time offsets affect the

parameters of the reconstructed tracks, and therefore the hit time residual distributions of all other DOMs. Therefore, the previous steps are repeated several times, iteratively updating the DOM time offsets using equation 8.10, until the fitted DOM time offsets converge to stable values. In order to assure convergence in a reasonable number of iterations, the DOM time offsets used in the first step need to be relatively close to the true DOM time offsets. Therefore, the DOM time offsets as determined using the L1 time difference distributions (section 8.1) have been used as a seed. With this starting point, about fifteen iterations are needed until the DOM time offsets convergence to stable values.

In figure 8.3, The hit time residual distribution of one DOM is shown in figure 8.3. To illustrate the calibration procedure, the time offset of this DOM is deliberately shifted by 10 ns with respect to the DOM time offset determined using the L1 time difference calibration method. The corresponding distribution is labelled "uncalibrated data". The distribution labelled "calibrated data" gives the distribution after the hit time residual calibration procedure. The observed distribution agrees with the expected distribution obtained from simulations, as one would expect after a successful DOM time offset calibration.

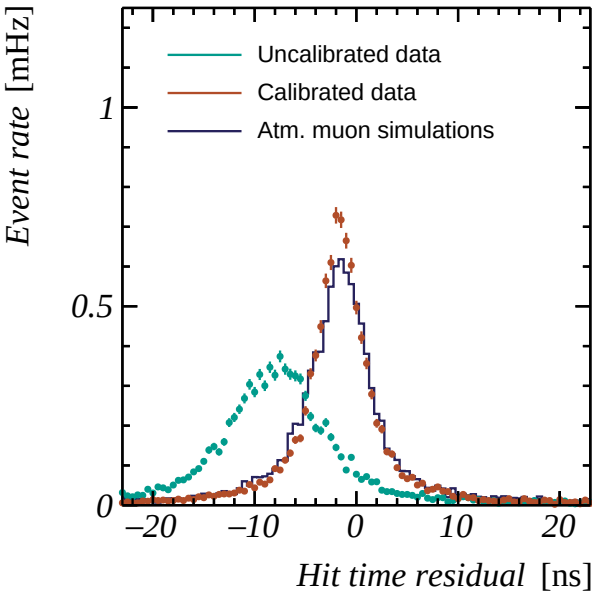


Figure 8.3: *The distribution of hit time residuals for DOM floor 12 on KM3NeT/ARCA DU 1. For illustrational purposes, the time offset of the DOM is deliberately put at a (fictional) miscalibration of −10 ns with respect to the initial calibration using L1 time difference distributions (see 8.1). The hit time residual distribution obtained from simulations is shown as well (see text).*

8.3 Detection Unit Length

In the previous sections, two routines to calibrate the DOM time offsets using atmospheric muons have been introduced. The technique based on the hit time difference of L1s recorded coincidentally on two DOMs will be referred to as the *L1dt calibration*. The calibration obtained from hit time residuals as the *HTR calibration*.

The DOM time offsets determined in-situ with the L1dt and HTR calibration routines are cross-checked with the results from the nanobeacon LED flashers (NB calibration, see section 3.8) in figure 8.4 (left).

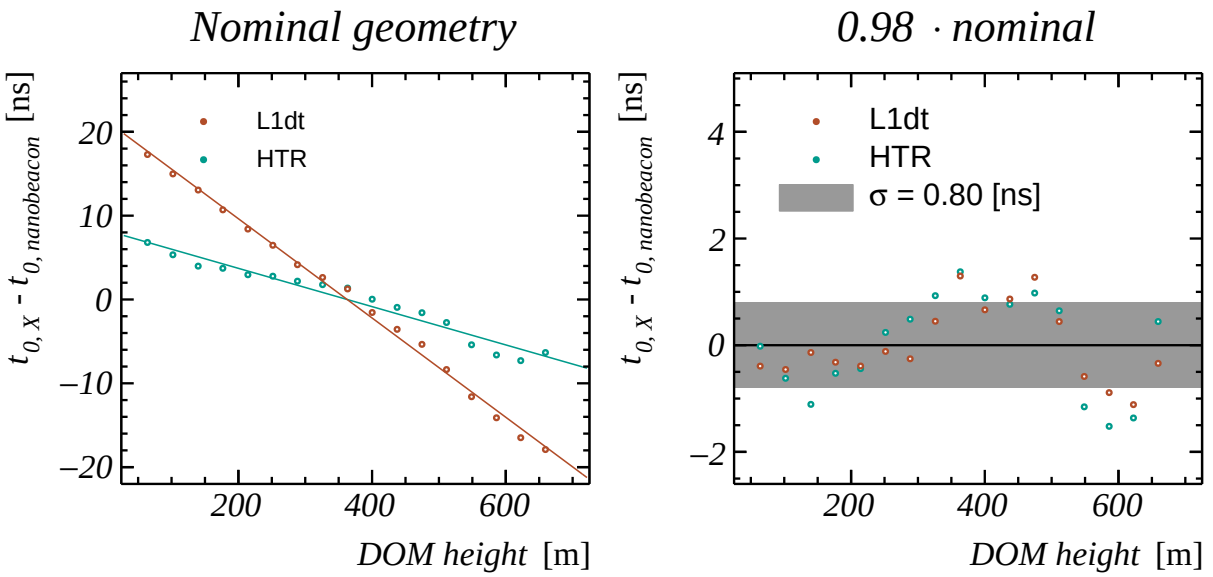


Figure 8.4: ***Left:*** *The DOM time offsets of DU1 as function of the height along the DU. The fitted DOM time offsets of all in-situ methods are compared to those obtained from the nanobeacon calibration. The average of the DOM time offsets is subtracted from all calibration methods.* ***Right:*** *Same as left figure, but calibrated using models based on simulations with a detector in which all distances between the DOMs are scaled (shortened) by 2 %.*

The time calibration based on the time differences between L1s as well as the one based on hit time residuals show a systematic linear trend. Such a trend indicates either an error in the fit procedure, or a systematic deviation in the simulations caused by (for example) incorrect DOM positions. In case the DOM positions are indeed incorrect, the observed linear trend could be explained by an overall scaling of the detection unit, i.e. the height difference between each pair of neighbouring DOM is under/overestimated by a certain length.

Due to the fact that the in-situ acoustic position system is not completely functional, it is not possible to verify this hypothesis with a direct measurement of the DOM heights. However, discussions with the technical crew revealed that the positions of the DOMs along the detection unit are estimated using the pressure gauge of the ROV inspecting the DUs after deployment [210]. To convert the pressure into a position measurement, the density of freshwater, rather than salt water has been used. Using the correct density of salt water results in a detection unit that is approximately 2 % shorter than assumed.

To check the hypothesis of this scaled DU, the L1dt calibration method has been modified to fit the scaling of the detection unit length in addition to the DOM time offsets. A best fit is obtained for a detection unit that is 1.9 % shorter than assumed (e.g. the distance between each pair of neighbouring DOMs is 35.32 m rather than 36 m). The simulations discussed in chapter 4 have been repeated to obtain new models for the L1dt calibration and HTR calibration. In figure 8.4 (right), the DOM time offsets obtained from the L1dt, HTR and NB calibration routines using this updated geometry are shown. With the 1.9 % shorter detection unit, the determined time offsets using atmospheric

muons agree within a standard deviation of 0.8 ns with the nanobeacon calibration.

8.4 Inter-DU Time Calibration

The inter-DU time offsets are known with picoseconds accuracy (see section 3.8). These offsets have been used to test a L1dt calibration procedure in which the relative DOM time offsets are fixed within the same DU, and only the inter-DU time offsets are left free in the fit. This procedure has been applied to the data from the first two KM3NeT/ARCA detection units. Both procedures are compatible, given an estimated accuracy of the inter-DU L1dt calibration of about 10 ns[211].

Chapter 9

Single Line Track Reconstruction

Up to May 2016, KM3NeT/ARCA consisted of a single detection unit. In order to study the flux of atmospheric muons, and identify potential neutrino candidate events, an algorithm to reconstruct (atmospheric) muon tracks using data from a single detection unit has been developed (see section 5.3.4). In this chapter, results of these two studies will be presented. In this, the DOMs are assumed to have nominal positions (as discussed in section 3.5).

The results in this chapter are obtained from triggered events in the approximately 471 hours of data taken between 23rd of December 2016 and the 13th of January 2017 (see sections 6.2 and 7.1), as well as from approximately 465 hours of simulations of atmospheric muon events. These simulations incorporate realistic PMT efficiencies, determined using the inter-PMT calibration procedure outlined in chapter 6.

9.1 Reconstruction Resolution

In figures 9.1 (left) and 9.1 (right), the estimated zenith angle (using the reconstruction algorithm described in section 5.3.4) of simulated atmospheric muon events is compared with the true zenith angle.

From these figures, it is clear that the algorithm is not always reconstructing the correct track parameters; differences of up to 160 ° between the reconstructed and true zenith angle exist. The causes of these differences will be investigated in this section, and corresponding event selection criteria will be introduced to reject these *mis-reconstructed tracks*. In this, a mis-reconstructed track is defined as a track with a reconstructed zenith angle differing more than 20 ° from the true zenith angle. A fraction of 11.3 % of all reconstructed atmospheric muon tracks is mis-reconstructed.

9.1.1 Event Selection

Three selection stages are used to reject mis-reconstructed tracks:

Stage 1: Fit quality The events with the largest difference between the reconstructed and true zenith angle contain multiple outliers. Consequently, the fit quality of these events

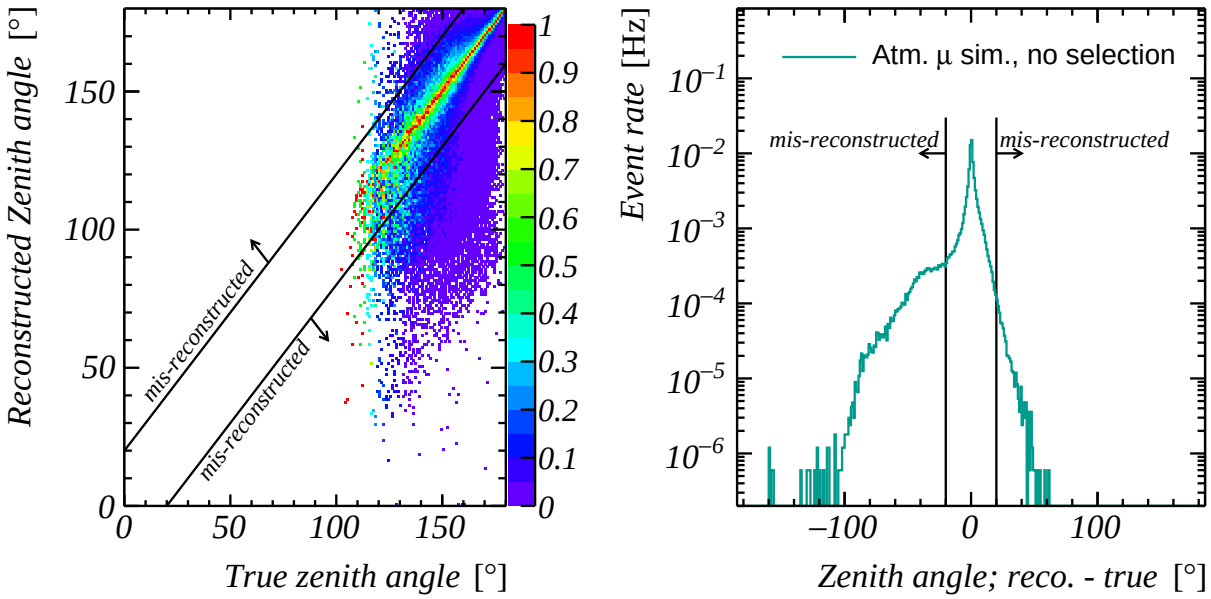


Figure 9.1: ***Left:*** *The distribution of the reconstructed zenith angle versus the true zenith angle of simulated atmospheric muon events. The entries in each vertical slice are scaled such that the maximum is equal for each slice.* ***Right:*** *The distribution of the difference between the reconstructed zenith angle and the true zenith angle, as obtained from simulated atmospheric muon events.*

is very poor. In order to reject these events, a cut on the fit quality Q is introduced (see 5.3.4 for the definition of Q).

In figure 9.2, the distribution of the difference between the true and reconstructed zenith angle is given as function of the fit quality Q (before any selection criteria are applied). Based on this figure, a cut of $Q < -10^3$ is chosen. About 24 % of the mis-reconstructed events do not pass this criterion, while approximately 11 % of the all reconstructed tracks are rejected by this criterion.

Stage 2: Number of large hit time residual hits A second selection criterion focuses on the number of hits with hit time residuals greater than 100 ns. A reconstructed event is rejected in case more than one of these outliers is present in the hits used in the fit. The corresponding distribution is given in figure 9.2

Stage 3: Height of first light A typical example of a mis-reconstructed event that is not excluded by the first two selection stages, is shown in figure 9.3. The dots represent the snapshot and triggered hits in the triggered event, the lines the expected arrival time of unscattered Cherenkov light (using equation 5.19) for the reconstructed and true track parameters. In this event, the uppermost DOM registers a hit with an arrival time about 50 ns later than the expected arrival time of direct Cherenkov light. The muon track that agrees best with the described pattern is moving almost horizontal, resulting in a characteristic 'kink' at a height of about 500 m. This height corresponds to the height along the detection unit where light from the fitted track can be observed first; the height of first light (see section 5.3.4).

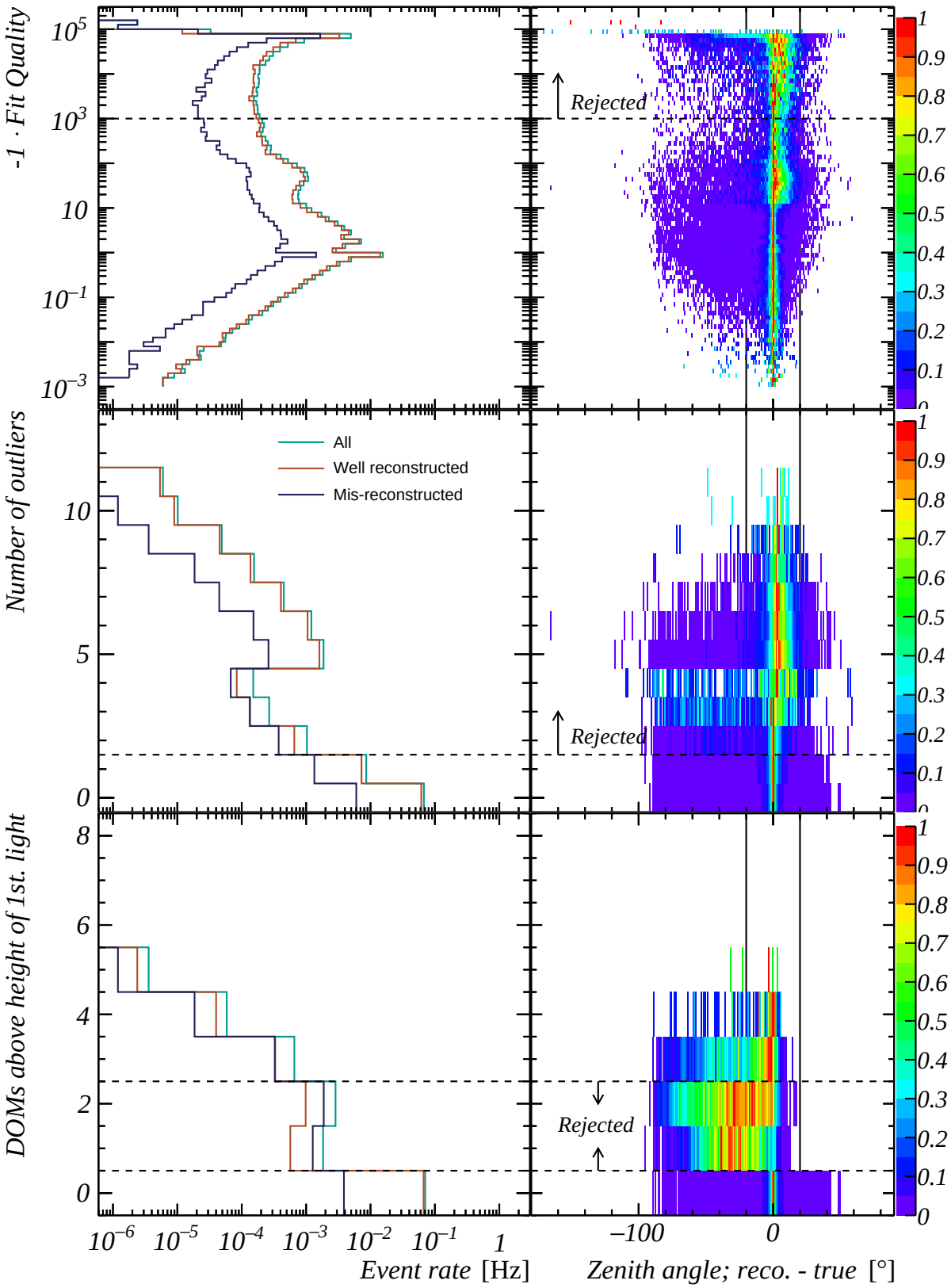


Figure 9.2: *The difference between the reconstructed and the true zenith angle for simulated atmospheric muon events. Given are the distributions before each selection stage discussed in the text; before stage 1: the fit quality (**top panels**), stage 2: the number of outliers (**middle row**) and stage 3: the number of hit DOMs above the height of first light (**bottom figures**). The entries of each vertical slice in the distributions in the right column are scaled such that the maximum of each slice equals one.*

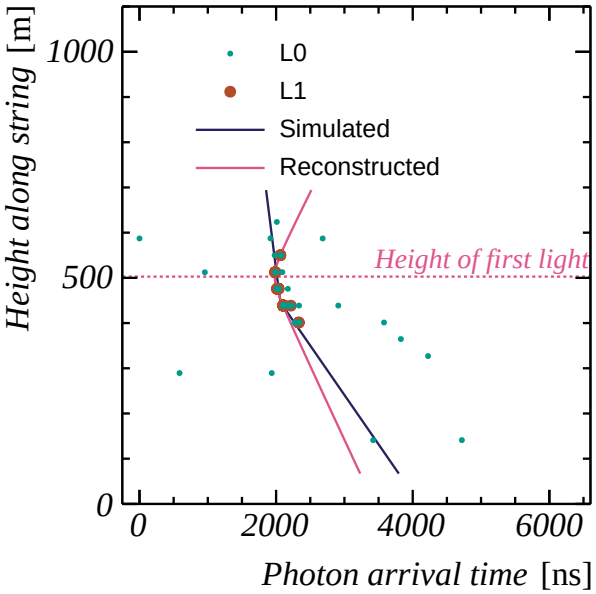


Figure 9.3: *An example of a mis-reconstructed event passing the first two event selection stages. Each dot indicates the measured time and height of a hit. All hits as well as the hits contributing to the trigger are shown. The solid lines give the arrival time of unscattered light from the true and reconstructed muon track.*

In figure 9.2, the difference between the reconstructed and true muon zenith angle is given as function of the number of hit DOMs above the height of first light. A straightforward choice for a selection criterion would be to reject all reconstructed events with one or more hit DOMs above the height of first light. Such a simple cut would, however, not only cut away a major fraction of the mis-reconstructed tracks, but all neutrino candidate events as well (see section 9.2). Therefore, the selection criterion of this stage is chosen to exclude all events with one or two hit DOMs above the height of first light, but keep all events with three or more DOMs above it.

In figure 9.4, the distribution of the difference between the reconstructed and true zenith angle is shown after each selection stage. Especially the cut on the height of closest approach is very effective to reject mis-reconstructed tracks. In table 9.1, the fraction of (mis-)reconstructed events after each selection stage is summarised.

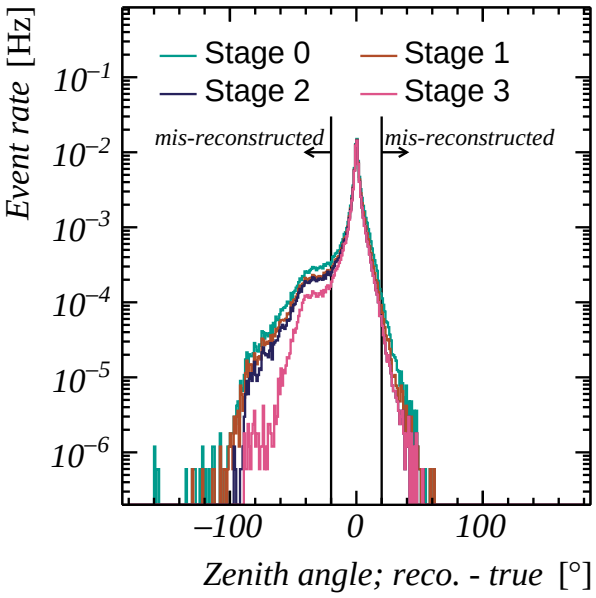


Figure 9.4: *The difference between reconstructed zenith angle and the true zenith angle of simulated atmospheric muon events. Four lines are shown, each corresponding to a stage in the event selection (see text), stage zero giving the distribution without any event selection criteria applied. The inset gives a zoom of the well-reconstructed events (maximum 20 ° difference between reconstructed and true zenith angle) with a linear vertical scale.*

Atmospheric muon simulations

Selection criterion			$\lvert\Delta\theta\rvert > 20°$		$\theta_{reco} \leq 90°$	
	$N_{ev.}$	$f_{sel.}$ [%]	$N_{ev.}$	$f_{sel.}$ [%]	$N_{ev.}$	$f_{sel.}$ [%]
none	153716	100.0	17426	11.3	1498	0.975
$Q \geq -1000$	136943	89.1	13197	8.6	1060	0.690
< 2 outliers	128635	83.7	11500	7.5	791	0.515
0 or ≥ 3 DOMs above $h_{1^{st}light}$	121004	78.7	6556	4.3	28	0.018

Data

Selection criterion			$\lvert\Delta\theta\rvert > 20°$		$\theta_{reco} \leq 90°$	
	$N_{ev.}$	$f_{sel.}$ [%]	$N_{ev.}$	$f_{sel.}$ [%]	$N_{ev.}$	$f_{sel.}$ [%]
none	180101	100.0			1421	0.789
$Q \geq -1000$	154720	85.9			829	0.460
< 2 outliers	144189	80.1			627	0.348
0 or ≥ 3 DOMs above $h_{1^{st}light}$	137350	76.3			20	0.011

Table 9.1: *Selected number of reconstructed events ($N_{ev.}$) and selected fraction of reconstructed events ($f_{sel.}$) after the various event selection criteria. The numbers in the upper table correspond to simulated events in 465.34 hours of simulated lifetime. The bottom table gives the same values for data (470.80 hours). The columns corresponds to all, mis-reconstructed and reconstructed as upward going selected events, respectively.*

In figure 9.5, the distributions of the fitted zenith angle, distance and height of closest approach and the fit quality of the selected events are shown. Even though the shape of the distributions match reasonably well, the fit quality of the reconstructed tracks in simulated events is slightly better than those of reconstructed track in data. One possible explanation could be that the detection lines are not straight (as assumed in the fit), but slightly curved as a result of sea currents (see section 3.5).

9.2 Neutrino Candidates

In this section, events that may originate from (muon)neutrinos will be identified. Since atmospheric muons cannot propagate through the Earth, up-going tracks indicate the presence of muon (anti)neutrinos interacting in the vicinity of the detector, but could correspond to mis-reconstructed events as well.

Applying the selection cuts presented in the previous section, 28 out of all simulated atmospheric muon events are (mis)reconstructed as up-going, corresponding to a rate of $1.67 \cdot 10^{-5}$ Hz. Since no neutrino interactions are present in the simulation of atmospheric muons, these events must be mis-reconstructed. To discard these events, all events with less than five hit DOMs above the height of earliest light are rejected.

This selection criterion is very strict, as none of the upward reconstructed simulated atmospheric muons pass this criterion. However, neither do any of the events observed in real data. When extending the data sample to all runs taken between the 23rd of

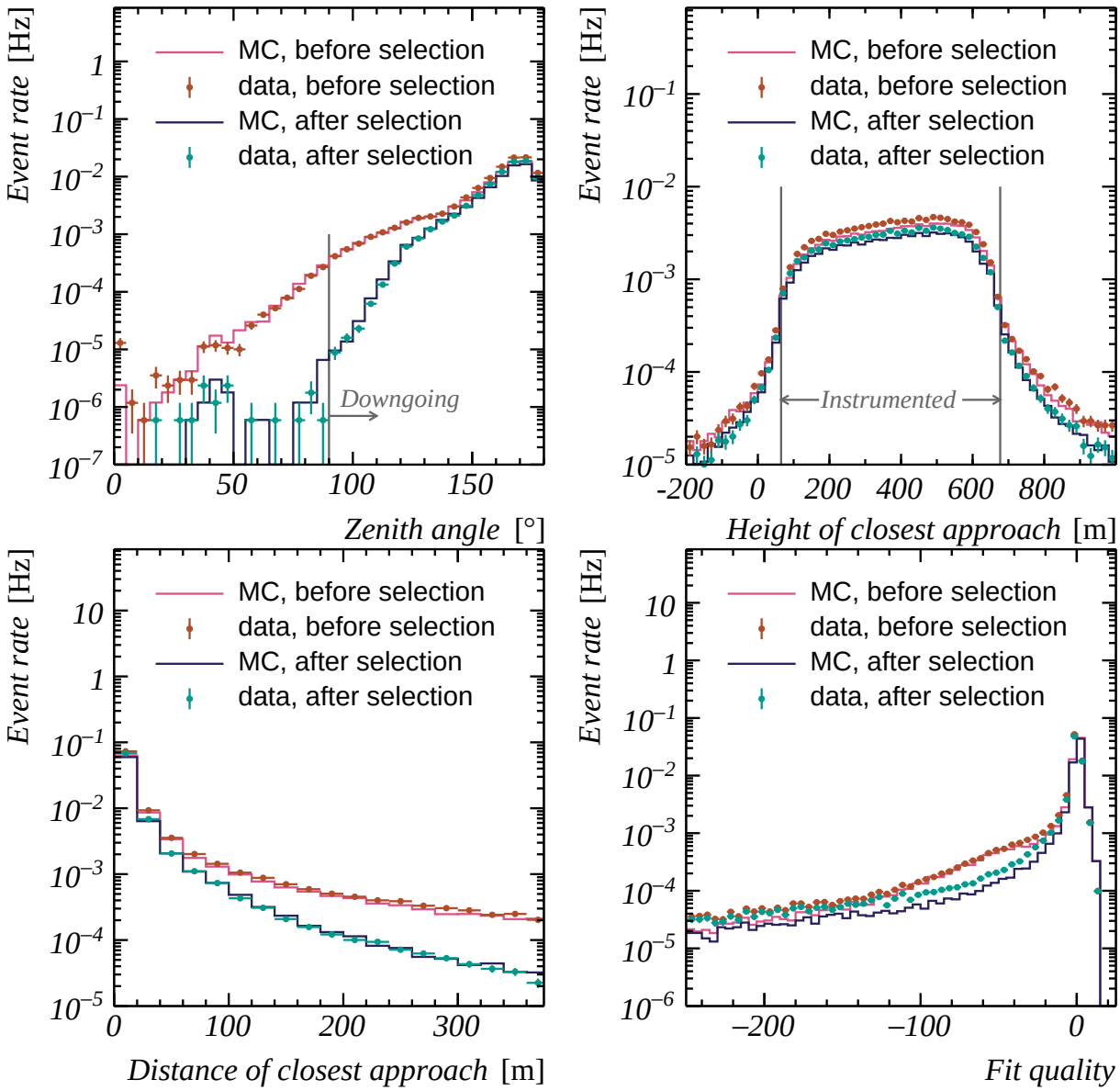


Figure 9.5: *The distributions of the reconstruction parameters before and after event selection criteria from both simulations and real data. Given are the reconstructed zenith angle (upward reconstructed events correspond to zenith angles smaller than 90 °), distance and height of closest approach and the fit quality. See section 5.3.4 for a description of these parameters.*

December 2016 and the 16 of April 2017 (2045.51 hours of data), 27 events do pass this criterion. Using the predicted number of zero events in the 465.34 hours of simulated data, a back-of-the-envelope calculation gives an indication of the probability $P_{\text{no }\nu}$ that all these events are mis-reconstructed atmospheric muon events:

$$P_{\text{no }\nu} < \int_0^{\text{inf}} e^{-465.34 \cdot R_{\mu,\uparrow}} \cdot P(27|2045.51 \cdot R_{\mu,\uparrow}) dR_{\mu,\uparrow}, \tag{9.1}$$

$$P_{\text{no }\nu} < \int_0^{\text{inf}} e^{-465.34 \cdot R_{\mu,\uparrow}} \cdot \frac{1}{27!} \cdot (2045.51 \cdot R_{\mu,\uparrow})^{27} \cdot e^{-2045.51 \cdot R_{\mu,\uparrow}} dR_{\mu,\uparrow}, \tag{9.2}$$

$$P_{\text{no }\nu} < 1.57 \cdot 10^{-6}. \tag{9.3}$$

in which $R_{\mu,\uparrow}$ denotes the (true) rate of selected atmospheric muon events per hour, reconstructed as up-going. The first term in the integral gives the probability that no upward reconstructed atmospheric muon is present in the simulations, the second term the Poisson probability that all 27 observed events are caused by upward reconstructed atmospheric muons. The result ($P_{\text{no }\nu} < 1.57 \cdot 10^{-6}$) does not include a study of systematic uncertainties. It should therefore only be interpreted as an indication.

The event displays of four selected neutrino candidates are shown in figures 9.6. It is left up to the reader to judge whether these events are genuine neutrino events or not.

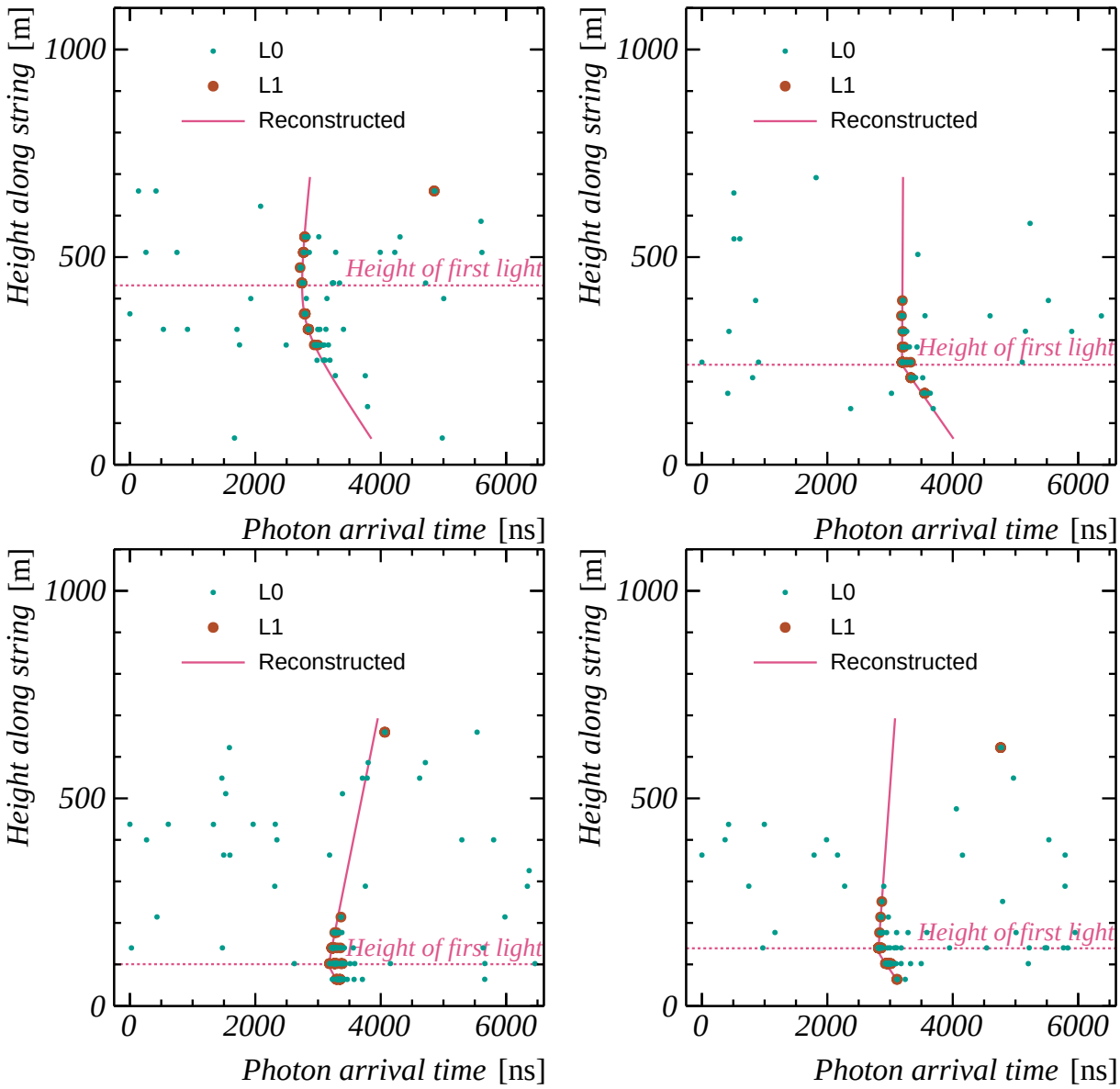


Figure 9.6: *Some selected event displays of possible neutrino events. Each dot indicates the measured time and height of a hit. All hits as well as the hits contributing to the trigger are shown. The solid line gives the arrival time of unscattered light from the reconstructed muon track.*

Part III

Neutrino Source Search

Abstract

A novel statistical method to identify and study cosmic neutrino sources will be introduced in this part. Event likelihood 'landscapes' are used to extract maximum information from the data in order to maximise the sensitivity of the method. These landscapes are combined with improved parameterisations of the expected flux, detection efficiency and detected event 'signatures'. The novel approach to neutrino source searches is very flexible, in the sense that any potentional source hypothesis can be tested, combining data taken with any detector configuration.

The event likelihood landscapes replace the reconstruction algorithms used in conventional neutrino source searches. These landscapes contain almost all available information, making event selection criteria unneccessary. Another advantage is that the novel method can easily be extended to describe any interaction signature, such as for example tau neutrino interactions, and include more information, such as the time-over-threshold.

Chapter 10

A Novel Approach

One of the main objectives of the KM3NeT collaboration is the identification of the source(s) of cosmic neutrinos (see chapters 1 to 3). Identifying such a cosmic neutrino source provides direct evidence for the origin of high-energy cosmic rays, and can help to understand the mechanism that accelerates cosmic rays to (ultra) high energies. The main challenge in the observation of such a source lies in the fact that only a few events in the (background-dominated) data sample are expected to originate from such a source.

A novel method to identify a possible neutrino source has been developed. The novel method aims at extracting the maximum amount of information out of each detected event. This maximises the potential to identify the source(s) of cosmic neutrinos, as will be discussed in the first section of this chapter. In the subsequent two sections, the novel method will be motivated and explained. The implementation of the novel method is the subject of the next chapters.

10.1 Hypothesis Testing

Neutrino source searches are based on the concepts of a *hypothesis test*. The hypothesis test is the statistical method used to make a judgement which of the two hypotheses represents the (unknown) true hypothesis best. In this, the null hypothesis ($\mathcal{H}_0$) corresponds to the background flux; originating solely from muons and neutrinos produced in the atmosphere. The alternative hypothesis ($\mathcal{H}_1$) relates to the background flux with an additional contribution from an assumed cosmic neutrino source. The data, used to make this judgement, consists of all detected events.

Given that either $\mathcal{H}_0$ or $\mathcal{H}_1$ is true, there are four possible outcomes of the test, summarised in table 10.1. In two scenarios, a hypothesis is erroneously accepted to be true. Of these two misjudgements, the incorrect acceptance of $\mathcal{H}_1$ (a *false positive* or *type I error*) is generally considered to be the worst mistake to make. Therefore, the criterion used in hypothesis tests is defined such that the corresponding probability is lower than a predefined limit (the *significance level*). In high-energy physics, commonly used significance levels are $\alpha \approx 0.27\%$ (a '3σ' level) and $\alpha \approx 0.000057\%$ (a '5σ' level). The probability $1 - \beta$ that no *false negative*, or *type II error*, is made is called the *power*, or the *sensitivity*, of the test. In general, one wants to use the most powerful hypothesis test for a given significance level.

	Accept $\mathcal{H}_0$ (reject $\mathcal{H}_1$)	Accept $\mathcal{H}_1$ (reject $\mathcal{H}_0$)
$\mathcal{H}_0$ = true	True negative ☺ $P = 1 - \alpha$ "Confidence level"	False positive Type I error $P = \alpha$ "Significance level"
$\mathcal{H}_1$ = true	False negative Type II error $P = \beta$	True positive ☺ $P = 1 - \beta$ "Power" or "sensitivity"

Table 10.1: *Possible scenarios in hypothesis testing. Each row corresponds to a true hypothesis, each column to one of the decisions made.*

10.1.1 The Neyman-Pearson Lemma

Any criterion used in hypothesis testing to differentiate between $\mathcal{H}_0$ and $\mathcal{H}_1$ can be written as a subset Ω_{accept} of all possible values of a single value $\lambda(\mathcal{D}) \subseteq \mathbb{R}$ (called the *test statistic*) assigned to each possible outcome (the data $\mathcal{D}$) of the experiment. The most powerfull hypothesis test for hypotheses which are uniquely defined (so-called simple hypotheses) was derived by Neyman, Pearson and Pearson[195, 212]. This hypothesis test uses a *likelihood ratio test statistic*:

$$\lambda(\mathcal{D}) \equiv \frac{\mathcal{L}(\mathcal{D}|\mathcal{H}_1)}{\mathcal{L}(\mathcal{D}|\mathcal{H}_0)}, \tag{10.1}$$

where $\mathcal{L}(\mathcal{D}|\mathcal{H}_1)$ and $\mathcal{L}(\mathcal{D}|\mathcal{H}_0)$ denote the conditional likelihood to observe the outcome of the experiment, given that $\mathcal{H}_1$ or $\mathcal{H}_0$ is true, respectively. These likelihoods will be discussed in the next section. The acceptance region of the Neyman-Pearson test statistic is given by:

$$\begin{cases} \lambda(\mathcal{D}) \in \Omega_{accept} & \text{if } \lambda(\mathcal{D}) > \lambda_{crit.} \quad \Rightarrow \text{accept } \mathcal{H}_1 \\ \lambda(\mathcal{D}) \notin \Omega_{accept} & \text{if } \lambda(\mathcal{D}) \leq \lambda_{crit.} \quad \Rightarrow \text{reject } \mathcal{H}_1 \end{cases}, \tag{10.2}$$

where $\lambda_{crit.}$ is chosen such that:

$$\alpha = \int_{\lambda_{crit.}}^{\infty} \frac{dP(\lambda|\mathcal{H}_0)}{d\lambda} d\lambda. \tag{10.3}$$

The function $\frac{dP(\lambda|\mathcal{H}_0)}{d\lambda}$ gives the probability density to observe a data set with a test statistic λ under the $\mathcal{H}_0$ hypothesis. An illustration of the Neyman-Pearson test is given in figure 10.1. A conceptual proof that the likelihood ratio provides the most sensitive hypothesis test is given in appendix E.

10.1.2 Composite Hypotheses

Often, the $\mathcal{H}_1$ (and sometimes the $\mathcal{H}_0$) hypothesis is not uniquely defined, but consists of a set of hypotheses, depending on one or more parameter(s) θ. A commonly encountered

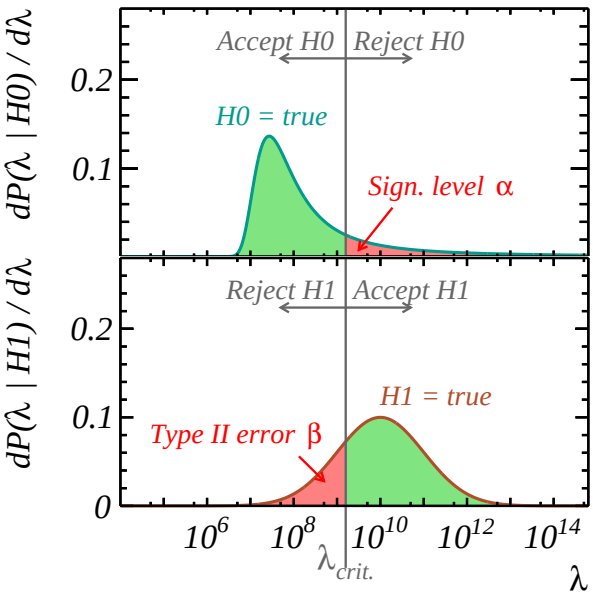


Figure 10.1: *An illustration of the Neyman-Pearson hypothesis test; the upper panel gives the probability density of the likelihood ratio test statistic λ in case the data is distributed under the $\mathcal{H}_0$hypothesis. The lower panel gives the same distribution for the case where $\mathcal{H}_1$is true. The acceptance region is indicated, as well as the probabilities associated with a false outcome of the test.*

example of such a *composite hypothesis* is given by a signal flux with an unknown location or total flux. When the $\mathcal{H}_1$ hypothesis is composite, the maximum likelihood ratio for any θ

$$\lambda(\mathcal{D}) \equiv \frac{\max_\theta \left[\mathcal{L}(\mathcal{D}|\theta, \mathcal{H}_1)\right]}{\mathcal{L}(\mathcal{D}|\mathcal{H}_0)}. \tag{10.4}$$

is often used as a test statistic. This test statistic is called the *maximum likelihood ratio*. It is proven to provide the most powerful test in case the data $\mathcal{D}$ consists of a scalar (for example the total number of events), and if the maximum likelihood ratio increases (i.e. $\mathcal{H}_1$ is increasingly likely) with increasing $\mathcal{D}$[213].

10.1.3 The Likelihood Function

The likelihood function $\mathcal{L}(\mathcal{D}|\mathcal{H})$ in the likelihood ratio test statistic (equation 10.1) denotes the likelihood to observe a certain set of data $\mathcal{D}$ for a given hypothesis $\mathcal{H}$. Like the test statistic, any function can be used. However, the power of the hypothesis test increases with the amount of available information that is included in the likelihood function[195]. In the following, the likelihood function $\mathcal{L}(\mathcal{D}|\mathcal{H})$ containing maximum information will be derived.

Counting experiment One of the simplest functions that can be used as an event likelihood is based on the probability to observe n detected events given an expected number of detected events $\mu_{\mathcal{H}}^{det}$ for hypothesis $\mathcal{H}$:

$$\mathcal{L}(n|\mathcal{H}) = \frac{e^{-\mu_{\mathcal{H}}^{det}} \cdot (\mu_{\mathcal{H}}^{det})^n}{n!}. \tag{10.5}$$

The corresponding hypothesis test is called a 'counting experiment'.

Binned likelihood In a binned likelihood function, the detected events are categorised in a set of M bins, each covering a subspace ω_j of the parameter space $\omega = \{\omega_1, \omega_2, \cdots, \omega_M\}$ of one (or more) parameter(s) $\boldsymbol{x} \in \omega$ describing each observed interaction. The event likelihood function is given by the product of M independent counting experiments, one for each bin:

$$\mathcal{L}(\{n_1, n_2, ..., n_M\}|\mathcal{H}) = \prod_{\text{bin } j=1}^{M} \left[\frac{e^{-\mu_{j,\mathcal{H}}^{det}} \cdot \left(\mu_{j,\mathcal{H}}^{det}\right)^{n_j}}{n_j!} \right], \tag{10.6}$$

$$= e^{-\mu_{\mathcal{H}}^{det}} \times \prod_{\text{bin } j=1}^{M} \left[\frac{\left(\mu_{j,\mathcal{H}}^{det}\right)^{n_j}}{n_j!} \right], \tag{10.7}$$

with n_j and $\mu_{j,\mathcal{H}}^{det}$ the detected and expected number of detected events in bin j, respectively, and $\mu_{\mathcal{H}}^{det}$ the expected total number of detected events:

$$\begin{aligned} n_j &= \text{number of events out of } \mathcal{D} \text{ in } \omega_j, \\ \mu_{j,\mathcal{H}}^{det} &\equiv \int_{\omega_j} \frac{d\Phi^{det}}{d\boldsymbol{x}}(\boldsymbol{x}|\mathcal{H}) d\boldsymbol{x}, \\ \mu_{\mathcal{H}}^{det} &= \sum_j [\mu_{j,\mathcal{H}}] = \int_{\omega} \frac{d\Phi^{det}}{d\boldsymbol{x}}(\boldsymbol{x}|\mathcal{H}) d\boldsymbol{x}. \end{aligned} \tag{10.8}$$

The binned likelihood includes a significant amount of the information contained in the data. Nevertheless, information is lost in the precise location of each event within the bins.

Unbinned likelihood, perfect accuracy In order to include the maximum amount information of each observed event, the bins in a binned likelihood function can be taken infinitesimal small. Consequently, no information is lost in the binning of the events. Such a likelihood function is called an 'unbinned' likelihood.

Correcting for the $N!/\prod_j [n_j!]$ combinations to distribute N events over the bins $\{n_1, n_2, ..., n_M\}$, equation 10.7 can be written as:

$$\mathcal{L}(\mathcal{D}|\mathcal{H}) = \frac{e^{-\mu_{\mathcal{H}}^{det}}}{N!} \times \prod_{\text{event } i=1}^{N} \left[\frac{d\Phi^{det}}{d\boldsymbol{x}}(\boldsymbol{x}|\mathcal{H}) \right], \tag{10.9}$$

with the total number of detected events $\mu_{\mathcal{H}}^{det}$ given by equation 10.8.

Unbinned likelihood, with realistic detector resolution The likelihood functions discussed in the previous paragraphs are based on the assumption that the interaction parameters $\boldsymbol{x}$ are observed with absolute precision. In realistic cases, however, the observed parameters $\mathcal{E}_i$ of each detected event are distributed according to an 'event likelihood function' $\mathcal{L}(\mathcal{E}_i|\boldsymbol{x})$, giving the probability that each event $\mathcal{E}$ is caused by a certain (set of) interaction parameter(s) $\boldsymbol{x}$. The likelihood to observe the data $\mathcal{D} = \{\mathcal{E}_1, \mathcal{E}_1, \cdots, \mathcal{E}_N\}$ for a flux $\frac{d\Phi^{det}}{d\boldsymbol{x}}(\boldsymbol{x}|\mathcal{H})$ under hypothesis $\mathcal{H}$ is therefore given by:

$$\mathcal{L}\left(\mathcal{D}|\mathcal{H}\right) = \frac{\exp\left(-\mu_{\mathcal{H}}^{det}\right)}{N!} \times \prod_{i=1}^{N} \left[\int \mathcal{L}(\mathcal{E}_i|\boldsymbol{x}) \cdot \frac{d\Phi^{det}}{d\boldsymbol{x}}(\boldsymbol{x}|\mathcal{H}) d\boldsymbol{x} \right]. \tag{10.10}$$

10.1.4 The most powerful test statistic

The combination of a likelihood ratio test statistic (equation 10.1) and an unbinned event likelihood (equation 10.10) is proven to give the most powerful hypothesis test. The corresponding test statistic is given by:

$$\lambda(\mathcal{D}) \equiv \exp\left(-\mu^{det}_{\mathcal{H}_1} + \mu^{det}_{\mathcal{H}_0}\right) \times \prod_{\text{event } i=1}^{N} \left[\frac{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)}{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)}\right], \tag{10.11}$$

where $\mathcal{L}(\mathcal{E}_i|\mathcal{H})$ is the 'test statistic integral' of event i, defined as

$$\mathcal{L}(\mathcal{E}_i|\mathcal{H}) \equiv \int \mathcal{L}(\mathcal{E}_i|\boldsymbol{x}) \cdot P^{det}(\boldsymbol{x}) \cdot \frac{d\Phi}{d\boldsymbol{x}}(\boldsymbol{x}|\mathcal{H}) d\boldsymbol{x}. \tag{10.12}$$

The used parameters are:

- $\boldsymbol{x}$ are the parameters describing each interaction (for example the position and time of the interaction).
- $\mathcal{E}$ denotes the set of observed quantities of a detected event.
- $\mathcal{L}(\mathcal{E}|\boldsymbol{x})$ is the likelihood to detect an event with observables $\mathcal{E}$ given a set of interaction parameters $\boldsymbol{x}$. This likelihood is closely related to the likelihood function used in event reconstruction (chapter 5).
- $\frac{d\Phi}{d\boldsymbol{x}}(\boldsymbol{x}|\mathcal{H})$ is the flux or interactions, as expected from a hypothesis $\mathcal{H}$.
- $P^{det}(\boldsymbol{x})$ gives the probability to detect an interaction at $\boldsymbol{x}$.
- $\mu^{det}_{\mathcal{H}}$ is the total expected number of detected events given hypothesis $\mathcal{H}$:

$$\mu^{det}_{\mathcal{H}} \equiv \int P^{det}(\boldsymbol{x}) \cdot \frac{d\Phi}{d\boldsymbol{x}}(\boldsymbol{x}|\mathcal{H}) d\boldsymbol{x}. \tag{10.13}$$

10.2 Conventional Neutrino Source Searches

Before discussing the novel approach, some commonly used strategies in conventional neutrino source searches will be discussed in this section. Three recently published studies will serve as a reference. The first two studies, using data from the ANTARES and the IceCube detector, respectively, are discussed in a paper by Albert et al.[214]. Both studies aim at the identification of a point-like or an extended neutrino source, with an unbroken energy spectrum located anywhere in the southern hemisphere. These studies will be referenced as ANT and IC, respectively. A third analysis (referred to as KM3) gives the foreseen sensitivity of the KM3NeT/ARCA detector to a set of galactic candidate sources, using models for their respective energy spectra [34]. The main characteristics of these studies are summarised in table 10.2, and will be discussed in the following.

Reconstruction The first step in each of the three analyses is the reconstruction stage. Reconstruction algorithms similar to the ones discussed in chapter 5 are used to estimate the most likely neutrino interaction parameters (position, direction and energy), the associated errors and fit quality of each event.

	ANT	IC	KM3
Source	Point-like in southern hemisphere		Galactic source
Source parameters	Source location, spectral index and flux normalisation.		Flux normalisation.
Spectrum	$dN/dE \propto E^{-\gamma}$		Source-specific models
Event selection	Classical cuts	BDT	Classical cuts
Flavour I.D.			None
Observables used in test statistic.	Neutrino direction, energy proxy and angular error estimate.		RDF

Table 10.2: *An overview of the different choices made in the neutrino source searches discussed by Albert et al.[214] and Aiello et al.[34] (see text). The abbreviations BDT and RDF stand for two machine learning algorithms; a boosted decision tree and a random decision forest, respectively.*

Event selection Background and mis-reconstructed events are rejected in an event selection stage. The criteria used in this event selection are either based on cuts on the reconstructed parameters (ANT and KM3), or (in IC) based on the output of a machine learning algorithm (a boosted decision tree, or BDT).

Flavour identification For ANT and IC, different algorithms are used to reconstruct each event, each of which is optimised for either a shower- or a track-like signature. In order to categorise each event, a step called flavour identification is implemented. The flavour identification in ANT and IC is combined with the event selection.

Test statistic In all analyses, a likelihood ratio test statistic is used (equation 10.1). In ANT and IC, the likelihood function used in the test statistic depends on the best fit neutrino direction and its estimated uncertainty, as well as a proxy of the neutrino energy (the reconstructed energy in case of IC, the number of hit PMTs in ANT). In KM3, a single observable is used, namely the output of a machine learning algorithm (more specifically a random decision forest, or RDF), trained to differentiate cosmic neutrino events from the backgrounds.

In all studies, the expected flux and detection efficiency are obtained from simulations, and parameterised as a function of the neutrino direction and energy proxy (ANT and IC), or the value of the RDF algorithm. The test statistic used in all three analyses does not include an integral as in equation 10.12. Instead, only the value of the expected flux times the event likelihood at the observed $\boldsymbol{x}_s$ is used (in this sense giving an approximate value of the full integral).

Test statistic distribution The test statistic distributions under the $\mathcal{H}_0$ and $\mathcal{H}_1$ hypotheses (needed to determine λ_{crit} and the sensitivity) are evaluated by means of pseudo-experiments: a large number of sets of randomly distributed observables are sampled from

the likelihood function. The test statistic of each of these pseudo-data sets is evaluated, resulting in a distribution of the test statistic.

10.2.1 Potential Improvements

There are a few points where information is lost in the conventional neutrino source searches:

1. All information contained in each event (i.e. the set of observed hits) is reduced to a few observables (four parameters in ANT and IC, one parameter in KM3).

2. The test statistic integral (equation 10.12) is approximated by a single point on a (relatively simple) parameterisation of the integrand, rather than using the full event likelihood 'landscape'.

3. An event selection is applied to reject background events. Since the event selection is not 100% perfect, signal events will also be rejected in this step.

These assumptions and approximations are implemented to minimise the computing time. However, since information is lost in these steps, the used hypothesis test is not as powerful as it potentially can be.

10.3 A Novel Approach

Philosophy A neutrino source search is a trade-off between the amount of information used (thus maximising the sensitivity) and the computing time required to perform the analysis. The method developed for this work is designed to approach the theoretical best sensitivity as close as possible. Assumptions and approximations are only made in case these are computationally inevitable. The risk in this extreme approach is that the novel method is computationally too demanding, and can therefore be applied only to a small fraction of the data. This is however not a critical risk, since the novel method can be used as an additional step after conventional methods, to 'squeeze' all information out of a the few events that are most likely to originate from a potential cosmic neutrino source. This possibility will be discussed in further detail in section 14.3.

Outline The basic concept of the novel approach is relatively simple. A likelihood ratio test statistic

$$\lambda(\mathcal{D}) \equiv \frac{\mathcal{L}(\mathcal{D}|\mathcal{H}_1)}{\mathcal{L}(\mathcal{D}|\mathcal{H}_0)}, \tag{10.14}$$

is combined with an unbinned likelihood function

$$\mathcal{L}\left(\mathcal{D}|\mathcal{H}\right) = \frac{\exp\left(-\mu_{\mathcal{H}}^{det}\right)}{N!} \times \prod_{\text{event } i=1}^{N} \left[\mathcal{L}(\mathcal{E}_i|\mathcal{H})\right], \tag{10.15}$$

and test statistic integral

$$\mathcal{L}(\mathcal{E}_i|\mathcal{H}) \equiv \int \mathcal{L}(\mathcal{E}_i|\boldsymbol{x}_s) \cdot P^{det}(\boldsymbol{x}_s) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) d\boldsymbol{x}_s. \tag{10.16}$$

This test statistic is proven to give the most powerful hypothesis text (see section 10.1), thus providing the best possible sensitivity to a potential cosmic neutrino source.

The most significant difference between the novel method and conventional approaches is that the test statistic integral $\mathcal{L}(\mathcal{E}_i|\mathcal{H})$ is explicitly evaluated for each event. Rather than using an appropriate form of the event likelihood function $\mathcal{L}(\mathcal{E}_i|\boldsymbol{x}_s)$ provided by reconstruction algorithms, the full event likelihood 'landscape' is used in the integral. Consequently, all information is included in the test statistic, and there is no need for event selection.

Event signature parameterisation An additional difference with respect to conventional approaches is that any signature with a single shower, a single track or a combination of these two are treated identically. This in contrast to the conventional methods, with a specific event reconstruction, flavour identification and test statistic for each event signature. To this extent, each event signature is parameterised by a shower and a muon track, both produced at the event signature position $\vec{x}_s$ at event signature time t_s, propagating along a common axis with direction $\hat{d}_s$. Two additional parameters are used to parameterise the energy E_μ of the muon and the visible shower energy $E_S^{\text{vis.}}$. The visible shower energy is defined as the energy of all light-emitting particles in the vicinity of $\vec{x}_s$ other than muons or tau-leptons. An overview of the relation between the different neutrino interaction channels and the visible shower energy and track energy is given in figure 10.2. The event signature parameterisation does not (yet) describe tau-like event signatures.

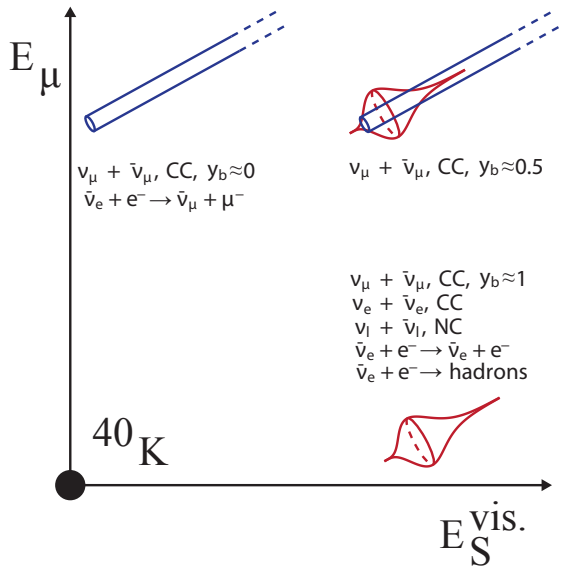


Figure 10.2: *A schematic overview of the neutrino interaction channels in relation to the visible shower energy E_S^{vis} and muon energy E_μ used in the event signature parameterisation used in the novel approach. The Bjorken-y variable of the interaction (discussed in appendix A) is denoted by y_b.*

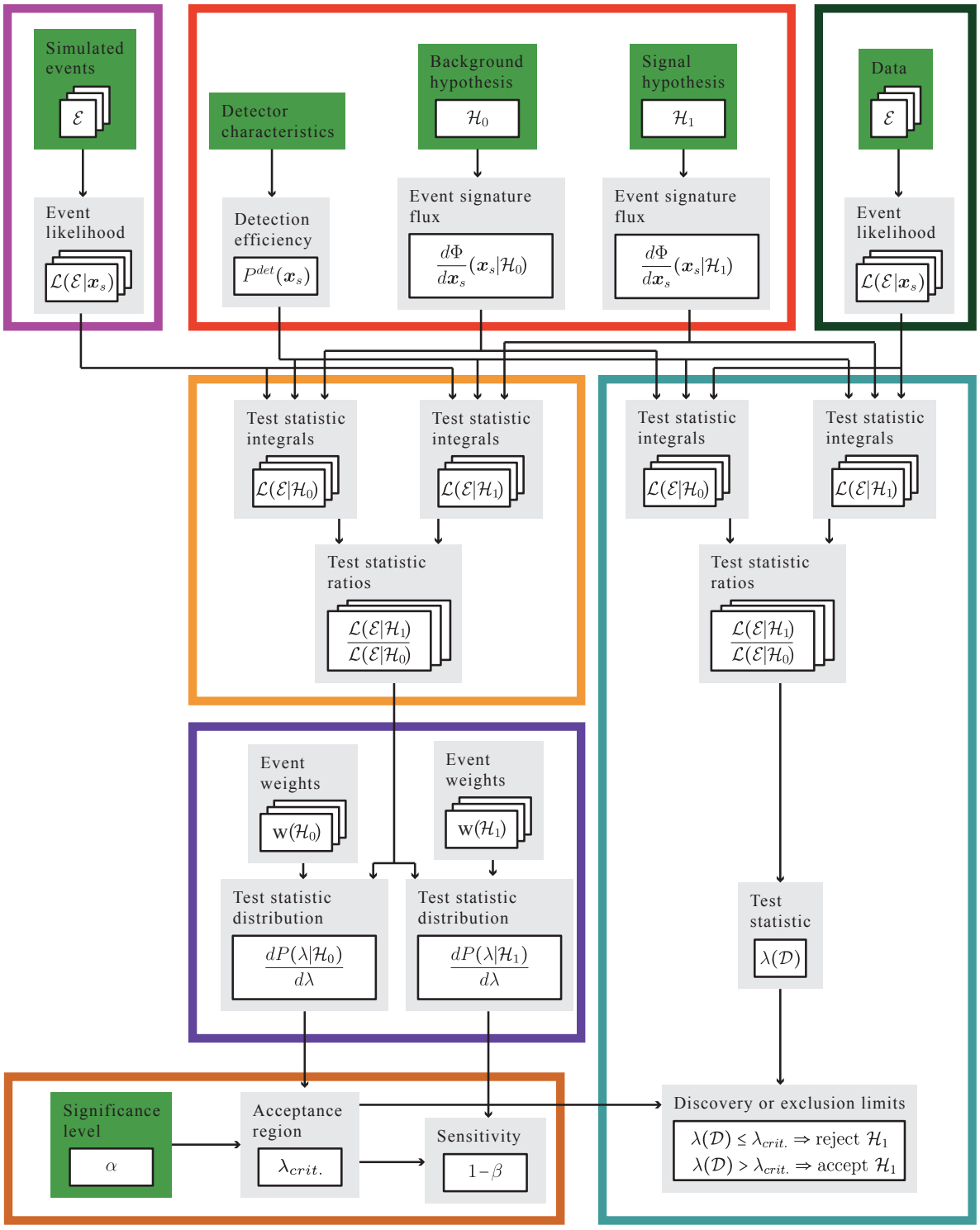


Figure 10.3: *An outline of the novel approach to neutrino source searches. The steps and quantities are discussed in the text.*

10.3.1 Testing a Potential Neutrino Source

The test of a potential $\mathcal{H}_1$ hypothesis consists of the following steps (see figure 10.3 for a schematic overview):

1. Hypothesis-independent tables Two routines are used to evaluate the function values of the detection efficiency $P^{det}(\boldsymbol{x}_s)$ and event likelihood $\mathcal{L}(\mathcal{E}_i|\boldsymbol{x}_s)$ at a large set of points covering the phase-space of $\boldsymbol{x}_s$. A third routine is used to calculate the probability $\frac{\partial^2 P}{\partial E_S^{\text{vis.}} \partial E_\mu}(E_S^{\text{vis.}}, E_\mu|E_\nu)$ that a neutrino with energy E_ν interacting through each of the different interaction channels gives an event signature with energies $E_S^{\text{vis.}}$ and E_μ. The evaluated function values are stored to disc, and will be referred to as the detection efficiency table, event likelihood tables (one table per event) and event signature probability tables (one table per interaction channel), respectively.

2. Test statistic integrals The stored tables evaluated in the previous step are used in a fourth routine, which calculates the test statistic integral $\mathcal{L}(\mathcal{E}_i|\mathcal{H})$ for both the $\mathcal{H}_0$ and a large set of $\mathcal{H}_1$ hypotheses in parallel. The hypothesis-dependent event signature flux $\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H})$ is calculated using the stored event signature probability tables and the expected neutrino interaction flux for each of the hypotheses.

3. Acceptance region The acceptance region is determined by simulating a large number of pseudo-data sets, each consisting of a number of simulated events. Event weights (see appendix C) are used to pick the events, such that the selected events represent the data taken during the lifetime of the detector under the assumption that $\mathcal{H}_0$ is true. The test statistic integrals $\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)$ and $\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)$ of the corresponding events, calculated in steps 1 and 2, are used to calculate the test statistic of each generated pseudo-data set $\mathcal{D}^{\text{pseudo}}$:

$$\lambda(\mathcal{D}^{\text{pseudo}}) = \exp\left(-\mu_{\mathcal{H}_1}^{det} + \mu_{\mathcal{H}_0}^{det}\right) \times \prod_{\text{event } i=1}^{N} \left[wi \cdot \left(\frac{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)}{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)} \right) \right], \tag{10.17}$$

where w_i is the event weight under the assumption that $\mathcal{H}_0$ is true. The acceptance region of the hypothesis test (by means of critical test statistic $\lambda_{crit.}$) can be obtained from the distribution of these test statistics. Note that this acceptance region is specific for each tested $\mathcal{H}_1$ hypothesis, and in case multiple (or composite) hypotheses are tested, the test statistic has to be calculated accordingly.

4. Sensitivity In order to calculate the sensitivity of the experiment to the $\mathcal{H}_1$ hypothesis, the same steps taken in step 3 are repeated, with the difference that the pseudo-data sets are sampled under the assumption that the $\mathcal{H}_1$ hypothesis is true. The sensitivity is given by the integral from $\lambda_{crit.}$ to infinity of the corresponding test statistic distribution.

5. Observed data The procedure to calculate the test statistic of the observed data is very similar to the previous steps. For each observed event, the event likelihood table is obtained and combined with the detection efficiency and event signature flux tables obtained in steps 1 and 2 to calculate the test statistic integral (equation 10.16) for both

$\mathcal{H}_0$ and $\mathcal{H}_1$. These values are used to calculate the test statistic according to equation 10.11. In case the observed test statistic exceeds λ_{crit}, a discovery can be claimed.

The event likelihood tables (in steps 1 and 5) are (foreseen to be) the most CPU-intensive to obtain. Once these are obtained, it is relatively straightforward to test any $\mathcal{H}_1$ hypothesis.

10.4 Advantages and Disadvantages

The challenges of the new method lie mainly in the required computing power. Significant computing time is needed to evaluate the test statistic integral (in particular the event likelihood function $\mathcal{L}(\mathcal{E}_i|\boldsymbol{x}_s)$). An additional difficulty is the more complex parameterisation of each of the ingredients (each depending on at least six event signature parameters). The complexity of these functions make it also hard to visualise these functions, which make cross-checks and verification more difficult.

In contrast to the functions, the basic principle of the novel approach is very simple and intuitive; a hypothesis can be accepted (i.e. a neutrino source can be claimed to exist) in case sufficient events have a large event likelihood in the region(s) of $\boldsymbol{x}_s$ where the expected flux from the source $P^{det}(\boldsymbol{x}_s) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}_1)$ is large (say, high-energy events from a certain point in the sky). The event signature flux $\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H})$ describes all physics (i.e. the source flux, neutrino propagation and interaction probabilities), the detection efficiency $P^{det}(\boldsymbol{x}_s)$ the performance of the detector and the event likelihood $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)$ the observed data.

The parameterisations used in conventional neutrino source searches are generally optimised for 'typical' events, assuming a Gaussian distributed event likelihood function close to its maximum. Events where this is not the case are typically rejected by the event selection. In the novel method, no assumption is made on the shape of the event likelihood function. Consequently, there is no need for an event selection; the probability that an event is mis-reconstructed or caused by background is described by the event likelihood function, which acts as a weight for the event, rather than a binary decision whether to include the event in the analysis or not. The same principle applies to events with multiple maxima in the event likelihood landscape, each of which contributes to the corresponding region of $\boldsymbol{x}_s$.

The probability to identify a neutrino source depends mainly on the accuracy of the probability density function describing the neutrino direction, energy and flavour. In conventional methods, multiple algorithms are used to determine the most likely neutrino energy and flavour. In the novel method, this information is fully described by the dependence of the event likelihood function on the visible shower energy and muon energy. Consequently, there is no need for different reconstruction algorithms or a flavour identification step.

10.5 Outline of this Part

Each of the steps of the novel approach, introduced in section 10.3.1, will be discussed in detail in the following chapters. The main focus will be on the first two steps, obtaining

the tables and combining them to calculate the test statistic integrals $\mathcal{L}(\mathcal{E}_i|\mathcal{H})$, since these two steps differ most from conventional neutrino sources. To this extent, the remainder of this thesis is structured as follows. The ingredients of the novel method will be discussed in chapter 11 (the expected event signature flux $\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H})$ and detection efficiency $P^{det}(\boldsymbol{x}_s)$) and in chapter 12 (the event likelihood $\mathcal{L}(\mathcal{E}_i|\mathcal{H})$). The evaluation of the test statistic integral ($\mathcal{L}(\mathcal{E}_i|\mathcal{H})$, e.g. equation 10.16) is the subject of chapter 13. The results will be discussed and an outlook will be given in chapter 14.

Chapter 11

Detectable Event Signature Flux

In the previous chapter, a novel approach to neutrino source searches has been introduced. The test statistic integral (equation 10.16) forms the backbone of this approach. An important ingredient of this integral is the detectable event signature flux, which is defined as the product of the event signature flux $\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H})$ and the detection efficiency $P^{det}(\boldsymbol{x}_s)$. The event signature flux and the detection efficiency will be discussed in this chapter.

11.1 Event Signature Flux

The event signature flux is defined as the total rate of neutrino interactions leading to an event signature with visible shower energy $E_S^{vis.}$ and muon energy E_μ:

$$\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) \equiv \frac{d^7\Phi}{dE_S^{vis}dE_\mu d^2\Omega d^3\vec{x}}(E_S^{vis}, E_\mu, \hat{d}_\nu, \vec{x}_\nu|\mathcal{H}) \quad \left[\left(s\cdot sr\cdot m^3\cdot GeV^2\right)^{-1}\right]. \tag{11.1}$$

The expected event signature flux depends on the (hypothesis specific) expected flux of neutrinos at the surface of the Earth, the fraction of these neutrinos that reach the detector, the probability that a neutrino interacts in the vicinity of the detector and the kinematic properties of this interaction. In the following, each of these aspects will be discussed in more detail. The calculated event signature flux $\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H})$ will be discussed in section 11.1.4.

11.1.1 Neutrino Flux

The expected flux of neutrinos for a certain hypothesis $\mathcal{H}$ is defined at the Earths' atmosphere.

Atmospheric neutrino background flux For the background-only hypothesis ($\mathcal{H}_0$), the expected neutrino flux is given by the sum of the conventional and prompt components of the atmospheric neutrino flux (see section 2.2.1). The used fluxes are given by Honda et al.[90, 215] (conventional component) and by Gauld et al. [92, 216] (prompt component), with the correction proposed by Gaisser et al. [60]. The corresponding fluxes are discussed in section 2.2.1. In both cases, the flux depends on the energy and zenith angle of the neutrino.

Non-neutrino backgrounds In addition to the (irreducible) background from atmospheric neutrinos, events caused by ^{40}K -decays, bioluminescence and atmospheric muons contribute to the detected event rate. A possible treatment of these backgrounds will be discussed in section 14.2.1. For now, the background-only hypothesis $\mathcal{H}_0$ is assumed to consist only of atmospheric neutrinos.

Cosmic neutrinos The signal hypothesis flux ($\mathcal{H}_1$) consists of the atmospheric neutrino flux, with an additional cosmic neutrino signal. As discussed in section 2.3, some physically motivated cosmic neutrino fluxes are 1) a diffuse flux of neutrinos with an unbroken power law spectrum (equation 2.7) with a $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$ flavour ratio, 2) a point-like neutrino source with an unbroken power law spectrum 3) a point-like source with a known emission spectrum and 4) a collection of point sources, sharing the same characteristics.

In this work, a set of 121 benchmark $\mathcal{H}_1$ hypotheses are used. The cosmic flux component of each of these fluxes is described by an unbroken energy spectrum

$$\frac{d}{dE_\nu}\left[\Phi_{\nu+\bar{\nu},e+\mu+\tau}\right] \equiv \Phi^{100\ \mathrm{TeV}}_{\nu+\bar{\nu},e+\mu+\tau} \cdot \left(\frac{E_\nu}{[100\mathrm{TeV}]}\right)^{-\gamma} \quad [1/\mathrm{GeV}/\mathrm{cm}^2/\mathrm{s}/\mathrm{sr}], \tag{11.2}$$

with a $\nu_e : \nu_\mu : \nu_\tau = 1 : 1 : 1$ flavour ratio and equal amount of produced neutrinos and antineutrinos. The normalisation $\Phi^{100\ \mathrm{TeV}}_{\nu+\bar{\nu},e+\mu+\tau}$ and spectral indices γ of the used benchmark fluxes are chosen to be (roughly) compatible with the observed astrophysical neutrino flux (see figure 11.1).

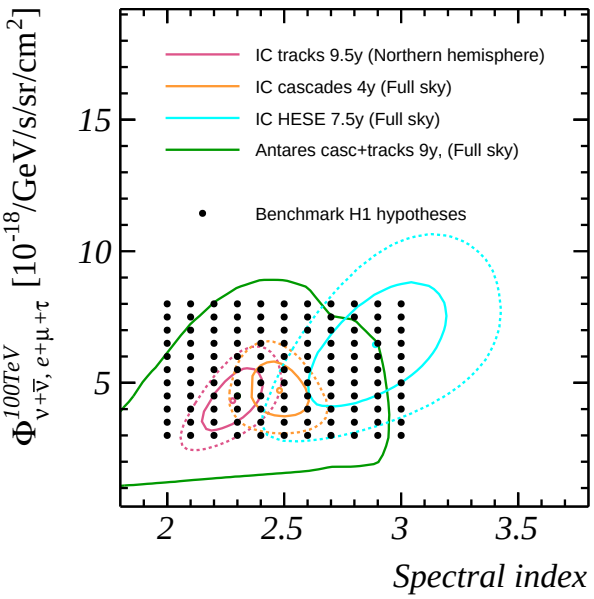


Figure 11.1: *Spectral indices γ and flux normalisation $\Phi^{100\ \mathrm{TeV}}_{\nu+\bar{\nu},e+\mu+\tau}$ of the 121 $\mathcal{H}_1$ hypothesis fluxes used in this work (see text), superimposed on measurements of the diffuse cosmic neutrino flux (see section 2.3.3).*

11.1.2 Earth Propagation

Neutrinos with a low energy can propagate almost unhindered through the Earth. The probability $P_\oplus$ that a neutrino with energy E_ν and zenith angle θ_ν traverses the Earth without interacting is given by:

$$P_{\oplus}(\theta, E_\nu) = \begin{cases} 1 & \text{if } \theta_\nu \geq 90°, \\ \exp\left(-\rho_{\oplus,tot.}(\theta_\nu) \cdot \sigma_{\text{abs.}}(E_\nu) \cdot N_A\right) & \text{else,} \end{cases} \tag{11.3}$$

where N_A is Avogadro's constant, $\rho_{\oplus,tot.}(\theta_\nu)$ is the column density of traversed matter and $\sigma_{\text{abs.}}$ gives the cross section of all interactions in which the neutrino is absorbed:

$$\sigma_{\text{abs.}} \equiv \begin{cases} \sigma_{\nu_l,CC} & \nu_e, \nu_\mu \text{ or } \nu_\tau \\ \sigma_{\bar{\nu}_e,CC} + 1/2 \cdot \sigma_{\bar{\nu}_e + e^- \to \text{hadrons}} & \bar{\nu}_e \\ \sigma_{\bar{\nu}_{\mu/\tau},CC} & \bar{\nu}_\mu \text{ or } \bar{\nu}_\tau \end{cases}$$

The cross section of $\bar{\nu}_e + e^- \to hadrons$ interactions is multiplied by 1/2 to account for the fact that an atom contains one electron for every two nucleons.

Neutral current interactions and neutrino oscillations Neutral current and $\bar{\nu}_e + e^- \to \bar{\nu}_l + l^-$ interactions produce a secondary neutrino with a slightly different direction and a lower energy than the primary neutrino. This secondary neutrino produced in neutral current interactions in the Earth can propagate to the detector, thus contributing to the total expected flux.

Since these interactions are not included in $\sigma_{\text{abs.}}$, the expected flux is overestimated at high energies and underestimated at low energies. In order to quantify the effects of neglecting neutral current interactions, figure 11.2 (right) shows the probability that a neutrino traversed the Earth but underwent at least one neutral current interaction in the Earth:

$$P_\oplus \cdot P_{\text{NC}} = P_\oplus \cdot \left(1 - \exp\left(-\rho_{\oplus,tot.}(\theta_\nu) \cdot \sigma_{\text{NC}}(E_\nu) \cdot N_A\right)\right), \tag{11.4}$$

in which σ_{NC} is the neutral current interaction cross-section. At most about 12 % of the neutrinos with an initial energy above 10^5 GeV that propagated through the Earth underwent at least one neutral current interaction. The energy loss and angle of deflection at these energies is relatively small (see appendix A).

Two additional contributions that are not taken into account in this work are neutrino oscillations in the Earth (this effect is negligible for neutrinos with energies above 100 GeV[135]) and tau regeneration. Tau regeneration is the process in which tau neutrino undergoes a charged current interaction in the Earth, with consecutive decay of the tau lepton to a secondary neutrino with less energy. Like neutral current interactions, tau regeneration can be considered as an energy loss process of neutrinos. This effect will be discussed in more detail in section 14.2.2.

The PREM density profile of the Earth The integrated column density $\rho_{\oplus,tot.}(\theta_\nu)$ of the Earth is calculated using a model of the density profile of the Earths' interior, called the *preliminary reference earth model* (PREM). This model is obtained from seismic data. A description of this model and the calculation of the column density are given in appendix F. In figure 14.9 (right), the integrated column density is shown. Using equation 11.3, the probability $P_\oplus(\theta_\nu, E_\nu)$ that a neutrino traverses the Earth without interacting is given in figure 11.2 (left).

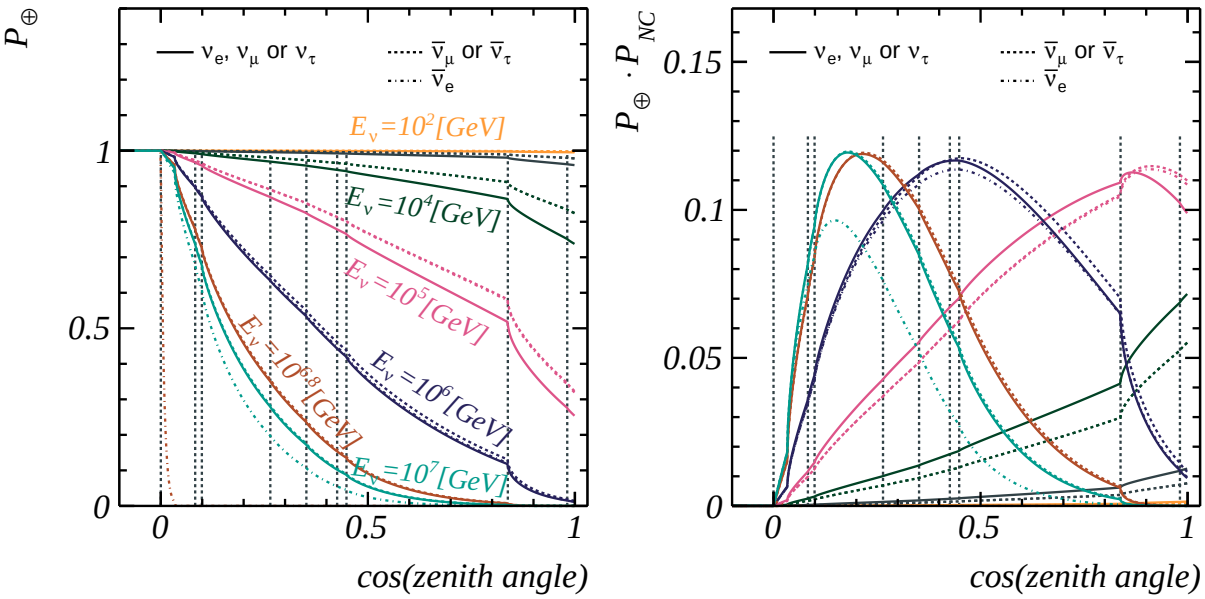


Figure 11.2: ***Left:*** *The probability that a neutrino propagates through the Earth without interacting through charged-current interactions. The probability is given for neutrinos (solid lines) and antineutrinos (dashed lines) of several energies. The vertical black dashed lines indicate the different density regions of the PREM.* ***Right:*** *The probability that a neutrino traverses the Earth (i.e. without interacting through a charged current interaction), but does interact (at least once) through the neutral current interaction. The used line colours and styles are the same as those used in the left figure.*

11.1.3 Neutrino Interaction Flux

The expected flux of interacting neutrinos in the vicinity of the detector $\frac{d^6\Phi_\nu^{int.}}{dE_\nu d^2\Omega d^3\vec{x}}$ (i.e. the number of interactions per GeV, second, m^3 and staradian) for a given hypothesis $\mathcal{H}$ is given by:

$$\frac{d^6\Phi_\nu^{int.}}{dE_\nu d^2\Omega d^3\vec{x}}(E_\nu, \hat{d}_\nu, \vec{x}_\nu|\mathcal{H}) = R_{\text{int.}}(E_\nu, \vec{x}_\nu) \cdot P_\oplus(E_\nu, \hat{d}_\nu) \cdot \frac{d^5\Phi_\nu^{atm.}}{dE_\nu d^2\Omega d^2A}(E_\nu, \hat{d}_\nu|\mathcal{H}), \quad (11.5)$$

where $\frac{d^5\Phi_\nu^{atm.}}{dE_\nu d^2\Omega d^2A}(E_\nu, \hat{d}_\nu|\mathcal{H})$ is the neutrino flux (per GeV, second, m^2 and staradian) at the Earths' atmosphere for hypothesis $\mathcal{H}$, $P_\oplus$ is the probability that a neutrino traverses the Earth and $R_{\text{int.}}(E_\nu, z_\nu)$ gives the number of neutrino interactions per meter of traversed matter. The latter quantity is given by the product of the density of the medium at the interaction position $\rho(z_\nu)$ (approximately 1.038 g/cm^3 in the seawater and 2.65g/cm^3 in the seabed), Avogadro's constant N_A and the interaction cross section $\sigma(E_\nu)$:

$$R_{\text{int.}}(E_\nu, \vec{x}_\nu) = \sigma(E_\nu) \cdot N_A \cdot \rho(z_\nu). \quad (11.6)$$

The interaction cross section $\sigma(E_\nu)$ depends on the neutrino type (flavour and neutrino/antineutrino) and interaction channel. The corresponding fluxes of interacting neutrinos are given in figure 11.3, for both the background only ($\mathcal{H}_0$) hypothesis and a benchmark signal flux ($\Phi^{100\,\text{TeV}}_{\nu+\bar{\nu},e+\mu+\tau} = 6 \cdot 10^{-18}$ [GeV/s/sr/cm^2], $\gamma = 2.5$). The neutrino interaction fluxes have some noteworthy features:

- The difference between the conventional atmospheric flux of $\bar{\nu}_e + \nu_e$ and the $\bar{\nu}_\mu + \nu_\mu$ flux is clearly observable in the $\mathcal{H}_0$ flux.
- The background-only flux of interacting tau neutrinos is zero, since no tau neutrinos are produced in cosmic ray interactions in the atmosphere (see section 2.2.1) and neutrino oscillations in the Earth are (assumed to be) negligible.
- The absorption of neutrinos in the Earth has an significant effect on the flux, especially for neutrinos with high energy propagating through the denser regions of the Earth ($\theta_\nu \approx 0^\circ$).
- The flux of electron antineutrinos interacting with electrons in the detector medium is generally negligible in comparison with neutrino-nucleon interactions. However, at energies close to 6.3 PeV (the Glashow resonance, see section 3.2.2), the $\bar{\nu}_e + e^-$ interaction cross section is strongly enhanced, and the corresponding flux of interacting neutrinos dominates the total flux. This is, however, only the case for downward and horizontally moving neutrinos ($\theta_\nu \gtrsim 90^\circ$), electron antineutrinos with these energies moving through the Earth are very likely to be absorbed in the Earth before reaching the detector.
- In the benchmark cosmic neutrino flux (i.e. the $\mathcal{H}_1 - \mathcal{H}_0$), (anti)neutrinos of different flavours are assumed to reach the Earth in equal numbers. Consequently, only the absorption of electron anti-neutrinos in the Earth through $\bar{\nu}_e + e^-$ interactions gives a noticeable difference between the fluxes of different flavours.
- The flux of interacting neutrinos from cosmic origin dominates the total flux at neutrino energies above approximately 1 TeV.

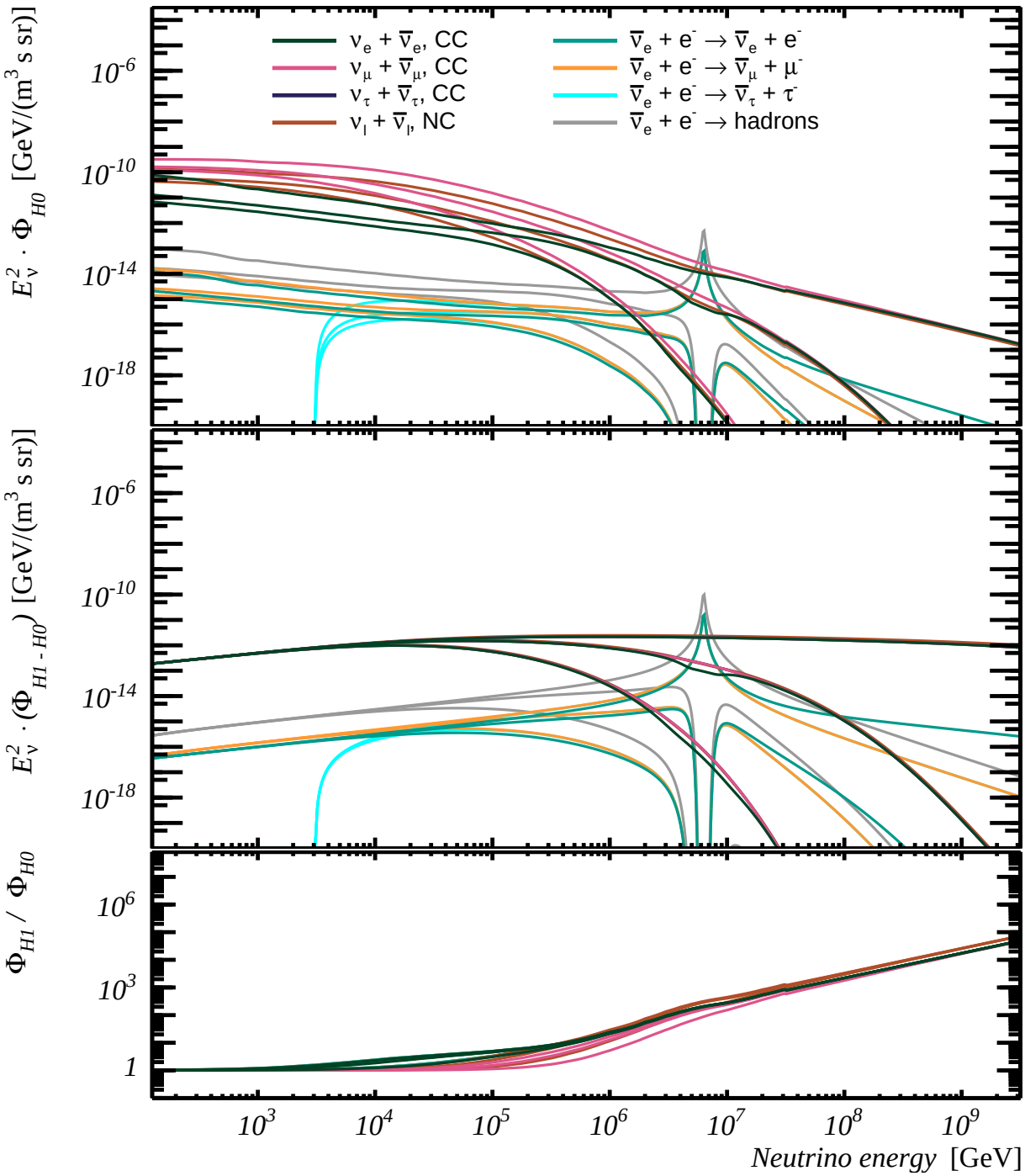


Figure 11.3: *The expected flux of interacting neutrinos* ($\frac{d^6\Phi_\nu^{int.}}{dE_\nu d^2\Omega d^3\vec{x}}$) *in seawater as a function of the neutrino energy. Each of the coloured lines gives the combined flux of neutrinos (and antineutrinos) giving an indifferentiable event signature (see text). Lines with the same colour correspond to neutrinos with different zenith angles:* $\theta_\nu = 0°$ *(neutrinos propagating right through the centre of the Earth),* $\theta_\nu = 60°$ *and* $\theta_\nu = 90°$ *(horizontally moving neutrinos). Of these, the line with the highest flux corresponds to* $\theta_\nu = 90°$ *, the line at the lowest flux to* $\theta_\nu = 0°$ *. The background-only* ($\mathcal{H}_0$) *flux (multiplied by* E_ν^2) *is given in the* ***upper*** *panel, a benchmark signal-only flux* ($\Phi^{100\ \mathrm{TeV}}_{\nu+\bar{\nu},e+\mu+\tau} = 6 \cdot 10^{-18}$ *[GeV/s/sr/cm*^{2}*], gamma* = 2.5*) is given in the* ***middle*** *figure. The* ***bottom*** *figure gives the ratio between the flux under the* $\mathcal{H}_1$ *and the* $\mathcal{H}_0$ *hypothesis.*

11.1.4 Event Signature Flux

The event signature flux $\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H})$ (not to be confused with the neutrino *interaction* flux discussed in the previous section) is defined as the rate of event signatures with energies $E_S^{\text{vis.}}$ and E_μ and a direction $\hat{d}_s$ occurring per unit volume $d^3\vec{x}$ and solid angle $d^2\Omega$. The event signature flux is summed over all 16 – 3 = 13 (anti)neutrino interaction channels in which no tau lepton is produced. Inclusion of charged-current tau (anti-)neutrino and $\bar{\nu}_e + e^- \to \bar{\nu}_\tau + \tau^-$ interactions will be discussed in section 14.2.

For each interaction channel, the event signature flux relates to the flux of neutrino interactions $\frac{d^6\Phi_\nu^{int.}}{dE_\nu d^2\Omega d^3\vec{x}}(E_\nu, \hat{d}_\nu, \vec{x}_\nu|\mathcal{H})$ (equation 11.5) via:

$$\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) = \int_0^\infty \frac{\partial^2 P}{\partial E_S^{\text{vis.}} \partial E_\mu}(E_S^{\text{vis.}}, E_\mu|E_\nu) \cdot \frac{d^6\Phi_\nu^{int.}}{dE_\nu d^2\Omega d^3\vec{x}}(E_\nu, \hat{d}_\nu, \vec{x}_\nu|\mathcal{H}) dE_\nu, \tag{11.7}$$

in which the *event signature probability* $\frac{\partial^2 P}{\partial E_S^{\text{vis.}} \partial E_\mu}(E_S^{\text{vis.}}, E_\mu|E_\nu)$ gives the probability density that a neutrino of energy E_ν interacting through this channel gives an event signature with $E_S^{\text{vis.}}$ and E_μ.

11.1.5 Event Signature Probability

The event signature probability density function $\frac{\partial^2 P}{\partial E_S^{\text{vis.}} \partial E_\mu}(E_S^{\text{vis.}}, E_\mu|E_\nu)$ depends on the neutrino energy, flavour and interaction channel. This probability density function depends on two distributions:

- The probability density function of the *Bjorken-y* variable (y_b). In the rest frame of the detector, y_b gives the ratio between the energy of the outgoing lepton or neutrino (in charged-current or neutral-current interactions, respectively) and the energy of the interacting neutrino. The probability density function of y_b for a given neutrino energy, denoted by $\frac{\partial P}{\partial y_b}(y_b|E_\nu)$, depends on the neutrino energy, flavour and interaction channel. The probability density function of y_b is proportional to the differential interaction cross section (see appendix A) and is shown in figure 11.4 (left).

- For hadronic showers; the probability density function $\frac{\partial P}{\partial F_S^{\text{vis.}}}(F_S^{\text{vis.}}|E_S)$ of the ratio $F_S^{\text{vis.}} \equiv E_S^{\text{vis.}}/E_S$ between the visible shower energy E_S^{vis} and the total (hadronic) shower energy E_S. This PDF depends only on the shower energy E_S. A two-dimensional histogram, obtained from simulations (see section 4.4.1) is used to parameterise $\frac{\partial P}{\partial F_S^{\text{vis.}}}(F_S^{\text{vis.}}|E_S)$. In figure 11.4 (right) this histogram is shown. The probability density is evaluated by using linear interpolation of the cumulative probability density function

$$P(F_S^{\text{vis.}} \leq x|E_S) \equiv \int_0^x \frac{\partial P}{\partial F_S^{\text{vis.}}}(F_S^{\text{vis.}}|E_S) dF_S^{\text{vis.}}. \tag{11.8}$$

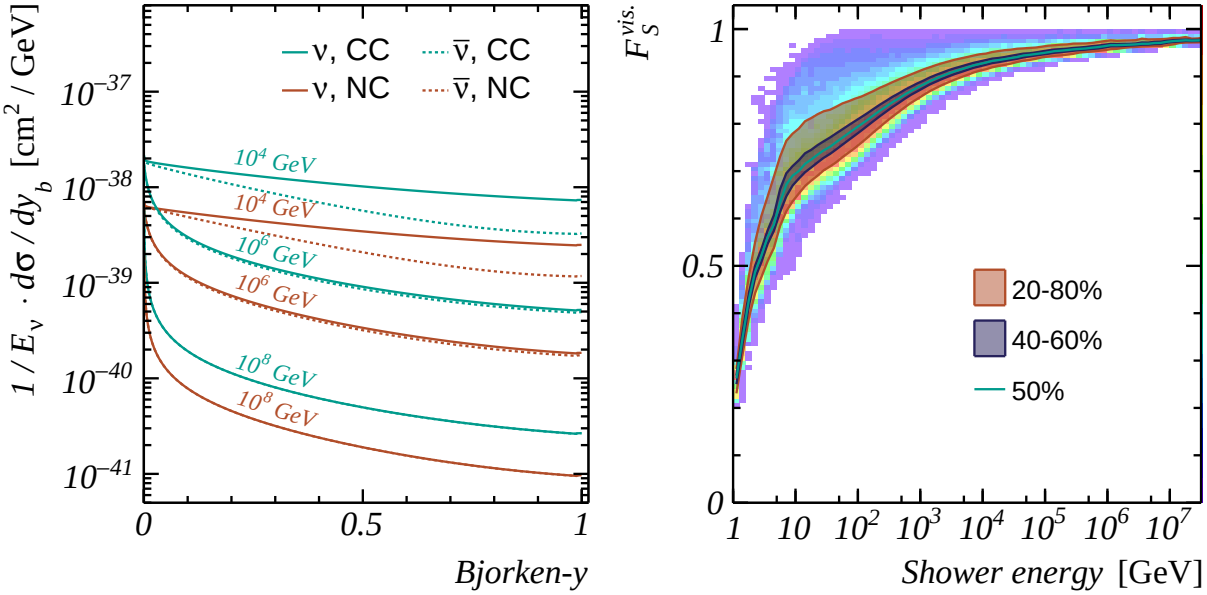


Figure 11.4: ***Left:*** *The differential neutrino-nucleus interaction cross section as function of the fraction of neutrino energy transferred to the nucleus (Bjorken-y, or y_b) for various energies of the interacting neutrino. Each distribution is divided by the neutrino energy.* ***Right:*** *The probability density of the ratio $F_S^{vis.} \equiv E_S^{vis}/E_S$ between the visible shower energy and the shower energy of hadronic showers. The maximum of each slice in the histogram is set to one. The coloured bands give the percentiles of the PDF.*

Interaction channel	E_S^{vis}/E_ν	E_μ/E_ν	$\frac{\partial^2 P}{\partial E_S^{\text{vis.}} \partial E_\mu}(E_S^{\text{vis.}}, E_\mu \vert E_\nu)$
ν_e or $\bar{\nu}_e$, CC	$y_b \cdot F_S^{\text{vis.}} + (1-y_b)$	0	$\frac{\delta(E_\mu=0)}{E_\nu} \cdot \int_0^1 \frac{\partial P}{\partial y_b}\left(y_b = \frac{E_\nu - E_S^{vis}}{E_\nu \cdot (1-F_S^{\text{vis.}})}, E_\nu\right) \cdot \frac{\partial P}{\partial F_S^{\text{vis.}}}\left(F_S^{\text{vis.}}, E_S = \frac{E_\nu - E_S^{vis}}{(1-F_S^{\text{vis.}})}\right) \cdot (E_\nu \cdot (1-F_S^{\text{vis.}}))^{-1} \; dF_S^{\text{vis.}}$
ν_μ or $\bar{\nu}_\mu$, CC	$y_b \cdot F_S^{\text{vis.}}$	$1-y_b$	$\frac{\partial P}{\partial y_b}\left(y_b = 1 - \frac{E_\mu}{E_\nu}, E_\nu\right) \cdot E_\nu^{-1} \cdot \frac{\partial P}{\partial F_S^{\text{vis.}}}\left(F_S^{\text{vis.}} = \frac{E_S^{vis}}{E_\nu - E_\mu}, E_S = E_\nu - E_\mu\right) \cdot (E_\nu - E_\mu)^{-1}$
ν_l or $\bar{\nu}_l$, NC	$y_b \cdot F_S^{\text{vis.}}$	0	$\delta(E_\mu = 0) \cdot \int_0^1 \frac{\partial P}{\partial y_b}\left(y_b = \frac{E_S^{vis}}{E_\nu \cdot F_S^{\text{vis.}}}, E_\nu\right) \cdot \frac{\partial P}{\partial F_S^{\text{vis.}}}\left(F_S^{\text{vis.}}, E_S = \frac{E_S^{vis}}{F_S^{\text{vis.}}}\right) \cdot (E_\nu \cdot F_S^{\text{vis.}})^{-1} \; dF_S^{\text{vis.}}$
$\bar{\nu}_e + e^- \to \bar{\nu}_e + e^-$	y_b	0	$\delta(E_\mu = 0) \cdot \frac{\partial P}{\partial y_b}\left(y_b = \frac{E_S^{vis}}{E_\nu}, E_\nu\right) \cdot E_\nu^{-1}$
$\bar{\nu}_e + e^- \to \bar{\nu}_\mu + \mu^-$	0	y_b	$\delta(E_S^{vis} = 0) \cdot \frac{\partial P}{\partial y_b}\left(y_b = \frac{E_\mu}{E_\nu}, E_\nu\right) \cdot E_\nu^{-1}$
$\bar{\nu}_e + e^- \to$ hadrons	$F_S^{\text{vis.}}$	0	$\delta(E_\mu = 0) \cdot \frac{\partial P}{\partial F_S^{\text{vis.}}}\left(F_S^{\text{vis.}} = \frac{E_S^{vis}}{E_\nu}, E_S = E_\nu\right) \cdot E_\nu^{-1}$

Table 11.1: *An overview of the relation between the visible shower energy E_S^{vis}, muon energy E_μ and neutrino energy E_ν for different neutrino flavours and interaction channels. The charged-current $\nu_\tau + \bar{\nu}_\tau$ and $\bar{\nu}_e + e^- \to \bar{\nu}_\tau + \tau^-$ interaction channels are not included in this work, and therefore not given in this table. The probability density functions $\frac{\partial P}{\partial y_b}(y_b|E_\nu)$ and $\frac{\partial P}{\partial F_S^{\text{vis.}}}(F_S^{\text{vis.}}|E_S)$ are discussed in the text. The Dirac delta function $\delta(x)$ is one in case $x = 0$ and zero otherwise.*

The relations between the event signature energies and the neutrino energy via y_b and $F_S^{\text{vis.}}$ for each of the interaction channels are given in table 11.1. For neutral current interactions, the event signature probability is independent of the flavour of the interacting neutrino, leading to a total of 13 – 4 = 9 different event signature probability density functions.

11.1.6 Event Signature Probability Tables

The event signature probability is calculated at a set of 102x102 fixed points covering $E_S^{\text{vis.}}$ and $E_\mu \in \{0, 10^0, 10^{0.1}, 10^{0.2}, \cdots, 10^{10}\}$ GeV each, and 101 variable abscissae in E_ν. The calculated values are stored to disk to be used in the evaluation of the event signature flux (see next subsection). The corresponding tables, one for each neutrino interaction channel, are called *event signature probability tables*. For numerical accuracy, the cumulative probability

$$P^{\text{cum.}}_{\text{evt. sig.}}(E_S^{\text{vis.}}, E_\mu | E_\nu) \equiv \int_0^{E_S^{\text{vis.}}, E_\mu} \frac{\partial^2 P}{\partial E_S^{\text{vis.}} \partial E_\mu}(E_S^{\text{vis., obs.}}, E_\mu^{\text{obs.}} | E_\nu) dE_S^{\text{vis., obs}} dE_\mu^{\text{obs.}}, \tag{11.9}$$

is calculated and stored, rather than the probability density. This cumulative probability increases with $E_S^{\text{vis.}}$ or E_ν and, consequently, decreases with increasing E_ν.

For the channels in which only a shower is produced, the event signature probability is non-zero only if $E_\mu = 0$ and $E_S^{\text{vis.}} \leq E_{nu}$. Consequently, the corresponding cumulative event signature probability $P^{\text{cum.}}_{\text{evt. sig.}}(E_S^{\text{vis.}}, E_\mu | E_\nu)$ is one in case $E_\nu \leq E_S^{\text{vis.}}$. Similarly, for $\bar{\nu}_e + e^- \to \bar{\nu}_\mu + \mu^-$ interactions, the event signature probability is non-zero only if $E_S^{\text{vis.}} = 0$ and $E_\mu \leq E_\nu$, therefore giving $P^{\text{cum.}}_{\text{evt. sig.}}(E_S^{\text{vis.}}, E_\mu | E_\nu) = 1$ in case $E_\nu \leq E_\mu$. Charged current ν_μ and $\bar{\nu}_\mu$ interactions form an exception to the previous points. This is easily explained by the fact that a large fraction of the neutrino energy goes either into the (hadronic) shower or to the muon, such that either one to two conditions $E_S^{\text{vis., obs.}} < E_S^{\text{vis.}}$ or $E_\mu^{\text{obs.}} < E_\mu$ is not met if $\min(E_S^{\text{vis.}}, E_\mu)\, E_\nu \leq E_S^{\text{vis.}} + E_\mu$. Both conditions are met in case $E_\nu \leq \min(E_S^{\text{vis.}}, E_\mu)$ and neither of the two conditions are met in case $E_\nu \gg E_S^{\text{vis.}} + E_\mu$.

For each interaction channel, $E_S^{\text{vis.}}$ and E_μ, the abscissae in E_ν are chosen to cover the cumulative event signature probability approximately isotropically. The evaluation of the cumulative probability and the choice of E_ν abscissae depends on the interaction channel:

- ν_e **and** $\bar{\nu}_e$**, CC**: First, a binary search is performed for the highest E_ν at which $P\left(F_S^{\text{vis.}} = E_S^{\text{vis.}}/E_\nu | E_S = E_\nu\right) > 0$. The neutrino energy is covered with 101 points linear in $\log(E_\nu)$ ranging from $E_S^{\text{vis.}}$ to the found upper limit. At each of these points, the cumulative event signature probability is given by the following integral:

$$P^{\text{cum.}}_{\text{evt. sig.}} = \int_0^1 P(F_S^{\text{vis.}} \leq 1 + \frac{E_S^{\text{vis.}}/E_\nu - 1}{y_b} | E_S = E_\nu \cdot y_b) \cdot \frac{\partial P}{\partial y_b}(y_b | E_\nu) dy_b. \tag{11.10}$$

  .

- ν_μ **and** $\bar{\nu}_\mu$**, CC**: These two interaction channels are the only ones where both $E_S^{\text{vis.}} > 0$ and $E_\mu > 0$. For each of the points in $E_S^{\text{vis.}}, E_\mu$, the 101 neutrino energies are determined at which $P^{\text{cum.}}_{\text{evt. sig.}} = \{0, 0.01, 0.02, \cdots, 1\}$. The corresponding neutrino energies

are found by a binary search, in which $\log(E_\nu)$ is varied between $\log(\min(E_S^{\text{vis.}}, E_\mu))$ and $\log(10^{10})$ to solve

$$P_{\text{evt. sig.}}^{\text{cum.}} = \int_{y_b=1-\frac{E_\mu}{E_\nu}}^{1} \frac{\partial P}{\partial y_b}(y_b|E_\nu)dy_b \cdot P(F_S^{\text{vis.}} \leq \frac{E_S^{\text{vis.}}}{E_S}|E_S = E_\nu - E_\mu)dy_b \tag{11.11}$$

.

- ν **and** $\bar{\nu}$**, NC**: For this channel, the event signature probability is independent of the flavour of the neutrino. The neutrino energy is sampled with equidistant steps in $\log(E_\nu)$ between $\log(E_S^{\text{vis.}})$ and $\log\left(E_S^{\text{vis.}}/y_{b,min.}/F_{S,min.}^{\text{vis.}}\right)$, where $y_{b,min} = 10^{-10}$ and $F_{S,min.}^{\text{vis.}}) = 0.16$ are taken from the range of the used probability density functions. At each $E_S^{\text{vis.}}$ and E_ν, the cumulative event signature probability is calculated by solving the following integral:

$$P_{\text{evt. sig.}}^{\text{cum.}} = \int_{y_{b,min.}}^{1} P(F_S^{\text{vis.}} \leq E_S^{\text{vis.}}/E_S|E_S = E_\nu \cdot y_b) \cdot \frac{\partial P}{\partial y_b}(y_b|E_\nu)dy_b \tag{11.12}$$

.

- $\bar{\nu}_e + e^- \rightarrow \bar{\nu}_e + e^-$: The cumulative event signature probability $P(E_S^{\text{vis., obs.}} < E_S^{\text{vis.}}, E_\mu^{\text{obs.}} < E_\mu|E_\nu)$ is sampled at $\{0., 0.01, 0.02, \cdots, 1.\}$. The corresponding neutrino energy is determined using a binary search for the neutrino energy at which

$$P_{\text{evt. sig.}}^{\text{cum.}} = \int_{0}^{y_b=E_S^{\text{vis.}}/E_\nu} \frac{\partial P}{\partial y_b}(y_b|E_\nu)dy_b \tag{11.13}$$

.

- $\bar{\nu}_e + e^- \rightarrow \bar{\nu}_\mu + \mu^-$: The same as $\bar{\nu}_e + e^- \rightarrow \bar{\nu}_e + e^-$, but with E_ν corresponding to the point at which

$$P_{\text{evt. sig.}}^{\text{cum.}} = \int_{0}^{y_b=E_\mu/E_\nu} \frac{\partial P}{\partial y_b}(y_b|E_\nu)dy_b \tag{11.14}$$

.

- $\bar{\nu}_e + e^- \rightarrow$ **hadrons**: This channel is treated identically as the previous two $\bar{\nu}_e + e^-$ interaction channels, except for using

$$P_{\text{evt. sig.}}^{\text{cum.}} = P(F_S^{\text{vis.}} \leq x = E_S^{\text{vis.}}/E_\nu|E_S = E_\nu) \tag{11.15}$$

to determine the neutrino energy.

The cumulative probability functions and the evaluated abscissae are shown in figures 11.5 and 11.6.

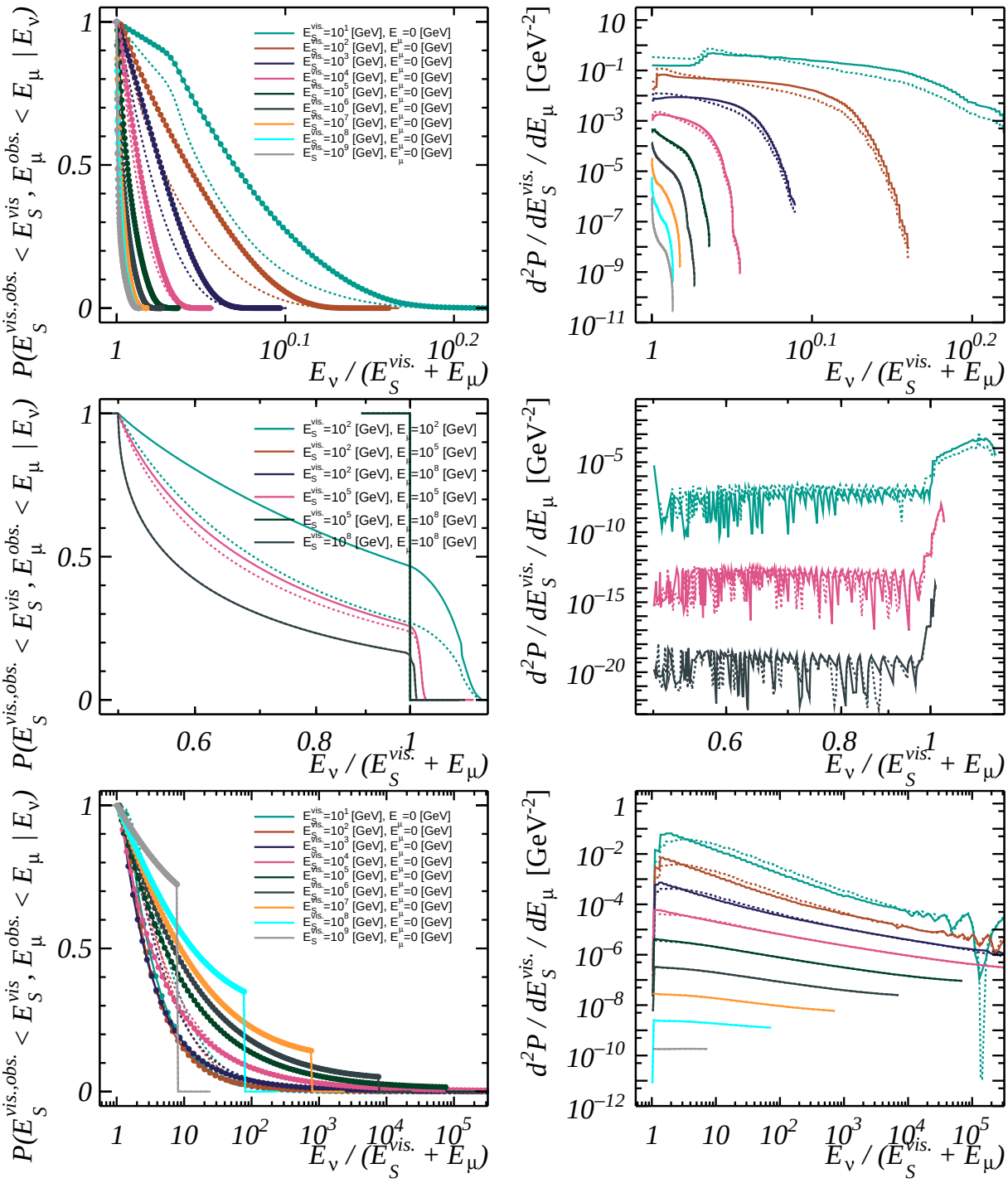


Figure 11.5: *The cumulative probability* $P(E_S^{\text{vis., obs.}} < E_S^{\text{vis.}}, E_\mu^{\text{obs.}} < E_\mu | E_\nu)$ *that the observed event signature energies* $E_S^{\text{vis., obs.}}$, $E_\mu^{\text{obs.}}$ *are smaller than* $E_S^{\text{vis.}}$, E_μ *(figures on the* ***left*** *side) and the probability density* $\frac{\partial^2 P}{\partial E_S^{\text{vis.}} \partial E_\mu}(E_S^{\text{vis.}}, E_\mu | E_\nu)$ *(****right*** *column) for a selection of different* $E_S^{\text{vis.}}$, E_μ *(each with a different charged-current* ν_e *and* $\bar{\nu}_e$ *interactions (top)). These probabilities are given as a function of the energy* E_ν *of the interacting neutrino, divided by the sum of (the various, but fixed)* $E_S^{\text{vis.}}$, E_μ*. Each row corresponds to a different interaction channel: charged-current* ν_e *and* $\bar{\nu}_e$ *interactions (****top****), charged-current* ν_μ *and* $\bar{\nu}_\mu$ *interactions (****middle****) and neutral current interactions of any flavour in the* ***bottom*** *figures. Neutrinos are distinguished from antineutrinos by* ***solid*** *(*ν*) and* ***dashed*** *(*$\bar{\nu}$*) lines.*

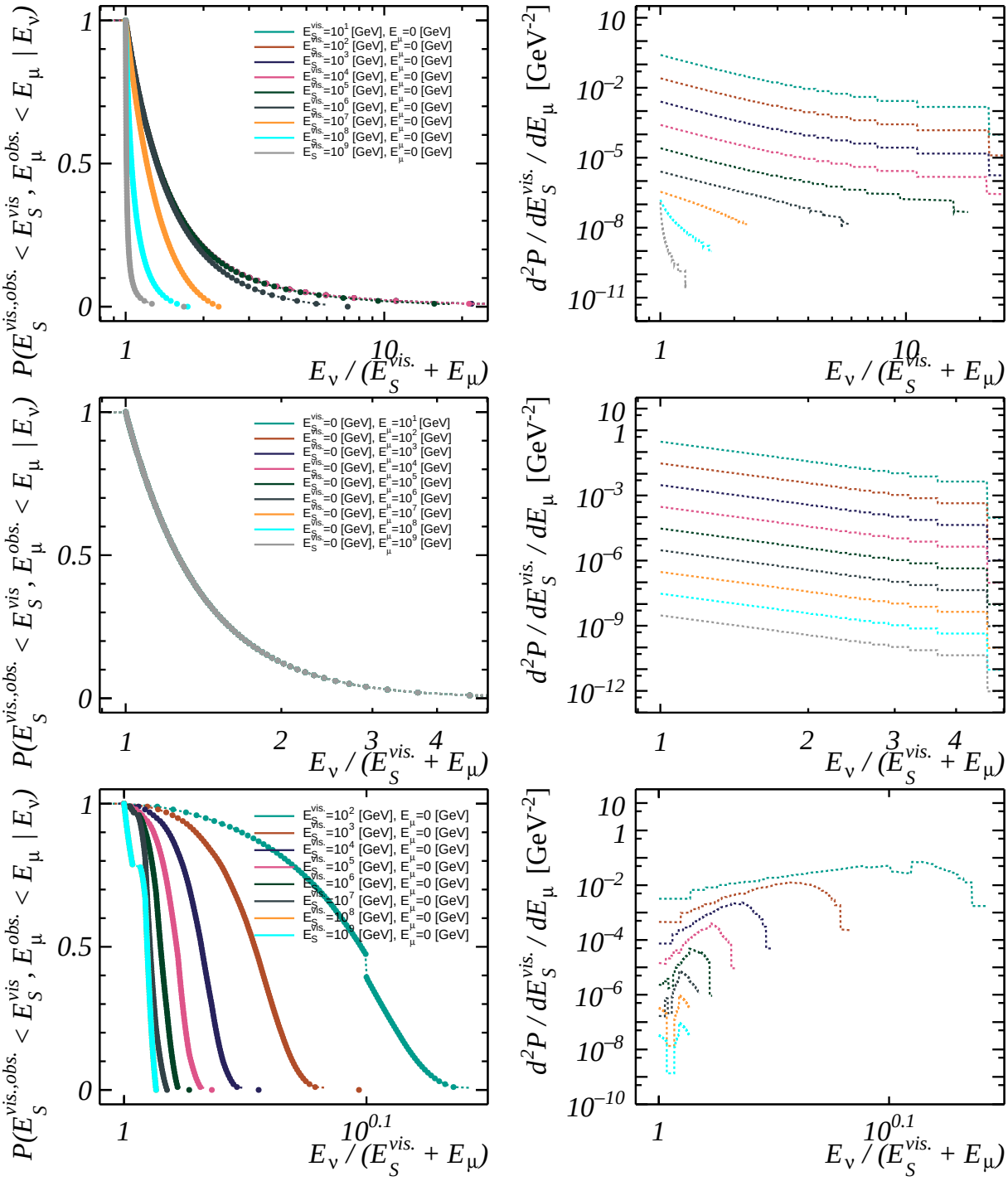


Figure 11.6: *Same as figures 11.5, but for the channels* $\bar{\nu}_e + e^- \to \bar{\nu}_e + e^-$ (***top*** *figures),* $\bar{\nu}_e + e^- \to \bar{\nu}_\mu + \mu^-$ (***middle***) *and* $\bar{\nu}_e + e^- \to$ *hadrons* (***bottom*** *figures).*

The relative spread of the observed energies typically decreases with increasing neutrino energy. This follows from the smaller spread in the distributions of y_b and $F_S^{\text{vis.}}$ at higher energies. For some points, the event signature probability is not a continuous function. These discontinuities are caused by the fact that probability density function used to describe hadronic showers ($\frac{\partial P}{\partial F_S^{\text{vis.}}}(F_S^{\text{vis.}}|E_S)$) is obtained from simulations, and inherently contains statistical fluctuations.

11.1.7 Event Signature Flux

The event signature probability tables are used to calculate the event signature flux for a given hypothesis, event signature direction, position and time in three steps:

First, the neutrino interaction flux (e.g. equation 11.5) of each of the interaction channels is calculated at $E_\nu \in \{10^{0.01}, 10^{0.02}, \cdots, 10^{10}\}$ GeV. A quadratic polynomial is used, taking $\log\left(\frac{d^6\Phi_\nu^{int.}}{dE_\nu d^2\Omega d^3\vec{x}}\right)$ and $\log(E_\nu)$ as abscissae, to evaluate the neutrino interaction flux at any E_ν. For neutral current interactions, the interaction fluxes of the three neutrino flavours are added.

The second step is to combine the cumulative event signature probability (stored in the event signature probability tables) with the neutrino interaction flux to calculate the cumulative event signature flux

$$\frac{d\Phi^{\text{cum.}}}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) \equiv \int_0^{E_S^{\text{vis.}}} \int_0^{E_\mu} \frac{d^7\Phi}{dE_S^{vis} dE_\mu d^2\Omega d^3\vec{x}}(E_S^{\text{vis., obs.}}, E_\mu^{\text{obs.}}, \hat{d}_\nu, \vec{x}_\nu|\mathcal{H}) dE_S^{\text{vis., obs.}} dE_\mu^{\text{obs.}}, \tag{11.16}$$

$$= \int_0^\infty P_{\text{evt. sig.}}^{\text{cum.}}(E_S^{\text{vis.}}, E_\mu|E_\nu) \cdot \frac{d^6\Phi_\nu^{int.}}{dE_\nu d^2\Omega d^3\vec{x}}(E_\nu, \hat{d}_\nu, \vec{x}_\nu|\mathcal{H}) dE_\nu. \tag{11.17}$$

The integrand of this integral is evaluated at all covered E_ν abscissae of both the event signature probability and neutrino flux. Linear interpolation between the logarithm of the integrand and $\log(E_\nu)$ is used to evaluate the integral.

In the final step, the cumulative event signature flux of the nine different interaction channels are added. Since the same abscissae in $E_S^{\text{vis.}}$ and E_μ are used for all channels, no interpolation is needed for this addition.

The result of these steps is a two-dimensional function of $\log\left(\frac{d\Phi^{\text{cum.}}}{d\boldsymbol{x}_s}\right)$ at $\log(E_S^{\text{vis.}}/\text{ GeV})$ and $\log(E_\mu/\text{ GeV}) \in \{-\inf, 0, 0.1, 0.2, \cdots, 10\}$. A two-dimensional quadratic polynomial is used to interpolate between the evaluated points. The event signature flux at any $\boldsymbol{x}_s$ is calculated by taking the derivative of this function (with the appropriate Jacobian).

In figure 11.7, the expected event signature fluxes of the $\mathcal{H}_0$ and the benchmark $\mathcal{H}_1$ hypotheses are given for an arbitrary event signature direction and position ($\theta_s = 45\circ$, $\phi_s = 0°$, $z = 1$[m]).

Muons produced in the seabed through charged-current $\nu_\mu + \bar{\nu}_\mu$ and $\bar{\nu}_e + e^- \to \bar{\nu}_\mu + \mu^-$ interactions can propagate through the rock and produce a detectable signature. The corresponding event signature is characterised by $E_S^{vis} = 0$ (since light emitted by the shower produced at the interaction vertex is not detected), the point $\vec{x}_s$ where the muon

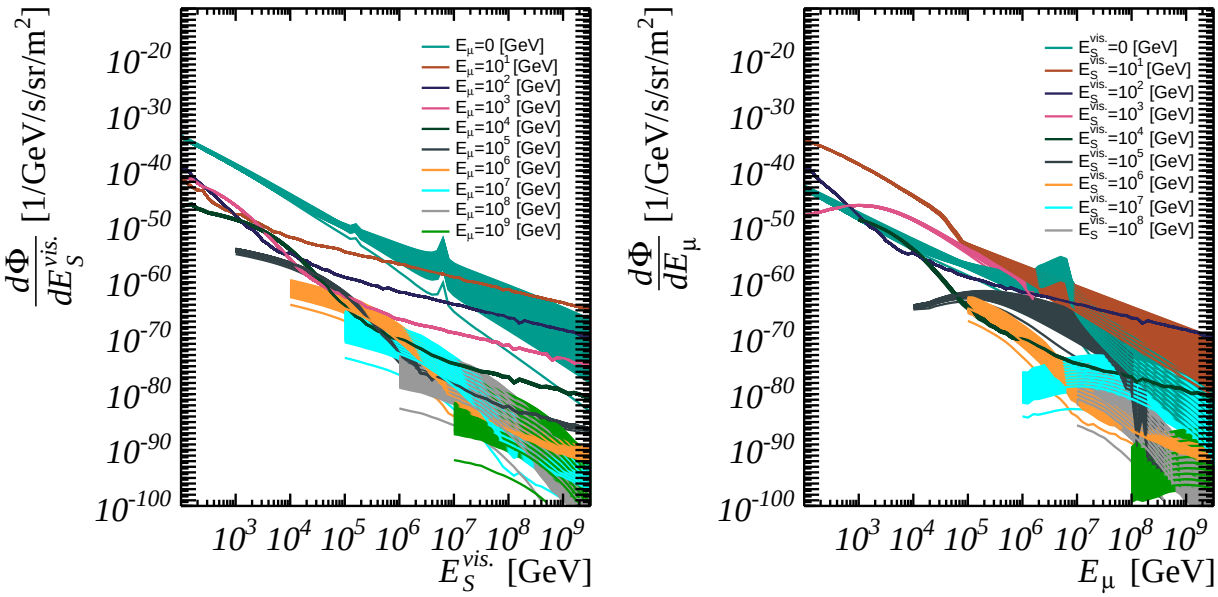


Figure 11.7: *The expected event signature flux as a function of the visible shower energy (**left**) and muon energy (**right**) for an arbitrary event signature direction and position ($\theta_s = 45°$, $\phi_s = 0°$, $z = 1[m]$). Several cross-sections of the (two-dimensional) flux are given, indicated by the different line colours. Each line corresponds to one hypothesis; the lowest line to the background-only $\mathcal{H}_0$ hypothesis, the other lines to the various benchmark $\mathcal{H}_1$ hypotheses (see section 11.1.1).*

leaves the rock and the energy E_μ of the muon at this point. The contribution of these muons to the three ingredients in the test statistic integral can be calculated as follows:

$$f(z_s = 0, E_S^{vis} = 0, E_\mu|\mathcal{H}) = \int_{-\infty}^{0}\int_{0}^{\infty} f(z_s < 0, E_S^{vis}, E_{\mu,0}(z_s)|\mathcal{H}) dE_S^{vis} dz_s, \tag{11.18}$$

where $E_{\mu,0}(z_s)$ is the initial energy of a muon produced at a depth z_s leaving the seabed with an energy E_μ:

$$E_{\mu,0}(z_s) = \left(\frac{A}{B} + E_\mu\right) \cdot e^{B \cdot |z_s|/\cos(\theta_s)} - \frac{A}{B}, \tag{11.19}$$

which follows from equation 3.15.

11.2 Detection Efficiency

The detection efficiency $P^{det.}(\boldsymbol{x}_s)$ is defined as the probability that an event signature with parameters $\boldsymbol{x}_s$ gives a detected (i.e. triggered) event. An estimate $\tilde{P}^{det.}(\boldsymbol{x}_s)$ of the detection efficiency has been obtained by the fraction of the number of triggered events k out of N simulated events, randomly generated with a signature parameters $\boldsymbol{x}_s$:

$$\tilde{P}^{det}(\boldsymbol{x}_s) \equiv \frac{k(\boldsymbol{x}_s)}{N(\boldsymbol{x}_s)}. \tag{11.20}$$

The detection efficiency may vary over time, as the DOMs can move under the influence of sea currents and the PMTs loose efficiency over time (some may even stop working completely). In this section, a method will be discussed to obtain $\tilde{P}^{det}(\boldsymbol{x}_s)$ for any KM3NeT detector geometry, with any PMT characteristics.

11.2.1 Coordinates

In this work, this method is applied to a single KM3NeT/ARCA building block with nominal PMT characteristics and default geometry (see section 3.5).

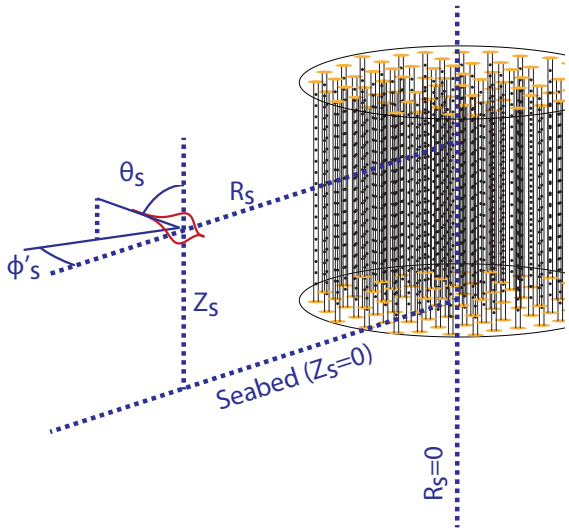


Figure 11.8: *An overview of the cylindrical coordinates z_s, R_s, θ_s and ϕ'_s used in the parameterisation of the detection efficiency.*

The instrumented volume of a KM3NeT/ARCA building block is approximately cylindrical, motivating the use of cylindrical coordinates. Using the point $\vec{x}_0 = [x_0, y_0, z_0]$ on the seabed straight below the centre of the smallest cylinder enclosing the instrumented volume, the event signature position $\vec{x}_s = [x_s, y_s, z_s]$ and direction $\hat{d}_s$ are given by the following four coordinates (see figure 11.8):

$$\begin{aligned} z_s &\equiv z_s - z_0, \\ R_s &\equiv \sqrt{(x_s - x_0)^2 + (y_s - y_0)^2}, \\ \theta_s &\equiv \arccos\left(\hat{d}_{s,z}\right), \\ \phi'_s &\equiv \arccos\left(\frac{1}{R_s \cdot \sin(\theta_s)} \cdot \begin{bmatrix} x_s - x_0 \\ y_s - y_0 \end{bmatrix} \cdot \begin{bmatrix} \hat{d}_{s,x} \\ \hat{d}_{s,y} \end{bmatrix}\right), \end{aligned} \tag{11.21}$$

11.2.2 Event Simulation

For each abscissa in $\boldsymbol{x}_s$ at which $\tilde{P}^{det}$ is evaluated, a set of N event signatures are simulated. For each simulated event, a random interaction position $\vec{x}_s$ is picked from the horizontal circle given by z_s and R_s. Light emission from the event signature (including stochastic energy losses along the muon track), the propagation of light through the seawater and

the detector response is simulated using slightly adapted versions of the JSirene and JTE simulation packages (see sections 4.4.1 and 4.5).

11.2.3 Number of Simulated Events

For a given k and N, the probability density of the true detection efficiency follows a beta distribution. The mean of this distribution is used as an estimate of the true detection efficiency (e.g. equation 11.20), the variance of this beta distribution:

$$Var\left[P^{det}\right] = \frac{(k+1)\cdot(N-k+1)}{(N+2)^2\cdot(N+3)}, \tag{11.22}$$

gives a measure of the accuracy of $\tilde{P}^{det}(\boldsymbol{x}_s)$. Events are simulated until $Var\left[P^{det}\right]$ is less than $4\cdot10^{-4}$. In order to reach this accuracy, the number of simulated events ranges from $N = 44$ in case $k = 0$ ($\tilde{P}^{det} = 0$) or $k = N$ ($\tilde{P}^{det} = 1$) to $N = 622$ if $k = N/2$ ($\tilde{P}^{det} = 0.5$).

The number of emitted photons from an event signature increases with E_S^{vis} and E_μ. Suppose the detection efficiency is estimated to be one ($k = N = 44$) at a certain $\boldsymbol{x}_s$, then the detection efficiency at all $\boldsymbol{x}_s$ with the same event signature position and direction but with higher $E_S^{vis.}$ and/or E_μ are be assumed to be one as well. This assumption is used in the following way: For each position and direction, $\tilde{P}^{det}(\boldsymbol{x}_s)$ is evaluated in order of increasing $E_S^{vis.}$ and E_μ. In case $\tilde{P}^{det}(\boldsymbol{x}_s) = 1$ for two subsequent abscissae in either $E_S^{vis.}$ or E_μ, the detection efficiency of all abscissae with higher $E_S^{vis.}$ and E_μ are assumed to be one as well. This approach leads to a significant reduction in the required time to obtain the detection efficiency table, since high-energy events are generally the most CPU-intensive to simulate.

In figure 11.9, the number of simulated events N and the corresponding $\tilde{P}^{det}(\boldsymbol{x}_s)$ are given as a function of $E_S^{\text{vis.}}$ and E_μ for one (arbitrary) point in z_s, R_s, θ_s and ϕ'_s. The principles used to minimise the number of simulated events is reflected in the fact that N is maximum if $\tilde{P}^{det}(\boldsymbol{x}_s) \approx 0.5$, decreases to 44 at higher energies ($\tilde{P}^{det}(\boldsymbol{x}_s) = 1$), and drops to zero for abscissae at even higher energies.

11.2.4 Detection Efficiency Table

The detection efficiency table contains (rather than just the estimated $\tilde{P}^{det}$), the number of simulated events N and number of triggered events k for a large set of abscissae in z_s, R_s, θ_s, ϕ'_s, E_S^{vis} and E_μ. Storing both N and k gives a measure of the accuracy of $\tilde{P}^{det}(\boldsymbol{x}_s)$ for later use. A second advantage of storing both N and k is that, if needed, the estimated detection efficiency $\tilde{P}^{det}(\boldsymbol{x}_s)$ can be refined at a later stage, including the results already evaluated.

The used grid of abscissae is summarised in table 11.2. In total, 161 317 800 points are used.

In figures 11.10, the estimate detection efficiency $\tilde{P}^{det}$ is given for the set of sampled event signature positions, directions and energies. The detection efficiency follows a rather intuitive function; the volume in which shower-like events are detected increases with the energy (either E_S^{vis} or E_μ, see top and middle row). For muon-like events, the detection

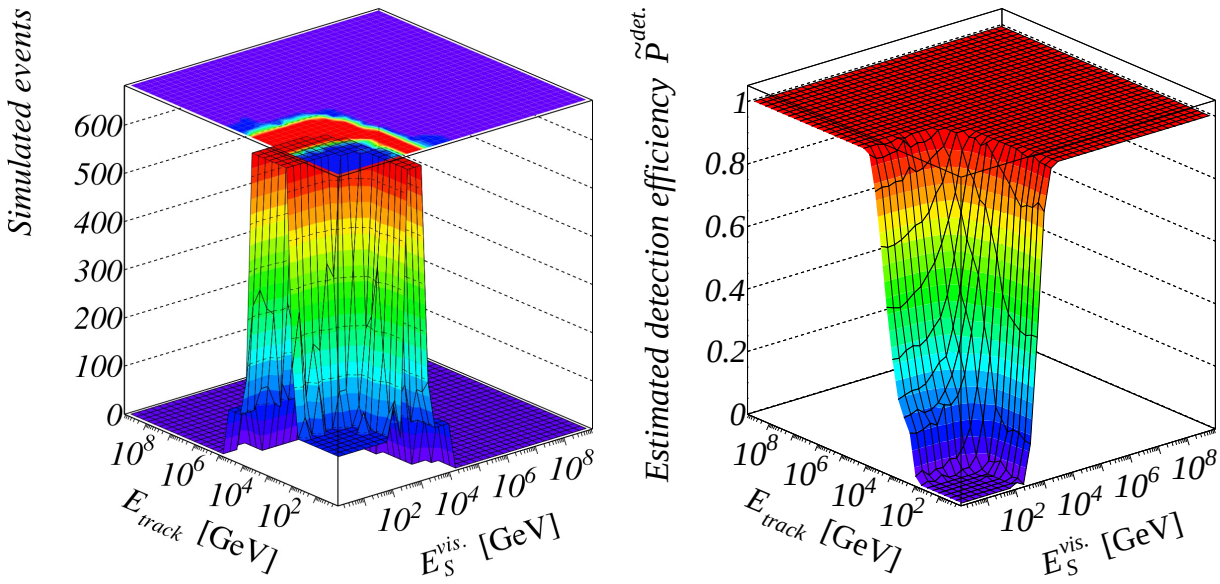


Figure 11.9: *The number of simulated events (**left**) and the estimated detection efficiency (**right**) as a function of $E_S^{vis.}$ and E_μ for an event signature at $R_s = 150$m, $z_s = 275$m, $\theta_s = 28.65°$ and $\phi'_s = 74.48°$.*

	Unit	Grid	No. abs.
Position $\vec{x}_s$	[m]	$R_s \in \{0, 50, 100, \cdots, 1200\}$ $z_s \in \{0, 25, \cdots, 100, 150, \cdots, 1400\}$	$25 \cdot 31 = 775$
Direction $\hat{d}_s$	[°]	$\theta_s \in \frac{180}{N_\theta - 1} \cdot \{0, 1, \cdots, N_\theta - 1\}$ $N_\theta = 1 + \lfloor \frac{1}{2} + \frac{180}{14 \cdot \Delta\hat{d}} \rfloor$, $\Delta\hat{d} = 10° \Rightarrow N_\theta = 14$ $\phi'_s \in \frac{180}{N_{\phi'} - 1} \cdot \{0, 1, \cdots, N_{\phi'} - 1\}$ $N_{\phi'} = 1 + \lfloor \frac{1}{2} + \frac{\pi \cdot \sin(\theta_s)}{2 \cdot \sqrt{1 - \cos(\Delta\hat{d})}} \rfloor$	118
Energy $E_S^{vis.}$, E_μ	[GeV]	$E_S^{vis.} \in \{0, 10^0, 10^{0.25}, 10^{0.5}, 10^{0.75}, \cdots, 10^{9.5}, 10^{9.75}, 10^{10}\}$ $E_\mu \in \{0, 10^0, 10^{0.25}, 10^{0.5}, 10^{0.75}, \cdots, 10^{9.5}, 10^{9.75}, 10^{10}\}$	$42 \cdot 42 = 1764$
Total			161 317 800

Table 11.2: *The grid of abscissae used for the detection efficiency table. The floor function of x is denoted by $\lfloor x \rfloor$.*

efficiency is maximal in case the track passes through the (centre of the) instrumented volume (bottom row).

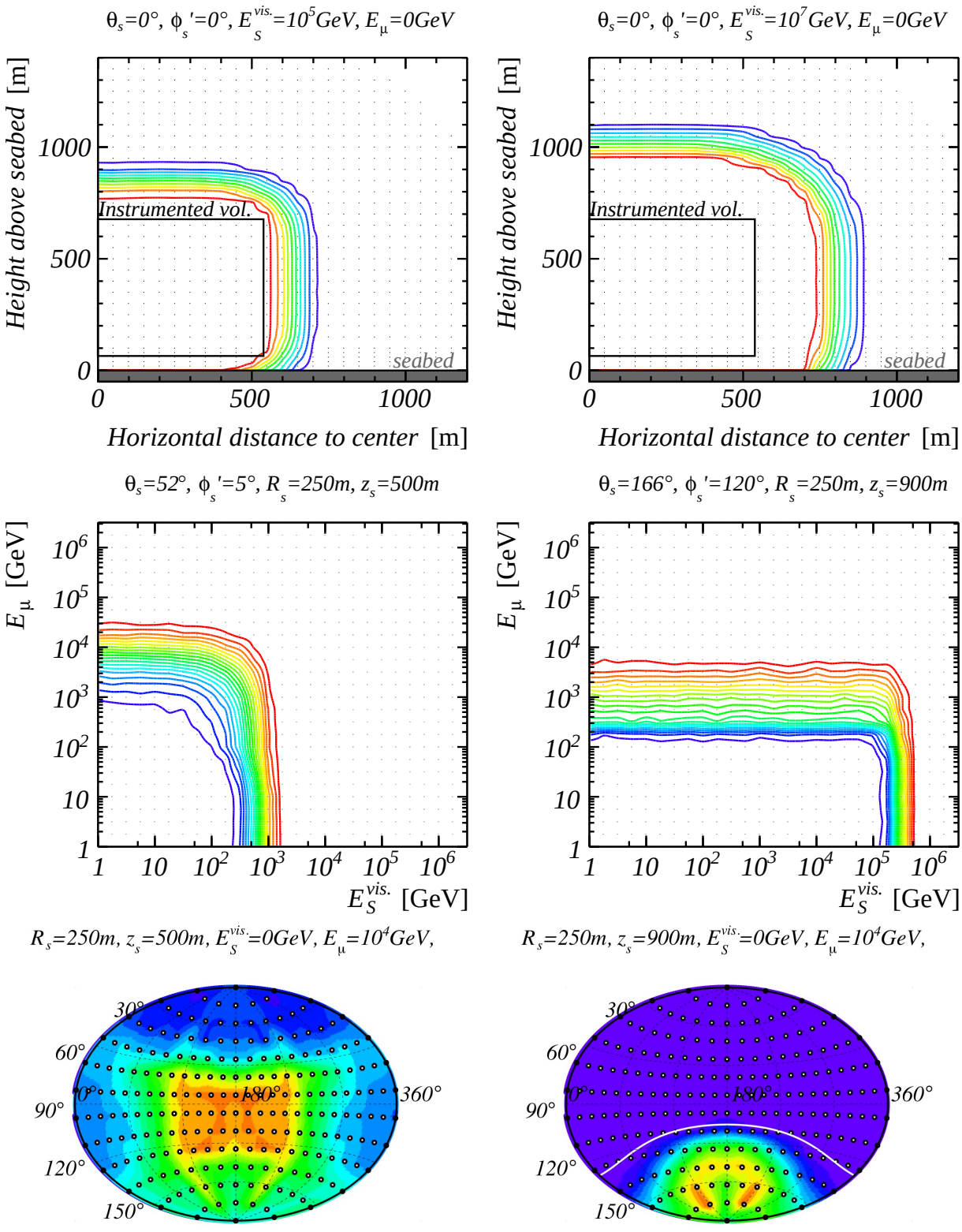


Figure 11.10: *The detection efficiency of one KM3NeT/ARCA building block with nominal PMT efficiencies and positions. The detection efficiency is estimated from simulations at the abscissae indicated by the title of and the dots in each figure, linear interpolation is used in between these abscissae. See text for the coordinates used. The cylinder enclosing the instrumented volume is indicated by the black rectangle in the upper two panels and the white line in the direction landscape in the lowest right corner.*

11.2.5 Interpolation

A linear interpolation between the ordinate values at the abscissae has been used to calculate the detection efficiency at any $\boldsymbol{x}_s$, using the JPP interpolation routines [193, 217]. A linear interpolation is the least sensitive to the statistical fluctuations, inherited from the procedure to estimate the detection efficiency.

Chapter 12

Event Likelihood

The event likelihood $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)$ is defined as the likelihood to observe a certain detected event (with observed L0s $\mathcal{E}$), caused by an event signature with parameters $\boldsymbol{x}_s$. As discussed in section 10.1.3, any function can be used. However, the event likelihood function that gives the most powerful hypothesis test, includes all available information of each detected event $\mathcal{E}$, and approximates the true probability density function to observe $\mathcal{E}$ for a given $\boldsymbol{x}_s$ as good as possible. The term 'likelihood' is used to underline the fact that $\boldsymbol{x}_s$ is varied, rather than $\mathcal{E}$. The used event likelihood function is normalised: $\int \mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)d\mathcal{E} = 1$, and is thus a probability density function in case $\mathcal{E}$ would be varied.

12.1 Definition

The event likelihood used in this work is based on the following information:

- The fact whether a PMT is 'hit' or not. In this, a PMT is 'hit' in case it recorded at least one L0 in a certain time window $[t_{min}, t_{max}]$. The time window used will be discussed in section 12.1.1.
- In case a PMT i is hit; the time t_i of the first detected L0 in the time window $[t_{min}, t_{max}]$.

Differentiating between hit and non-hit (n.h.) PMTs, the used event likelihood function is given by:

$$\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s) \equiv \mathcal{L}^{\text{hit PMTs}}(\mathcal{E}|\boldsymbol{x}_s) \times \mathcal{L}^{\text{n.h. PMTs}}(\mathcal{E}|\boldsymbol{x}_s), \tag{12.1}$$

with:

$$\mathcal{L}^{\text{hit PMTs}}(\mathcal{E}|\boldsymbol{x}_s) \equiv \prod_{\text{hit PMTs}} \left[\mathcal{L}_i^{\text{hit}}(\boldsymbol{x}_s)\right], \tag{12.2}$$

and

$$\mathcal{L}^{\text{n.h. PMTs}}(\mathcal{E}|\boldsymbol{x}_s) \equiv \prod_{\text{non-hit PMTs}} \left[\mathcal{L}_i^{\text{n.h.}}(\boldsymbol{x}_s)\right], \tag{12.3}$$

where $\mathcal{L}_i^{\text{n.h.}}(\boldsymbol{x}_s)$ is the likelihood that a PMT is not hit

$$\mathcal{L}_i^{\text{n.h.}}(\boldsymbol{x}_s) \equiv e^{-n_i(t_{min.},t_{max}|\boldsymbol{x}_s)}, \tag{12.4}$$

and $\mathcal{L}_i^{\text{hit}}(\boldsymbol{x}_s)$ the likelihood that the first L0 on a hit PMT is detected (in a small time window of $\Delta t = 1$ ns centred) at time t_i:

$$\mathcal{L}_i^{\text{hit}}(\boldsymbol{x}_s) \equiv e^{-n_i(t_{min.},t_i-\Delta t/2|\boldsymbol{x}_s)} \cdot \left(1 - e^{-n_i(t_i-\Delta t/2,t_i+\Delta t/2|\boldsymbol{x}_s)}\right). \tag{12.5}$$

Both probabilities ($\mathcal{L}_i^{\text{n.h.}}(\boldsymbol{x}_s)$ and $\mathcal{L}_i^{\text{hit}}(\boldsymbol{x}_s)$) are derived from Poisson distributions, depending on $\boldsymbol{x}_s$ via the expected number of photoelectrons on a PMT i between time t_a and t_b:

$$n_i(t_a, t_b|\boldsymbol{x}_s) \equiv \int_{t_a}^{t_b} \dot{n}_i(t|\boldsymbol{x}_s)dt, \tag{12.6}$$

Where the time-dependent expected number of photoelectrons $\dot{n}_i(t|\boldsymbol{x}_s)$ is given by the sum over photoelectrons caused by ^{40}K decays, as well as those induced by the shower and muon from the event signature:

$$\dot{n}_i(t|\boldsymbol{x}_s) = \dot{n}_{^{40}K} + \dot{n}_{S,i}(t|\boldsymbol{x}_s) + \dot{n}_{\mu,i}(t|\boldsymbol{x}_s), \tag{12.7}$$

with

$$\dot{n}_{^{40}K} \equiv R_{^{40}K}[\text{Hz}] \cdot 10^{-9}[\text{s / ns}], \tag{12.8}$$

the (time-dependent and PMT-specific) expected rate of L0s from ^{40}K decays (about 6.7 kHz per PMT, see paragraph 6.2). The time-dependent expected number of photoelectrons on PMT i from the shower $\dot{n}_{S,i}(t|\boldsymbol{x}_s)$ and the muon $\dot{n}_{\mu,i}(t|\boldsymbol{x}_s)$ component of the event signature $\boldsymbol{x}_s$ are evaluated using (multidimensional) interpolation of tabulated values. These tables are used in the simulations (section 4.4.1) as well as in the the track reconstruction algorithm (section 5.3). These functions are parameterised as a function of the time difference $t - t_s$ with respect to the event signature time t_s, four parameters describing the position and orientation of the PMT, and the two event signature energies E_S^{vis} and E_μ. The relations between $\dot{n}_{S,i}(t|\boldsymbol{x}_s)$ and $\dot{n}_{\mu,i}(t|\boldsymbol{x}_s)$ and these parameters are discussed in appendix D.

Simplifications Using the approximations

$$n_i(t_i - \Delta t/2, t_i + \Delta t/2|\boldsymbol{x}_s) \approx \dot{n}_i(t) \cdot \Delta t, \tag{12.9}$$

and

$$n_i(t_{min.}, t_{max.}|\boldsymbol{x}_s) \approx n_i(-\infty, \infty|\boldsymbol{x}_s), \tag{12.10}$$

the likelihoods of the hit- and non-hit PMTs is approximately given by:

$$\mathcal{L}_i^{\text{n.h.}}(\boldsymbol{x}_s) \approx e^{-N_i(\boldsymbol{x}_s)}, \tag{12.11}$$

$$\mathcal{L}_i^{\text{hit}}(\boldsymbol{x}_s) \approx e^{-n_i(t_i|\boldsymbol{x}_s)+\frac{1}{2}\cdot\dot{n}_i(t_i|\boldsymbol{x}_s)} \cdot \left(1 - e^{-\dot{n}_i(t_i|\boldsymbol{x}_s)}\right). \tag{12.12}$$

where

$$N_i(\boldsymbol{x}_s) \equiv \int_{-\infty}^{\infty} \dot{n}_i(t|\boldsymbol{x}_s)dt, \tag{12.13}$$

$$n_i(t|\boldsymbol{x}_s) \equiv \int_{-\infty}^{t} \dot{n}_i(t|\boldsymbol{x}_s)dt, \tag{12.14}$$

The routine used to evaluate these quantities (see appendix D) allows for a significant reduction in required computing time to evaluate this approximate form of the event likelihood. In the remainder of this thesis, only this approximate form of the event likelihood will be used.

Log-likelihood The event likelihood $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)$ gives the likelihood that (of all possible events) one particular event $\mathcal{E}$ is observed. Since the number of possible events is very large, the event likelihood is typically very small. Therefore, it is often more convenient to use the natural logarithm of the likelihood:

$$\begin{aligned}
l(\mathcal{E}|\boldsymbol{x}_s) &\equiv \log\left(\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)\right), \\
&\equiv l^{\text{hit PMTs}}(\mathcal{E}|\boldsymbol{x}_s) + l^{\text{n.h. PMTs}}(\mathcal{E}|\boldsymbol{x}_s), \\
l^{\text{n.h. PMTs}}(\mathcal{E}|\boldsymbol{x}_s) &\equiv \sum_{\text{non-hit PMTs}} \left[l_i^{\text{n.h.}}(\boldsymbol{x}_s)\right], \\
l_i^{\text{n.h.}}(\boldsymbol{x}_s) &\approx -N_i(\boldsymbol{x}_s), \\
l^{\text{hit PMTs}}(\mathcal{E}|\boldsymbol{x}_s) &\equiv \sum_{\text{hit PMTs}} \left[l_i^{\text{hit}}(\boldsymbol{x}_s)\right], \\
l_i^{\text{hit}}(\boldsymbol{x}_s) &\approx -n_i(t_i|\boldsymbol{x}_s) + \frac{1}{2}\cdot\dot{n}_i(t_i|\boldsymbol{x}_s) + \log\left(1 - e^{-\dot{n}_i(t_i|\boldsymbol{x}_s)}\right).
\end{aligned} \tag{12.15}$$

12.1.1 Time Window for First Hit Selection

A PMT is 'hit' in case it recorded at least one L0 in a time window $[t_{min}, t_{max}]$. The time window should be as small as possible to maximise the probability that the first level zero hit in $[t_{min}, t_{max}]$ originates from the neutrino interaction and not from a ^{40}K decay. On the other hand, the time window should be sufficiently large to include a large fraction of the signal hits.

In reconstruction algorithms, such as the ones discussed in chapter 5, the time window differs for every PMT, and depends on the expected arrival time of light from a shower or muon on a PMT at a given position, time and direction. The advantage of this choice is that a relatively small time window can be used. A downside of this approach is that the number of hit and non-hit PMTs depends on the signature parameters $\boldsymbol{x}_s$. Consequently, the event likelihood function $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)$ is not a continuous function, thus complicating the use of it in numerical procedures. Additionally, the event likelihood function becomes biased, in the sense that some $\boldsymbol{x}_s$ have a higher likelihood only because a (background) hit is or is not included in the set of first hits.

In this work, the time window used to select the first L0s is the same for all PMTs. This time window is defined with respect to the time of the first detected L1 ($t_{\text{1st trig.}}$) of the triggered event (called the 'first triggered hit'), as identified by the clustering algorithms used in the event trigger (see section 3.7.4).

The distribution of the difference between the measured L0s hit times and $t_{\text{1st trig.}}$ is shown in figure 12.1. The arrival time distributions of L0s, from shower-like as well as muon-like events, peak at the time of the first triggered hit. The tails towards later times result mainly from the (scattered) paths of photons. Based on these distributions, a time window of $[t_{min}, t_{max}] = [t_{\text{1st trig.}} - 100\text{ns}, t_{\text{1st trig.}} + 1000\text{ns}]$ has been chosen. The distribution of the corresponding number of selected (i.e. first) L0s per event is given in figure 12.1 (right). The expected number of selected L0s from ^{40}K decays in this time window is $5 \cdot 10^3 \cdot 1100 \cdot 10^{-9} \cdot 31 \cdot 18 \cdot 115 \approx 353$, or about one L0 per 182 PMTs.

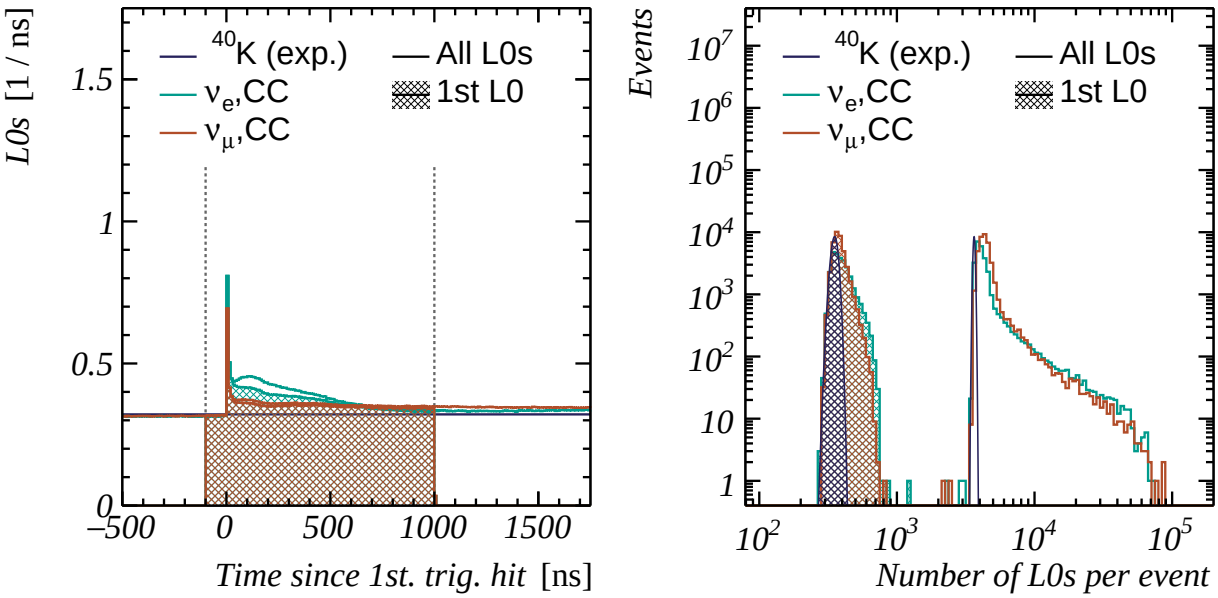


Figure 12.1: ***Left:*** *The number of L0s (in one KM3NeT/ARCA building block) as a function of the time of the hits with respect to the time of the first recorded triggered hit. Shown are the distributions from a set of simulated shower-like (charged-current ν_e+$\bar{\nu}_e$ interactions) and track-like (charged-current ν_μ+$\bar{\nu}_\mu$ interactions) events, as well as the expected L0 rate from ^{40}K decays (horizontal line). The time window used to select the first hits on each PMT is indicated by the dashed lines. In the* ***right*** *figure, the distributions of the total number of L0s per triggered event and the number of selected L0s per triggered event are given. The distributions referred to as ^{40}K (exp.) corresponds to the expected number of selected L0s from ^{40}K only.*

12.2 Event Likelihood 'Landscapes'

In figures 12.2, 12.3 and 12.4, an impression of some event (log-)likelihood 'landscapes' are given for a simulated shower-like event ($E_\mu \approx 0$, $E_S^{vis} \gg 0$), a track-like event ($E_\mu \gg 0$, $E_S^{vis} \approx 0$) and a hybrid event ($E_\mu \gg 0$, $E_S^{vis} \gg 0$). The relevant characteristics of these events are summarised in table 12.1. These landscapes are obtained by varying one or two parameters of $\boldsymbol{x}_s$, while keeping the other parameters fixed to their true values. The event signature position $\vec{x}_s$ and time t_s are parameterised with respect to the true values

$\vec{x}_s^{\,\text{true}}$ and t_s^{true} using the following quantities:

$$\begin{aligned}
\Delta t &\equiv t_s - t_s^{\text{true}}, \\
\Delta z &\equiv (\vec{x}_s - \vec{x}_s^{\,\text{true}}) \cdot \hat{d}_s^{\text{true}}, \\
\Delta x &\equiv \left(\vec{x}_s - \vec{x}_s^{\,\text{true}} - \Delta z \cdot \hat{d}_s^{\text{true}}\right) \cdot \hat{p}, \\
\Delta y &\equiv \sqrt{(\vec{x}_s - \vec{x}_s^{\,\text{true}})^2 - \Delta z^2 - \Delta x^2},
\end{aligned} \tag{12.16}$$

with $\hat{p}$ corresponding to an arbitrary direction perpendicular to the true signature direction $\hat{d}_s$. The quantity Δz is called the longitudinal position, the parameters Δx and Δy give the perpendicular position of the event signature. With these, the sampled event signature position $\vec{x}_s$ is given by:

$$\vec{x}_s = \vec{x}_s^{\,\text{true}} + \Delta z \cdot \hat{d}_s + \Delta x \cdot \hat{p} + \Delta y \cdot \left(\hat{d}_s \times \hat{p}\right), \tag{12.17}$$

Several features of the event likelihood function are present in these landscapes:

1. The time profiles ($l(\mathcal{E}|\boldsymbol{x}_s)$ as a function of Δt) have numerous local maxima. The locations of these correspond (approximately) to the t_s at which either $\dot{n}_{S,i}(t|\boldsymbol{x}_s)$ or $\dot{n}_{\mu,i}(t|\boldsymbol{x}_s)$ is maximum for one (or more) of the selected L0s. The global maximum of the event likelihood landscape occurs at the value of $\boldsymbol{x}_s$ where all these local maxima overlap. This maximum is generally found close to the true event signature parameters $\boldsymbol{x}_s^{\text{true}}$.

2. For $\Delta t \gg 0$, the assumed t_s is much later than t_s^{true}. Therefore, the majority of the selected hits have times before t_s, i.e. before any light was emitted. The corresponding likelihood converges to:

$$l_{\text{late } t_s} = l^{\text{n.h. PMTs}}(\mathcal{E}|\boldsymbol{x}_s) + l^{\text{hit PMTs}}(\mathcal{E}|^{40}K), \tag{12.18}$$

   with:

$$l^{\text{hit PMTs}}(\mathcal{E}|^{40}K) \equiv \sum_{\text{hit PMTs}} \left[-R_{^{40}K} \cdot (t_i - t_{min}) + \frac{R_{^{40}K}}{2} + \log\left(1 - \exp\left(-R_{^{40}K}\right)\right)\right]. \tag{12.19}$$

3. For $\Delta t \ll 0$, the situation is comparable. In contrast to $\Delta t \gg 0$, all hit PMTs should have recorded a hit (much) earlier than they actually did. Consequently, the likelihood is significantly lower than $l_{\text{late } t_s}$:

$$l_{\text{early } t_s} = l_{\text{late } t_s} - \sum_{\text{hit PMTs}i} \left[N_{S,i} + N_{\mu,i}\right]. \tag{12.20}$$

4. The event likelihood varies with Δz only in case a significant fraction of the total number of emitted photons originates from the shower. For the muon-like event signature (figure 12.3), this is not the case, and the time profiles at different Δz are nearly identical, except for a shift by $\Delta t = \Delta z/c$, with c the speed of light.

5. The energy landscapes (the event likelihood as a function of E_S^{vis} and E_μ, given in the left panel of the middle rows) are characterised by a relatively smooth increase of the likelihood towards the true energies of the signature, followed by a sharply decreasing likelihood at higher energies. The latter is caused by non-hit PMTs, which should have recorded a hit at these higher energies.

6. For events with $E_S^{vis} \gg 0$ (figures 12.2 and 12.4), the event likelihood is relatively large for a broad range of track energies E_μ. For the track-like signature (figure 12.3), the event likelihood is relatively large for a broad range of E_S^{vis}. The difference is caused by the spatial distribution of the hits; the event likelihood is high as long as the true shower or muon energy is close to the true energy, thus 'explaining' the detected hits.

7. The perpendicular event likelihood ($l(\mathcal{E}|\boldsymbol{x}_s)$ as a function of Δx and Δy, given in the right panel of the middle rows) is characterised by a complex landscape, with several distinct local minima. The locations of these minima correspond to $\boldsymbol{x}_s$ with the muon track passing close to one or more non-hit PMTs which, for the given $\boldsymbol{x}_s$, should have been hit. These minima are also present (but less prominent) in the longitudinal landscape (l versus Δz), for example in the top right panels of figures 12.2 at $\Delta z \approx -40$ m and figures 12.4 at $\Delta z \approx 25$ m, as well as in the direction landscapes (l as a function of $\hat{d}_s$, bottom figures).

Quantity	Unit	Event 1 Fig. 12.2	Event 2 Fig. 12.3	Event 3 Fig. 12.4
Interaction		ν_e, CC	ν_μ, CC	ν_μ, CC
E_ν	[GeV]	$32.4 \cdot 10^3$	$27.4 \cdot 10^3$	$35.3 \cdot 10^3$
y_b		0.313	0.982	0.360
E_μ	[GeV]	0	$26.9 \cdot 10^3$	$12.7 \cdot 10^3$
E_S	[GeV]	$32.4 \cdot 10^3$	502.0	$22.6 \cdot 10^3$
F_S^{vis}		0.818	0.395	0.793
E_S^{vis}	[GeV]	$28.4 \cdot 10^3$	198.3	$17.9 \cdot 10^3$
R_s	[m]	392.1	255.9	455.0
z_s	[m]	199.5	296.8	562.3
θ_s	[◦]	44.8	135.9	139.0
ϕ_s	[◦]	142.0	6.1	108.0
ϕ'_s	[◦]	192.7	231.2	51.8
Number of hit PMTs		2352	710	2068
$l(\mathcal{E}\|E_S^{vis.} = E_\mu = 0)$		−30499.7	−9321.2	−26950.1
$l(\mathcal{E}\|\boldsymbol{x}_s^{\text{true}})$		−28240.6	−8941.4	−25201.6

Table 12.1: *An overview of characteristics of the simulated events used to produce figures 12.2 to 12.4. The used quantities are described in the text. The quantities R_s, z_s and ϕ'_s give the event signature position and direction with respect to the instrumented volume, as defined in equation 11.21.*

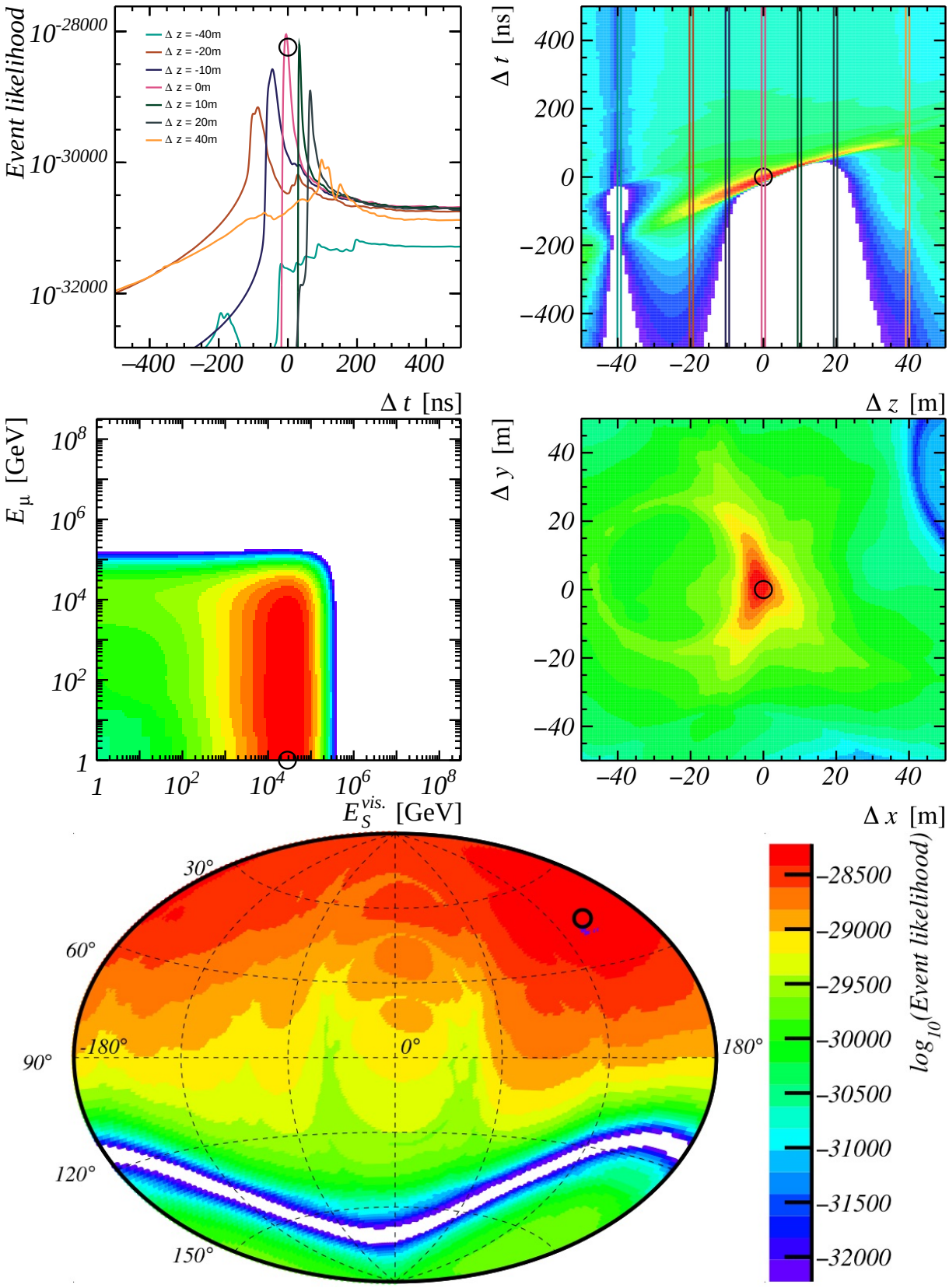


Figure 12.2: *The event likelihood 'landscapes' of a simulated shower-like event (event number one). All event signature parameters (see text for the definition) other than the ones indicated on the axes are kept fixed at the true signature parameters (see table 12.1). The true parameters are indicated by the black circles.*

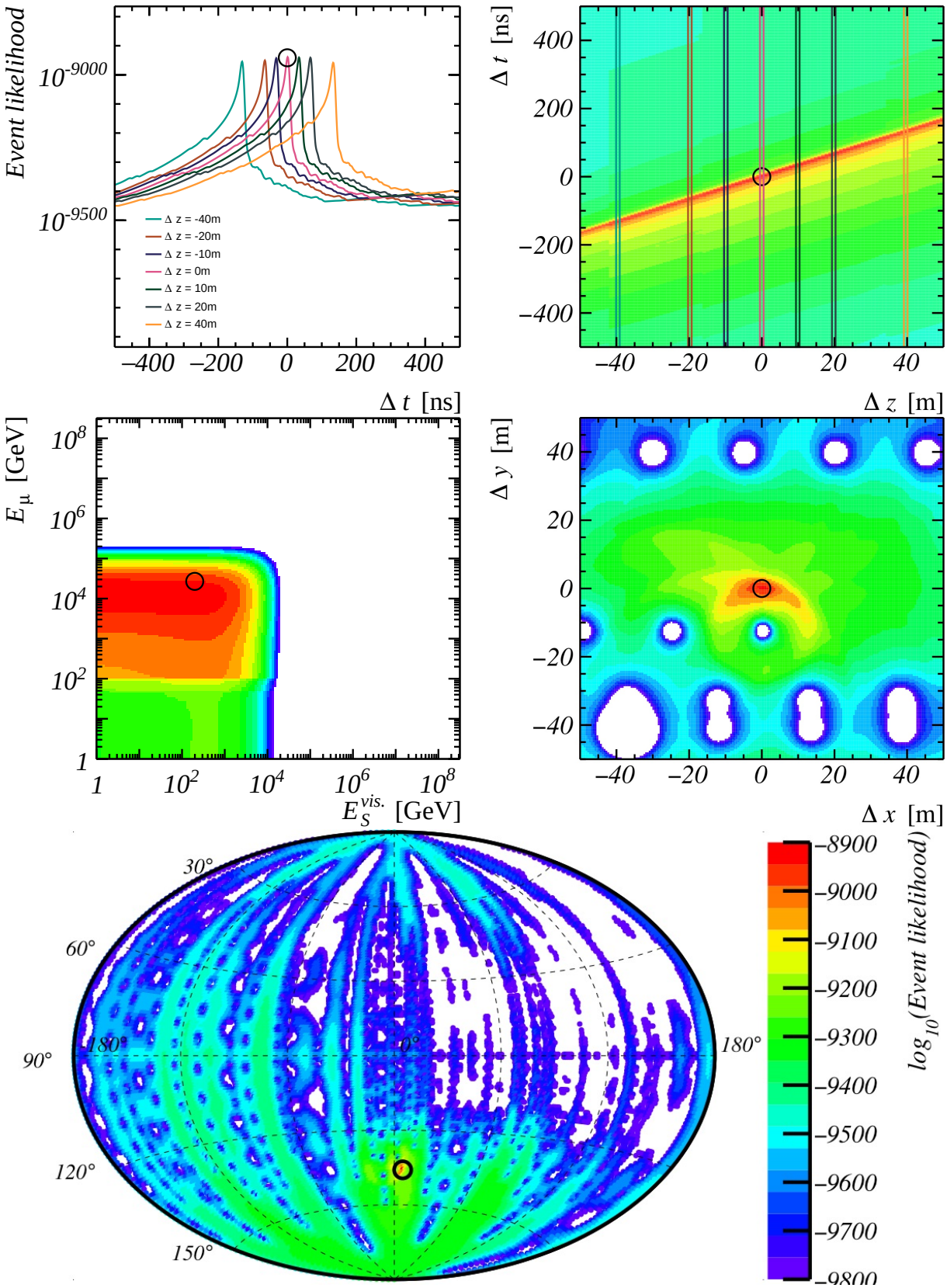


Figure 12.3: *The event likelihood 'landscapes' of a simulated track-like event (event number two). All event signature parameters (see text for the definition) other than the ones indicated on the axes are kept fixed at the true signature parameters (see table 12.1). The true parameters are indicated by the black circles.*

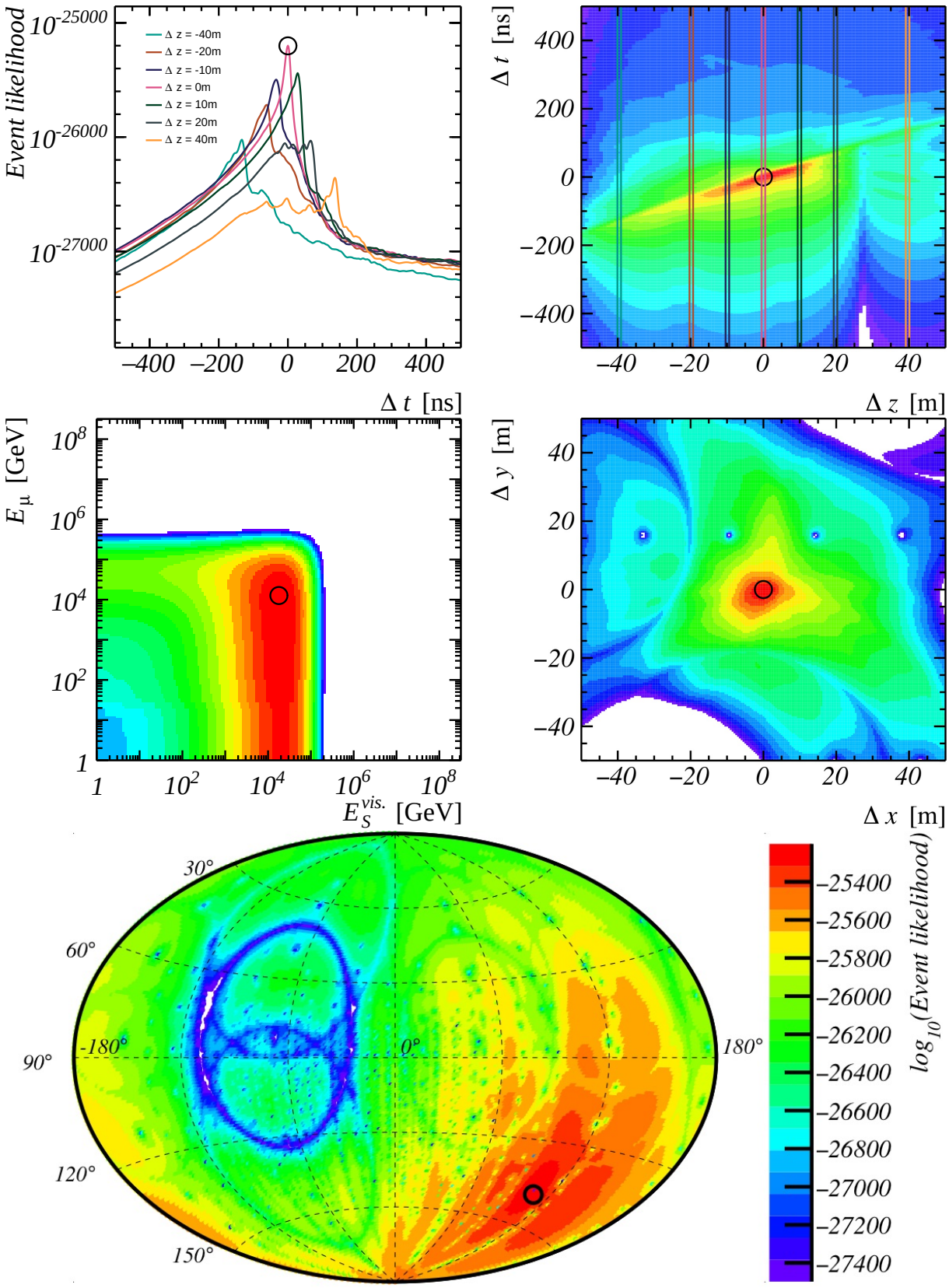


Figure 12.4: *The event likelihood 'landscapes' of a simulated shower-like event (event number three). All event signature parameters (see text for the definition) other than the ones indicated on the axes are kept fixed at the true signature parameters (see table 12.1). The true parameters are indicated by the black circles.*

12.3 Time-dependent Event Likelihood

The event likelihood varies only during a very short period around the time t_s^{max} at which $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)$ is maximum. The expected event signature flux and detection efficiency are constant in this short period. For earlier and later t_s, the event likelihood is constant in time. Because of this, it is sufficient to characterise the time-dependent event likelihood (figure 12.5) with four quantities:

- The time t_s^{max} at which $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)$ (or, equivalently, $l(\mathcal{E}|\boldsymbol{x}_s)$) is maximal.
- The log-likelihood $l_{\text{early } t_s}$ at very early t_s (equation 12.20).
- The log-likelihood $l_{\text{late } t_s}$ at very late t_s (equation 12.18).
- The logarithm of the time-integrated event likelihood, defined as

$$\log(\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)) \equiv \log\left(\int_{-\infty}^{\infty} \max\left[\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s) - \mathcal{L}_{\text{late } t_s}, 0\right] dt_s\right). \tag{12.21}$$

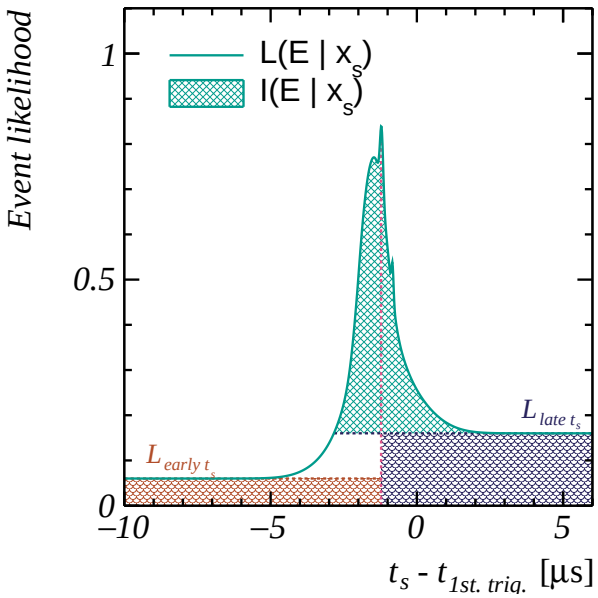


Figure 12.5: *An illustration of the quantities stored in the event likelihood tables ($l_{\text{early } t_s}$, $l_{\text{late } t_s}$, t_s^{max} and $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$). The shown event likelihood function $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ does not represent a real event likelihood function.*

With these four quantities, the time-integrated test statistic integrand between any two bounds $t_a \ll t_s^{max} \ll t_b$ can be approximated by:

$$\begin{aligned}
&\int_{t_a}^{t_b} \mathcal{L}(\mathcal{E}|\boldsymbol{x}_s) \cdot P^{det}(\boldsymbol{x}_s) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) dt_s \\
&\quad \approx \int_{t_a}^{t_s^{max}} \exp\left(l_{\text{early } t_s}\right) \cdot P^{det}(\boldsymbol{x}_s) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) dt_s \\
&\quad + \mathcal{I}(\mathcal{E}|\boldsymbol{x}_s) \cdot P^{det}(\boldsymbol{x}_s, t_s^{max}) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s, t_s^{max}|\mathcal{H}) \\
&\quad + \int_{t_s^{max}}^{t_b} \exp\left(l_{\text{late } t_s}\right) \cdot P^{det}(\boldsymbol{x}_s) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) dt_s
\end{aligned} \tag{12.22}$$

The calculation of $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ is the subject of the next section.

12.3.1 Numerical Integration

For a given $\vec{x}_s, \hat{d}_s, E_S^{vis}$ and E_μ, the event likelihood function is calculated at a set of interaction times $\{t_{s,1}, t_{s,2}, \cdots, t_{s,m}\}$. The numerical procedure used to calculate the time-integrated event likelihood $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ (e.g. equation 12.21) from these points is discussed here, whereas the method for determining the abscissae $\{t_{s,1}, t_{s,2}, \cdots, t_{s,m}\}$ is discussed in appendix G.

Interpolation At each abscissa $t_{s,i} \in \{t_{s,1}, t_{s,2}, \cdots, t_{s,m}\}$, the logarithm of the event likelihood $l_i \equiv l(\mathcal{E}|\vec{x}_s, \hat{d}_s, E_S^{vis}, E_\mu, t_{s,i})$ is calculated (using equation 12.15). Additionally, the derivatives $\dot{l}_i$ of l_i with respect to t_s are calculated:

$$\dot{l}(\mathcal{E}|\boldsymbol{x}_s) \equiv \frac{dl}{dt_s}(\mathcal{E}|\boldsymbol{x}_s), \tag{12.23}$$

$$= \frac{dl^{\text{hit PMTs}}}{dt_s}(\mathcal{E}|\boldsymbol{x}_s), \tag{12.24}$$

$$= \sum_{\text{hit PMTs k}} \left[\frac{dt}{dt_s} \cdot \frac{dl_k^{\text{hit}}}{dt}(\boldsymbol{x}_s) \right], \tag{12.25}$$

$$\approx -1 \cdot \sum_{\text{hit PMTs k}} \left[-\dot{n}_k(t_s) + \frac{1}{2} \cdot \ddot{n}_k(t_s) + \ddot{n}_k(t_s) \cdot \frac{1}{\exp\left(\dot{n}_k(t_s)\right) - 1} \right], \tag{12.26}$$

with $\dot{n}_k(t_s)$ and $\ddot{n}_k(t)$ given in equation 14.37. The factor -1 follows from the fact that $\dot{n}_k(t_s)$ is defined as a function of $t - t_s$.

Using the set of evaluated log-likelihoods $\{l_1, l_2, \cdots, l_m\}$ and derivatives $\{\dot{l}_1, \dot{l}_2, \cdots, \dot{l}_m\}$, the logarithm of the event likelihood function is approximated by a set of third-order polynomials, one for each pair of abscissae i, $i+1$:

$$\tilde{l}(t_s) \equiv \begin{cases} A_i \cdot \left(\frac{t_s - t_{s,i}}{t_{s,i+1} - t_{s,i}}\right)^3 + B_i \cdot \left(\frac{t_s - t_{s,i}}{t_{s,i+1} - t_{s,i}}\right)^2 \\ \qquad + \dot{l}_i \cdot \left(\frac{t_s - t_{s,i}}{t_{s,i+1} - t_{s,i}}\right) + l_i \quad \text{if } t_{s,i} \le t_s < t_{s,i+1} \end{cases} \tag{12.27}$$

with:

$$A_i \equiv \left(\dot{l}_i + \dot{l}_{i+1}\right) - 2 \cdot \left(l_{i+1} - l_i\right), \tag{12.28}$$

$$B_i \equiv 3 \cdot \left(\dot{l}_i + \dot{l}_{i+i}\right) - \left(2 \cdot l_{i+1} - l_i\right). \tag{12.29}$$

Gauss-Legendre integration The integral $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ is given by a sum over each pair of subsequent abscissae in t_s:

$$\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s) = \sum_{i=1}^{m-1} \left[\mathcal{I}_i(\mathcal{E}|\boldsymbol{x}_s)\right], \tag{12.30}$$

$$\mathcal{I}_i(\mathcal{E}|\boldsymbol{x}_s) \equiv \int_{t_{s,i}}^{t_{s,i+1}} \max\left[\exp\left(\tilde{l}(t_s)\right) - \exp\left(l_{\text{late } t_s}\right), 0\right] dt_s, \tag{12.31}$$

where $\tilde{l}(t_s)$ denotes the interpolated event log-likelihood given in equation 12.27.

The time-integrated event likelihood $\mathcal{I}_i$ in each interval i between $t_{s,i}$ and $t_{s,i+1}$ is evaluated using Gauss-Legendre integration[217]. In Gauss-Legendre integration, the integral over a function is calculated using a sum of the function values at a predefined set of k abscissae X_j in the range $[-1;1]$, each multiplied with a (precalculated) corresponding weight W_j:

$$\int_a^b f(x)dx \approx (b-a)\cdot\sum_{j=1}^{k}\left[W_j\cdot f(a+\frac{1+X_j}{2}\cdot(b-a))\right], \tag{12.32}$$

As discussed in section 12.1, the value of $\tilde{l}(t_s)$ is typically very small. In order to evaluate $\exp\left(\tilde{l}(t_s)\right)$ within the numerical accuracy of the computers used, the maximum of l_i and l_{i+1} is subtracted from $\tilde{l}(t_s)$ in the calculation of each $\mathcal{I}_i(\mathcal{E}|\boldsymbol{x}_s)$. This operation is compensated by adding the maximum of $(t_b-t_a)\cdot\exp(l_i)$ and $(t_b-t_a)\cdot\exp(l_{i+1})$ to the calculated $\mathcal{I}_i(\mathcal{E}|\boldsymbol{x}_s)$. The time-integrated event likelihood is then given by:

$$\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s) = \sum_{i=1}^{m-1}\left[\mathcal{I}_i(\mathcal{E}|\boldsymbol{x}_s)\right], \tag{12.33}$$

$$\begin{aligned}\mathcal{I}_i(\mathcal{E}|\boldsymbol{x}_s) \approx\ & (t_{s,i+1}-t_{s,i})\cdot\exp\left(\max\left[l_i,l_{i+1}\right]\right)\\ & \times\sum_{j=1}^{k}\left[W_j\cdot\max\left[\exp\left(\tilde{l}\left(t_{s,i}+\frac{1+X_j}{2}\cdot(t_{s,i+1}-t_{s,i})\right)-\max\left[l_i,l_{i+1}\right]\right)\right.\right.\\ & \left.\left.-\exp\left(l_{\text{late }t_s}-\max\left[l_i,l_{i+1}\right]\right),0\right]\right]\end{aligned} \tag{12.34}$$

Each $\mathcal{I}_i$ is evaluated using $k=100$ abscissae. A relatively high number is chosen since the computing time required to evaluate each Gauss-Legendre integral is insignificant when compared to all other steps, discussed in the remaining part of this chapter.

Accuracy The abscissae $\{t_{s,1},t_{s,2},\cdots,t_{s,m}\}$ used to sample the time-dependent event likelihood are chosen such that a minimum number of abscissae are needed for an accurate estimate of each $\mathcal{I}_i(\mathcal{E}|\boldsymbol{x}_s)$. This sampling consists of three steps. First, a (set of) start value(s) are determined. In the second step, abscissae are added around these points until all local maxima are characterised with at least 5 points in their vicinities. In the third step, abscissae are added in-between these points until the estimated accuracy of the evaluated $\mathcal{I}_i(\mathcal{E}|\boldsymbol{x}_s)$ is better than 0.1%. This sampling is described in more detail in appendix G.

On average, about 200 abscissae are needed to achieve the required accuracy of 0.1% in $\mathcal{I}_i(\mathcal{E}|\boldsymbol{x}_s)$ (see figure 12.6 (right)). Out of these points, about 40 pairs contribute more than 0.1% to the total integral. The relative error of a subset of the evaluated time-integrated event likelihoods $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ has been calculated by comparing the obtained results with an integral performed with a fine grid of t_s abscissae. The majority (about 95.9 %) of the evaluated integrals has a relative accuracy better than 10^{-3} (see figure 12.6 (right)).

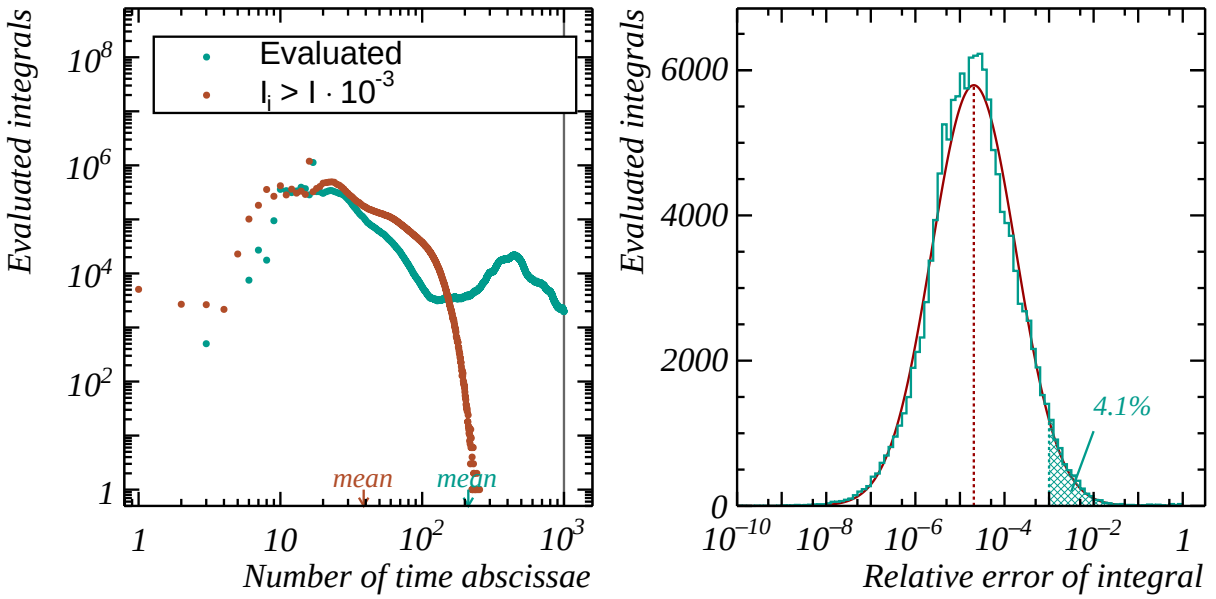


Figure 12.6: ***Left:*** *The distribution of the number of abscissae used to evaluate the time-integrated event likelihood $\mathcal{I}$ for a single simulated event. Two distributions of the number of abscissae are shown, the total number of abscissae at which the event likelihood is calculated (labelled 'evaluated'), and the number of (pairs of) abscissae that contribute at least 10^{-3} to the total time-integrated event likelihood $\mathcal{I}$ for any of the energy abscissae. The maximum allowed number of evaluated abscissae (1000) is indicated by the solid vertical line.* ***Right:*** *The relative accuracy of the numerically evaluated $\mathcal{I}$ with respect to the 'true' $\mathcal{I}$, evaluated using a detailed scan with a large number of abscissae.*

12.4 Event Likelihood Tables

The event likelihood table of an event is used to parameterise $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s)$. The event likelihood tables replace the reconstruction stage used in conventional neutrino source searches, and thus form a critical ingredient in the novel approach. Each table contains the values of $l_{\text{early } t_s}$, $l_{\text{late } t_s}$, t_s^{max} and $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ for a large number of abscissae covering the event signature direction $\hat{d}_s$, position $\vec{x}_s$ and energies E_S^{vis} and E_μ.

12.4.1 Abscissae

Ideally, the event likelihood tables cover the full phase-space of $\boldsymbol{x}_s$ with a large number of abscissae. Unfortunately, the evaluation of the event likelihood function is rather CPU-intensive. Evaluating the event likelihood of a 'typical' event at a single point in $\boldsymbol{x}_s$ takes about 150 milliseconds (see appendix app:EventLikelihoodSampling). Sampling the event likelihood of this event with, for example, ten points in each dimension of $\boldsymbol{x}_s$ would require a total computing time of approximately $10^8 \cdot 150[\text{ms}] \approx 5.7$ months.

A significant effort has been put into minimising the total amount of required CPU time, while covering the event likelihood landscapes with optimal accuracy. The main improvement is achieved by first evaluating the total expected number of photoelectrons $N_i(\boldsymbol{x}_s)$ for each PMT, and the time-dependent expected number of photoelectrons

$\dot{n}_i(t_i|\boldsymbol{x}_s)$ for each 'hit' PMT for a given event signature direction $\hat{d}_s$ and position $\vec{x}_s$. In this, the event signature direction is sampled using a grid of directions similar to the one given in table 11.2. For each direction, the perpendicular event signature position (see equation 12.16) is sampled first, followed by a scan over the longitudinal event signature position.

The abscissae describing the time-dependent expected number of photoelectrons of each hit PMT for each sampled $\hat{d}_s$ and $\vec{x}_s$ are copied to the RAM memory, and used in the sampling of the remaining event signature parameters E_S^{vis}, E_μ and t_s. The sampling of these parameters is also factorised. First, t_s is sampled. The values of $\dot{n}_i(t_i|\boldsymbol{x}_s)$ are determined, and used in a scan over a grid of energies in E_S^{vis} and E_μ.

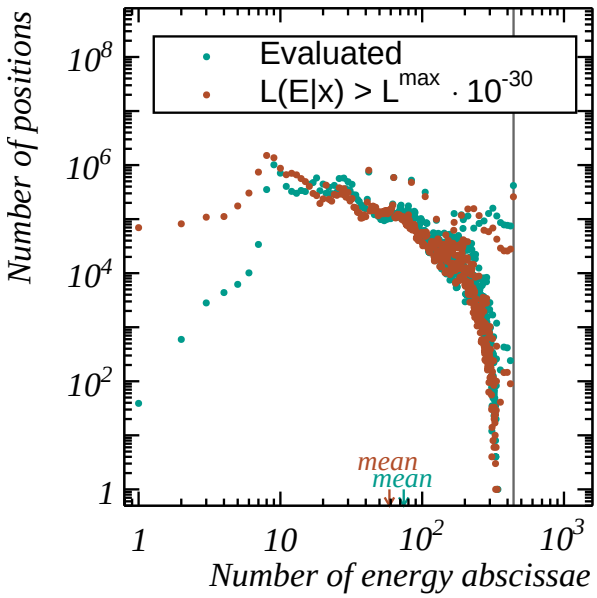
12.4.2 Output

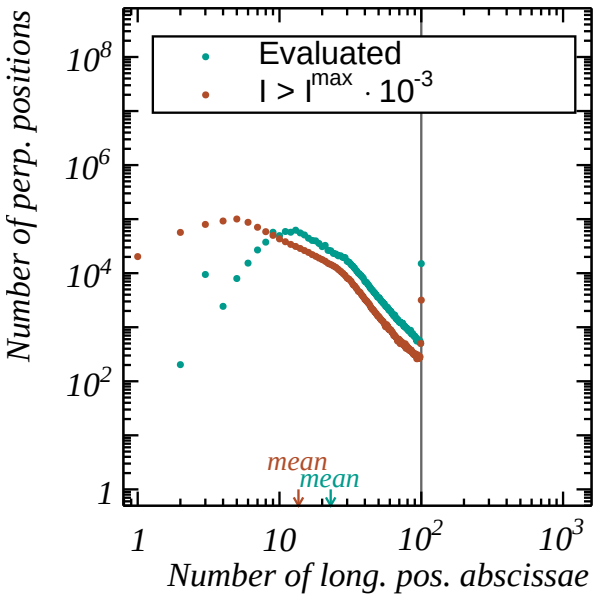
Each event likelihood table contains a header with an unique identifier of the event, the number of hit and non-hit PMTs in the triggered event, the time of the first triggered hit, and the time window used for the definition of these, the (estimated) rate of ^{40}K decays and two corresponding likelihoods: $l^{\text{hit PMTs}}(\mathcal{E}|^{40}K)$ and $l^{\text{non-hit PMTs}}(\mathcal{E}|^{40}K)$.

The main content of the event likelihood table are the calculated $l_{\text{early } t_s}$, $l_{\text{late } t_s}$, t_s^{max} and $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$. Any redundant information (such as the used grid of event signature energies $E_S^{\text{vis.}}$, E_μ and other scanned parameters of $\boldsymbol{x}$) is removed to minimise the size of the stored files.

12.4.3 Computational Requirements

A cluster of computers has been used to distribute the computational load required to obtain the event likelihood landscapes of a number of benchmark events. Each 'job', run on this cluster, consists of the scan over the event signature position, time and energy profiles for one of the event signature directions of a single event. The typical time consumption per sampled direction and the size of the corresponding output table is given in figure 12.8. Each job takes about 37 minutes to finish. For a fraction of the directions, the set upper limit of two hours computing time per job was not sufficient to obtain the event likelihood landscape. Typically, these directions deviate a lot from the true event signature direction, and the event likelihood of these landscapes is negligible in comparison to the event likelihood values of sampled directions close to the true direction. Therefore, the decision has been made not to re-run these jobs with a higher upper limit on the allowed CPU time.





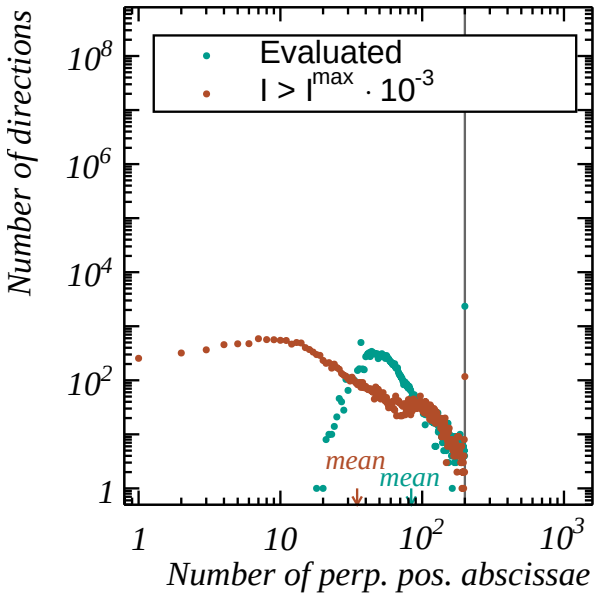


Figure 12.7: *The distribution of the number of abscissae at which the event signature energies E_S^{vis} and E_μ are evaluated for each point in $\vec{x}_s$ and $\hat{d}_s$ of a few benchmark events. The two-dimensional grid on which these abscissae lie consist of 441 points, indicated by the vertical line. The distribution indicated by $\mathcal{I} > \mathcal{I}^{max} \cdot 10^{-30}$ corresponds to the number of abscissae (out of the evaluated number of abscissae) at which the time-integrated event likelihood $\mathcal{I}$ exceeds 10^{-30} times the maximum $\mathcal{I}$ of all evaluated abscissae of each $\vec{x}_s$, $\hat{d}_s$. The number of abscissae used to sample the longitudinal (left figure) and perpendicular (right figure) event signature position. The maximum number of abscissae are indicated by the vertical lines.*

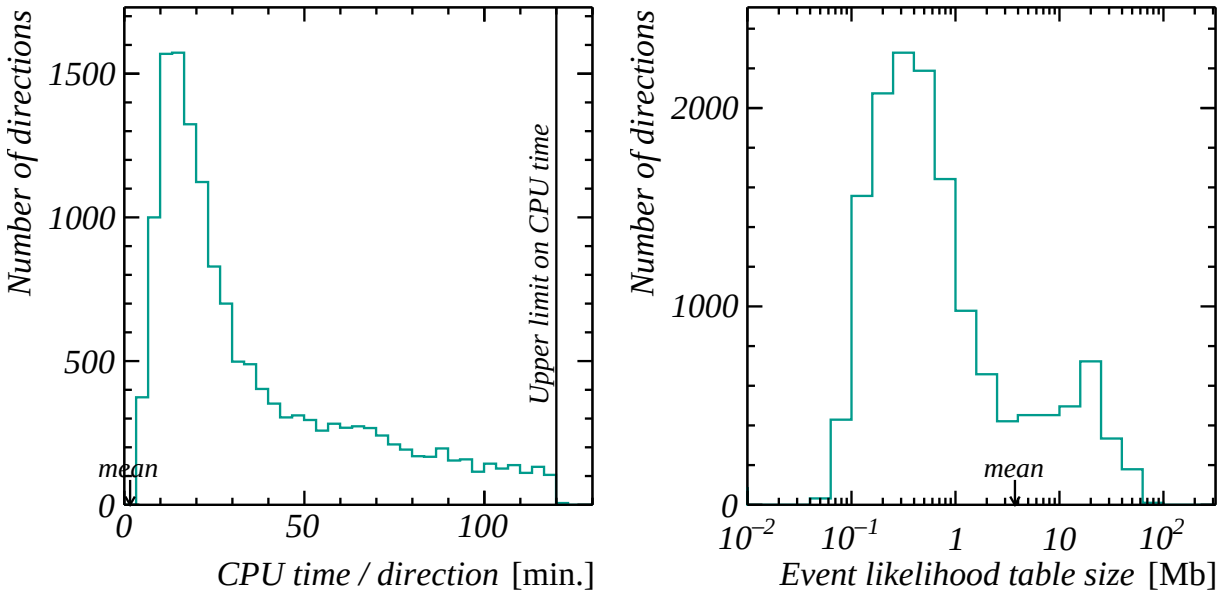


Figure 12.8: *The required CPU time (on a single 2 GHz processor) (**left figure**) and file size of the corresponding event likelihood table (**right figure**) per sampled direction of a few benchmark events. The mean of each distribution is indicated, as well as the imposed upper limit on the allowed CPU time of each 'job' to obtain the event likelihood table.*

Chapter 13

Evaluating the Test Statistic Integral

The backbone of the novel neutrino source search method is the test statistic integral

$$\mathcal{L}(\mathcal{E}_i|\mathcal{H}) \equiv \int \mathcal{L}(\mathcal{E}_i|\boldsymbol{x}_s) \cdot P^{det}(\boldsymbol{x}_s) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) d\boldsymbol{x}_s, \tag{13.1}$$

which is proportional to the probability that a detected event $\mathcal{E}_i$ results from a certain neutrino flux (i.e. a hypothesis $\mathcal{H}$). The subject of this chapter is the (numerical) evaluation of this integral.

13.1 Introduction

The (eight-dimensional) integrand of the test statistic integral is the product of three functions:

- The expected event signature flux for a hypothesis $\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H})$. Using the event signature probability tables discussed in section 11.1.4, this flux can be calculated for any neutrino interaction flux $\frac{d^6\Phi_\nu^{int.}}{dE_\nu d^2\Omega d^3\vec{x}}(E_\nu, \hat{d}_\nu, \vec{x}_\nu|\mathcal{H})$ at a certain neutrino energy, direction and position. The following neutrino fluxes are used as input: A background-only $\mathcal{H}_0$ hypothesis flux of atmospheric neutrinos, and 121 benchmark $\mathcal{H}_1$ hypothesis fluxes, each consisting of the atmospheric neutrino background flux with an additional diffuse astrophysical flux component. These fluxes are discussed in section 11.1.1.

- The probability $P^{det}(\boldsymbol{x}_s)$ that an event signature with parameters $\boldsymbol{x}_s$ is detected. The evaluation of this function has been discussed in section 11.2.4. The detector used in this work consists of a single KM3NeT/ARCA building block (115 DUs, running with stable performance over a period of ten years.

- The event likelihood 'landscape' $\mathcal{L}(\mathcal{E}_i|\boldsymbol{x}_s)$ of detected event i. The event likelihood is calculated from event likelihood tables, discussed in section 12.4.

The test statistic integral is integrated over all possible event signature parameters $\boldsymbol{x}_s$. For the event signature direction $\hat{d}_s$, this corresponds to the full solid angle. The range of the event signature position $\vec{x}_s$ is infinite, but is limited by the fact that the detection

efficiency converges to zero for positions far away from the instrumented volume. Similarly, the event signature energies E_S^{vis} and E_μ are limited by the detection efficiency converging to zero for low energies and the event signature flux converging to zero at high energies.

The integral limits of the event signature time are related to the definition of the event likelihood (see section 12.1) and detection efficiency (section 11.2), and correspond to the lifetime of the detector. This implies that, in theory, an event could be caused by a neutrino interaction happening long before (or after) the detected event. The probability for this to happen is typically very low ($l_{\text{early } t_s}$ and $l_{\text{late } t_s}$ are typically negligible in comparison with $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$) since all detected light should be caused by ^{40}K only, while the emitted light from the hypothetical neutrino interaction did not lead to a detected (i.e. triggered) event.

13.2 Numerical Evaluation

Computing an (eight-dimensional) integral is not particularly difficult if the integrand is computional inexpensive to evaluate. For the event likelihood function this is, unfortunately, not the case. Consequently, the integrand of the test statistic integral can only be evaluated at a relatively small number of points in each dimension. In order to interpolate between the sampled abscissae, a suitable parameterisation of the integrand is required, which complicates the evaluation of the integral.

An additional difficulty is caused by inaccuracies in the sampled event signature times to calculate the time-integrated event likelihood $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$. For some event signature positions, the sampled time abscissae do not cover the global maximum of the time-dependent event likelihood, resulting in a erroneous small time-integrated event likelihood $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$. Event signature positions for at which this is the case should be excluded from the evaluation of the integral.

The procedure used to calculate the test statistic integral under the $\mathcal{H}_0$ hypothesis (atmospheric neutrinos only) differs slightly from the procedure followed to calculate the test statistic integrals under the $\mathcal{H}_1$ hypotheses (an atmospheric neutrino flux with an additional diffuse cosmic signal) A schematic overview of these procedures is given in figure 13.1, and will be introduced in the following paragraphs. In the remainder of this chapter, each of the steps are discussed in more detail.

$\mathcal{H}_0$ hypothesis The calculation of the test statistic integral under the $\mathcal{H}_0$ hypothesis consists of the following steps:

1. For each sampled event signature position and direction, the three quantities ($\mathcal{L}_{\text{early } t_s}$, $\mathcal{L}_{\text{late } t_s}$ and $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$) describing the time-dependent event likelihood (see section 12.3) at each $\vec{x}_s$ and $\hat{d}_s$ are parameterised as a function of $E_S^{\text{vis.}}$ and E_μ.

2. The event signature probability tables are used to calculate the corresponding event signature flux $\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}_0)$, using the procedure described in section 11.1.7.

3. The parameterisations of the previous steps are used to calculate the time- and

energy integrated test statistic integral

$$\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H}_0) \equiv \int_0^\infty \int_0^\infty \int_{t_a}^{t_b} \mathcal{L}(\mathcal{E}|\boldsymbol{x}_s) \cdot P^{det}(\boldsymbol{x}_s) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}_0) dt_s dE_S^{\text{vis.}} dE_\mu, \tag{13.2}$$

where equation 12.22 is used to calculate the integral over t_s.

4. A three-dimensional multivariate normal distribution is fitted to each of the $\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H}_0)$ landscapes belonging to a single sampled event signature direction $\hat{d}_s$. An algorithm is used to exclude outliers (caused by the inaccurate sampling of t_s) from this fit. The fitted parameterisations are used to calculate the position-integrated test statistic integral

$$\kappa(\hat{d}_s|\mathcal{H}) \equiv \int_{-\infty}^{\infty} \kappa(\vec{x}_s, \hat{d}_s|\mathcal{H}) d^3\vec{x}_s. \tag{13.3}$$

5. The last step consists of a relatively simple fit of $\kappa(\hat{d}_s|\mathcal{H}_0)$ to calculate the test statistic integral

$$\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0) = \int_{4\pi} \kappa(\hat{d}_s|\mathcal{H}_0) d^2\hat{d}_s. \tag{13.4}$$

$\mathcal{H}_1$ hypotheses The test statistic integrals $\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)$ of the 121 benchmark $\mathcal{H}_1$ hypotheses are calculated in parallel with the calculation of $\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)$. The same steps as discussed in the previous paragraph are followed, but the outlier removal in the third step is not applied to these fits. Instead, the points $\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H}_1)$ at the same positions used in the fit of $\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H}_0)$ are used. This is done to minimise the probability that a large difference between $\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)$ and $\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)$ is caused by a different sampling, rather than by the underlying function.

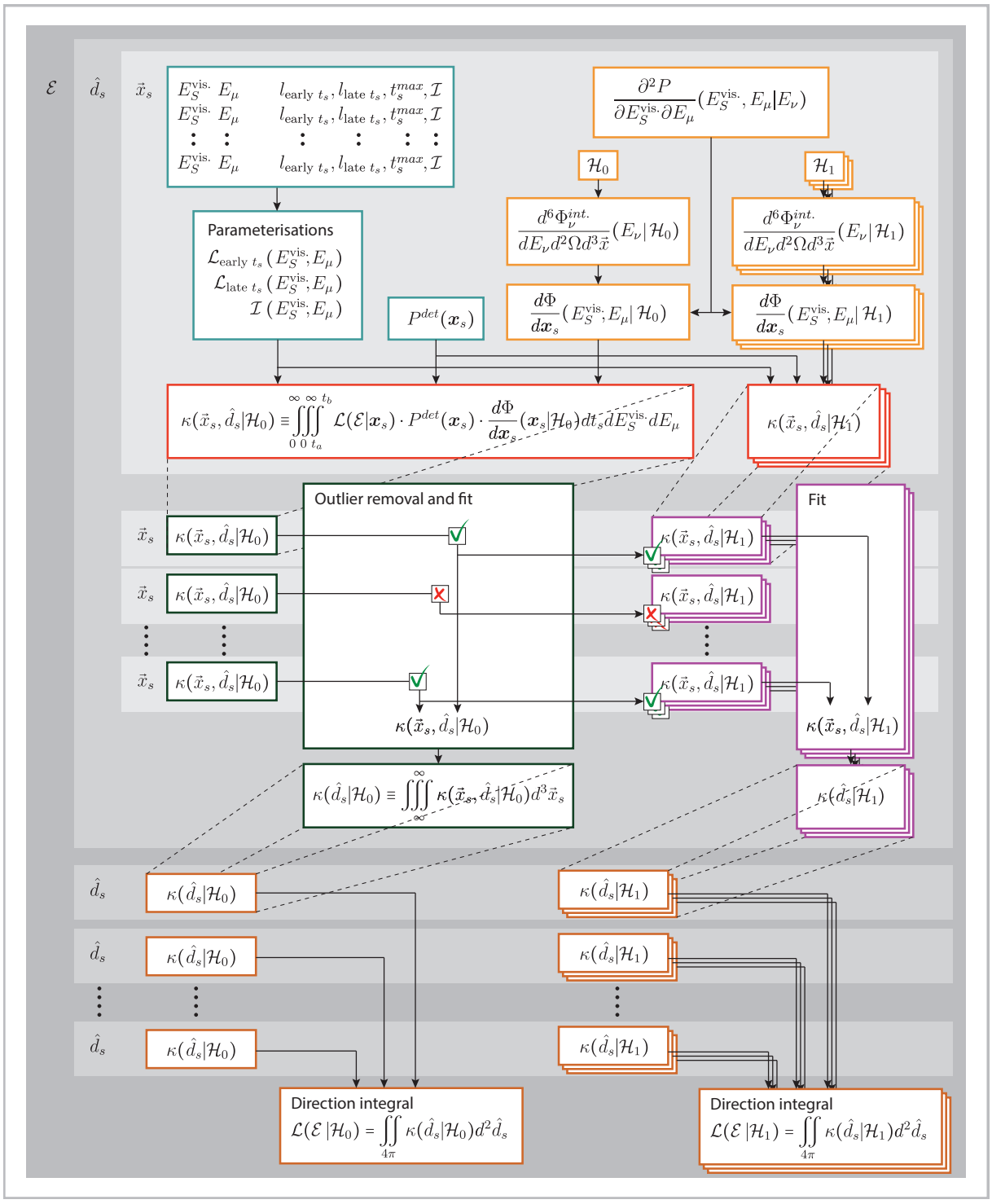


Figure 13.1: *An outline of the steps followed to calculate the test statistic integrals for the $\mathcal{H}_0$ and $\mathcal{H}_1$ hypotheses of a single event. The used symbols are discussed in the text. Each step (see text) is indicated with a different colour, the shaded areas correspond to steps in which the event signature position (light grey), direction (grey) or event (dark grey) is constant.*

13.2.1 Parameterisation of the time-dependent event likelihood

A two-dimensional spline is fitted to the logarithm of the time-dependent event likelihood $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ as a function of $\log_{10}(E_S^{\text{vis.}})$ and $\log_{10}(E_\mu)$. A fit of a relatively simple two-dimensional function, given by:

$$\mathcal{L} \approx \exp\left(-\left(a_{40K} + a_\mu \cdot E_\mu + a_S \cdot E_S^{\text{vis.}}\right)\right). \tag{13.5}$$

is used to parameterise the likelihood at early ($\mathcal{L}_{\text{early } t_s}$) and late ($\mathcal{L}_{\text{late } t_s}$) event signature time t_s. In figures 13.2, two examples of the fitted functions are given.

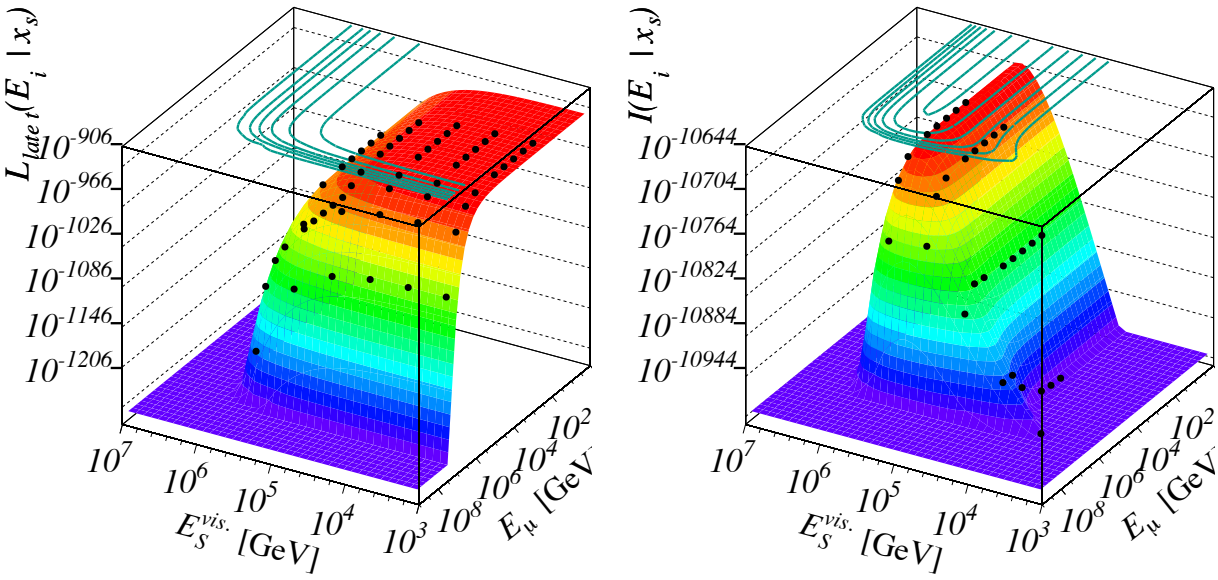


Figure 13.2: *An example of the fit to the log-likelihood $\mathcal{L}_{\text{late } t_s}$ (**left** figure) and $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ (**right** figure) as a function of the two event signature energies $E_S^{\text{vis.}}$ and E_μ. The points at which $\mathcal{L}_{\text{late } t_s}$ and $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ are calculated are indicated by the black points, the fitted functions are discussed in the text.*

13.2.2 Integral over event signature time and energies

The parameterisations of the time-dependent event likelihood are used to calculate the time-integrated event likelihood at any $E_S^{\text{vis.}}$ and E_μ according to equation 12.22. The integral over the event signature energies $\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H})$, i.e. equation 13.2, is calculated using Gauss-Legendre integration with 100×100 points in $\log(E_S^{\text{vis.}})$ and $\log(E_\mu)$ (see section 12.3.1 and Teukolsky et al.[217] for a description of Gauss-Legendre quadrature). This integral is evaluated for all $\mathcal{H}_0$ and $\mathcal{H}_1$ hypotheses in parallel, using the same parameterisations of the event likelihood, but with different input event signature fluxes.

13.2.3 Outlier removal and position-integrated test statistic

An example of the time- and energy-integrated test statistic integral $\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H}_0)$ as a function of the event signature position $\vec{x}_s$ for an arbitrary event and event signature

direction is given in figure 13.3. Close to the maximum, these points follow approximately a three-dimensional multivariate normal distribution

$$\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H}) \approx \kappa(\hat{d}_s|\mathcal{H}) \cdot \frac{1}{(2\pi)^{1.5}\,|\Sigma|^{0.5}} \cdot \exp\left(-\frac{1}{2}\,(\vec{x}_s - \vec{x}_s^{\,\text{max.}})^T \cdot \Sigma^{-1} \cdot (\vec{x}_s - \vec{x}_s^{\,\text{max.}})\right). \quad (13.6)$$

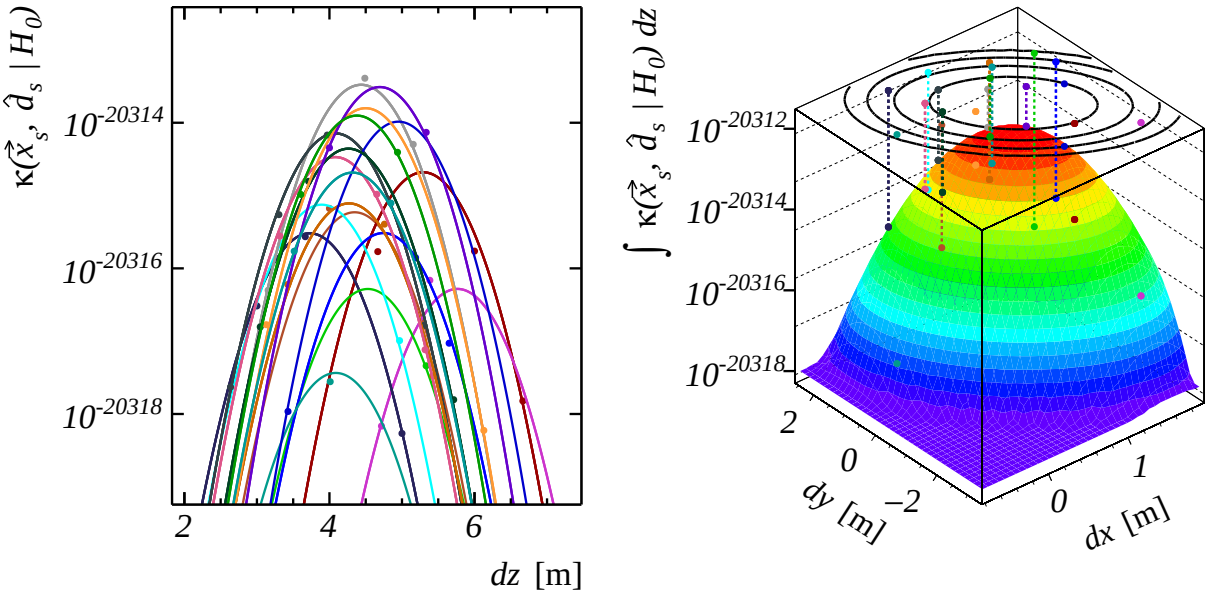


Figure 13.3: *An example of a fit of a three-dimensional multivariate normaldistribution to the energy and time-integrated test statistic (see text) of an arbitrary event at an arbitrary direction. The projections of the fit along this direction (in the* ***left*** *figure) correspond to the points with the same colour in the* ***right*** *figure, in which the projection perpendicular to the sampled direction is given. The point $dx = dy = dz = 0$ corresponds to the first sampled point of the position scan (see appendix G.)*

This distribution (10 free parameters) is fitted to a selection of the calculated $\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H})$ points. An outlier removal technique is implement to reject all points in the tails of the distribution. This outlier removal is also effective in removing all $\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H})$ points of which the t_s abscissae did not cover the global maximum of the time-dependent event likelihood, thus resulting in an erroneous (low) time-integrated event likelihood.

The aim of the outlier removal technique is to find the best combination of n out of N points with the best $\chi^2 = \sum_{i=1}^{n}\left[\left(\frac{f_i - y_i}{\sigma_i}\right)_i^2\right]$ per degree of freedom $n_{\text{d.o.f.}} = n - 10$. In order to find this combination of points, the following steps are followed:

1 All points except the 60 points with the highest $y_i = \kappa(\vec{x}_s, \hat{d}_s|\mathcal{H}_0)$ are excluded from the fit. The procedure allows for more than 60 points to be used, but the corresponding y_i are generally orders of magnitude smaller than the maximum.

2a The multivariate normal distribution is fit to a set of $11 \leq n \leq 60$ randomly selected points (each point having a 50% probability to be included), and the corresponding $Q = \chi^2/n_{\text{d.o.f.}}$ is calculated.

2b Next, the point with the lowest fit residual $\frac{f_i - y_i}{\sigma_i}$ from the remaining points is added to the set of randomly selected points, but only in case a new fit with this points leads to a better (i.e. lower) fit quality Q. This step is repeated until Q does not improve any further.

3 Steps 2a and 2b are repeated a 1000 times, each time using a different set of randomly selected n data points as a seed.

4 Out of all performed fits, the fit with the best Q is used as position fit.

In case $N < 15$ points are provided to the first step, there are less than 1000 unique combinations of points possible, and instead of randomly selecting combinations of points, all possible combinations are fitted.

The distribution of the fit quality of all sampled combinations of points for a single event and direction are displayed in figure 13.4 (left). The different bands correspond to different combinations of points in which one (or more) outliers are present.

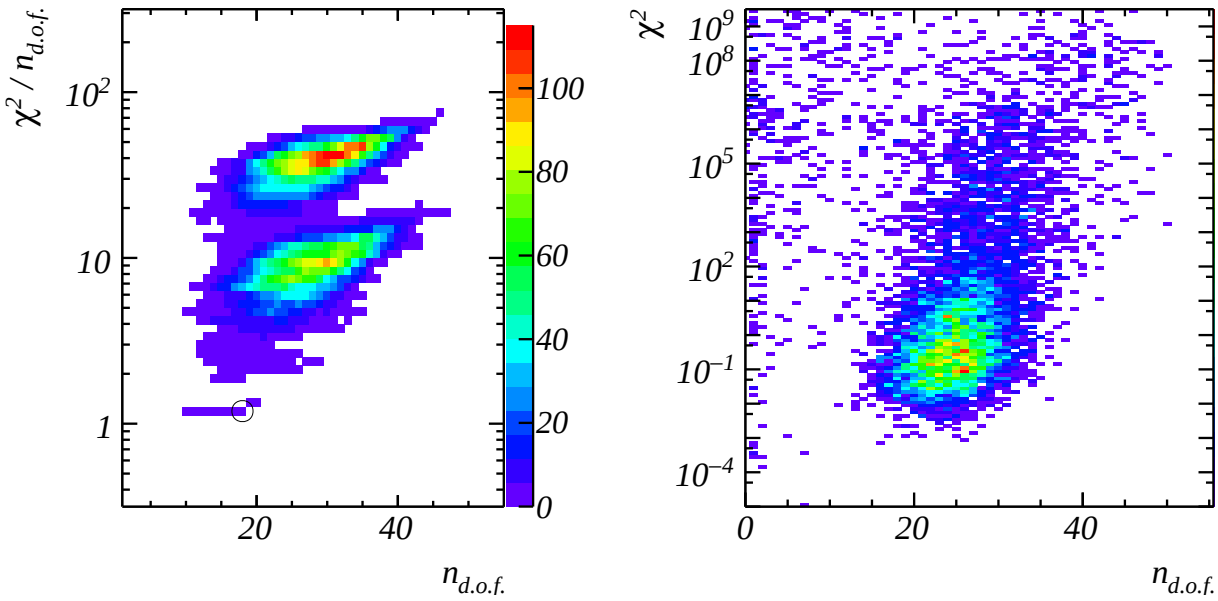


Figure 13.4: ***Left*** *figure: The distribution of the fit quality $Q = \chi^2/n_{d.o.f.}$ as a function of the number of degrees of freedom $n_{d.o.f.}$ used in each of the fits of the position profile with outlier removal (see text). This distribution is given for a single position fit for a single arbitrary event at an arbitrary direction. The fit with the best quality is indicated by the black circle. This point gives a single entry in the right figure, in which the quality of the position fit with the best Q is given for a large number of position fits.*

The normalisation $\kappa(\hat{d}_s|\mathcal{H})$ of the fitted distribution with the best fit quality Q is used as an estimate of the time-, energy- and position-integrated test statistic (equation 13.3).

$\mathcal{H}_1$ hypotheses For each $\mathcal{H}_1$ hypothesis, the position fit is applied to all points $\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H}_1)$ at the same positions as the points used in the fit of the $\mathcal{H}_0$ hypothesis. The fit quality of

the evaluated position integrals is shown in figure 13.4 (right). In general, the maximum of the fitted function agrees with the true event signature position (see figure 13.5).

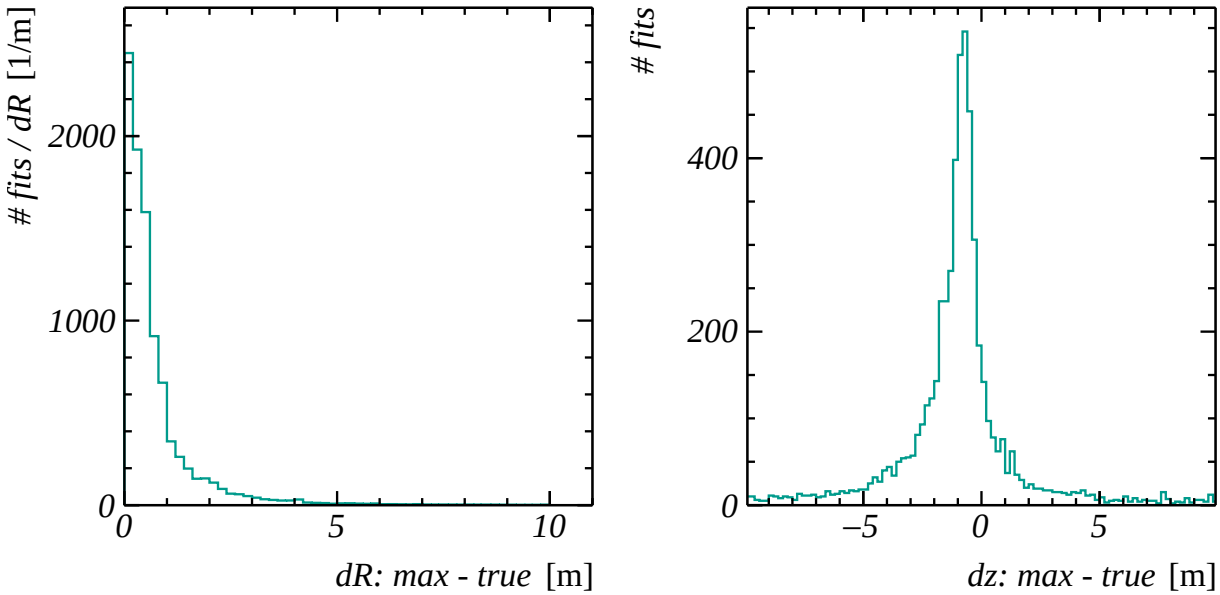


Figure 13.5: *The distance of the maximum of the fitted function to the time- and energy-integrated test statistic to the true event signature position along (**left** figure) and perpendicular (**right** figure) to the true event signature direction.*

13.2.4 Direction-integrated test statistic

Some examples of the direction-dependent 'landscapes' of the time- position- and energy-integrated test statistic $\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H}_0)$ (equation 13.3) are given in figure 13.6. The maximum of each landscape is located in the vicinity of the true direction of the simulated neutrino. However, due to the limited number of sampled directions, the region around the maximum is relatively sparsely sampled. A more refined sampling, in particular in the region close to the maximum, can relatively easily be achieved at the cost of more CPU time. For now, since the aim of this work is to show the feasibility of the method, the integral over the event signature directions is solved by Gauss-Legendre integration, using Delauney triangulation[218] to interpolate $\kappa(\vec{x}_s, \hat{d}_s|\mathcal{H})$.

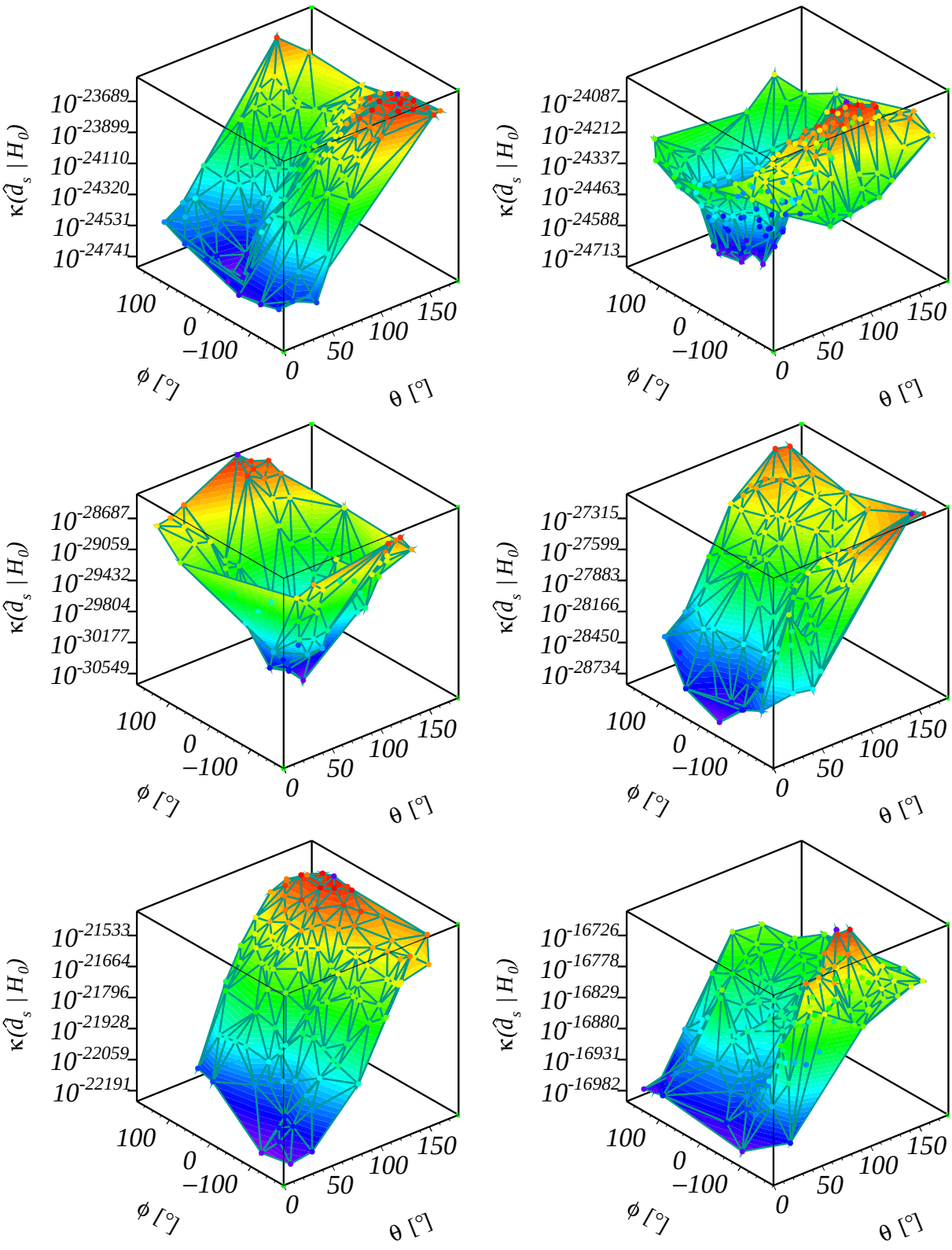


Figure 13.6: *The time, position and energy integrated test statistic (vertical axes) under $\mathcal{H}_0$ of six exemplary shower-like events, as a function of the event signature zenith (θ) and azimuth angle ϕ. Interpolation between the points at which this integral is calculated is achieved by Delauney triangulation [218] (see text).*

13.3 Results

Using the techniques discussed in the previous paragraphs, the test statistic integrals for each of the tested hypotheses (see section 11.1.1) have been calculated for a set of 60 simulated benchmark events. Triggered events originating from charged-current interactions of electron- or muon-neutrinos with a neutrino energy of at least 100 TeV are considered (28 and 32 events, respectively). These interactions are simulated from a $dN/dE \propto E^{-1.4}$ energy spectrum. The main characteristics of these events are summarised in table 14.5 in appendix H.

The ratio between the test statistic integral under any of the $\mathcal{H}_1$ hypotheses and the background-only hypothesis $\mathcal{L}(\mathcal{E}|\mathcal{H}_1)/\mathcal{L}(\mathcal{E}|\mathcal{H}_0)$ gives an estimate of how much more likely the event $\mathcal{E}$ is caused by a neutrino of cosmic origin than being caused by a neutrino produced in the Earths' atmosphere. To first order approximation, the event likelihood ratio $\mathcal{L}(\mathcal{E}|\mathcal{H}_1)/\mathcal{L}(\mathcal{E}|\mathcal{H}_0)$ is approximately equal to the flux ratio

$$\Psi_{\mathcal{H}_1/\mathcal{H}_0}(\boldsymbol{x}_{s,\,\mathrm{max}}|\mathcal{H}_1) \equiv \frac{\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_{s,\,\mathrm{max}}|\mathcal{H}_1)}{\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_{s,\,\mathrm{max}}|\mathcal{H}_0)} \tag{13.7}$$

at the event signature parameters $\boldsymbol{x}_{s,\,\mathrm{max.}}$ where the event likelihood $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s|)$ is maximal. The flux ratio $\Psi_{\mathcal{H}_1/\mathcal{H}_0}(\boldsymbol{x}_{s,\,\mathrm{max}}|\mathcal{H}_1)$ is, in turn, approximately proportional to the flux ratio

$$\Psi_{\mathcal{H}_1/\mathcal{H}_0}(\boldsymbol{x}_{s,\,\mathrm{true}}|\mathcal{H}_1) \equiv \frac{\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_{s,\,\mathrm{true}}|\mathcal{H}_1)}{\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_{s,\,\mathrm{true}}|\mathcal{H}_0)} \tag{13.8}$$

at the true (i.e. simulated) event signature parameters $x_{s,\,\mathrm{true}}$. The quantity

$$\zeta(\mathcal{E}_i|\mathcal{H}_1) \equiv \frac{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)/\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)}{\Psi_{\mathcal{H}_1/\mathcal{H}_0}(\boldsymbol{x}_{s,\,\mathrm{true}}|\mathcal{H}_1)} \tag{13.9}$$

indicates how much more '$\mathcal{H}_1$-like' the detected event is than the true event signature parameters $\boldsymbol{x}_{s,true}$. In this work, where each $\mathcal{H}_1$ hypothesis represents a diffuse flux, the detected event is '$\mathcal{H}_1$-like' in case the event signature energies $E^{\mathrm{vis.}}_{S,\,\mathrm{max}}$ and $E_{\mu,\,\mathrm{max}}$ at $\boldsymbol{x}_{s,\,\mathrm{max.}}$ are relatively high.

The evaluated likelihood ratios $\mathcal{L}(\mathcal{E}|\mathcal{H}_1)/\mathcal{L}(\mathcal{E}|\mathcal{H}_0)$ are shown as function of $\Psi_{\mathcal{H}_1/\mathcal{H}_0}(\boldsymbol{x}_{s,\,\mathrm{true}}|\mathcal{H}_1)$ in figure 13.7. Except for a few events in which the test statistic integrals of a few $\mathcal{H}_1$ hypotheses are not evaluated correctly, the distribution of the points follows from the characteristics of the event:

- All points are located at the point 1,1 in case the dominating fraction of the event likelihood of an event is located in the region of $\boldsymbol{x}_s$ where the $\mathcal{H}_0$ event signature flux dominates the total event signature flux (in this case low $E^{\mathrm{vis.}}_{S,\,\mathrm{max}}$ and $E_{\mu,\,\mathrm{max}}$). Such points are (deliberately) underrepresented in this figure, by selecting only events with $E_\nu > 100\mathrm{TeV}$.

- For some events, all points are distributed along the 'horizontal' line where $\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)/\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0) \approx 1$. Like the previous point, the corresponding event likelihood landscape is situated in the parameter space of $\boldsymbol{x}_s$ where $\mathcal{H}_0$ dominates the total flux (i.e. low energies).

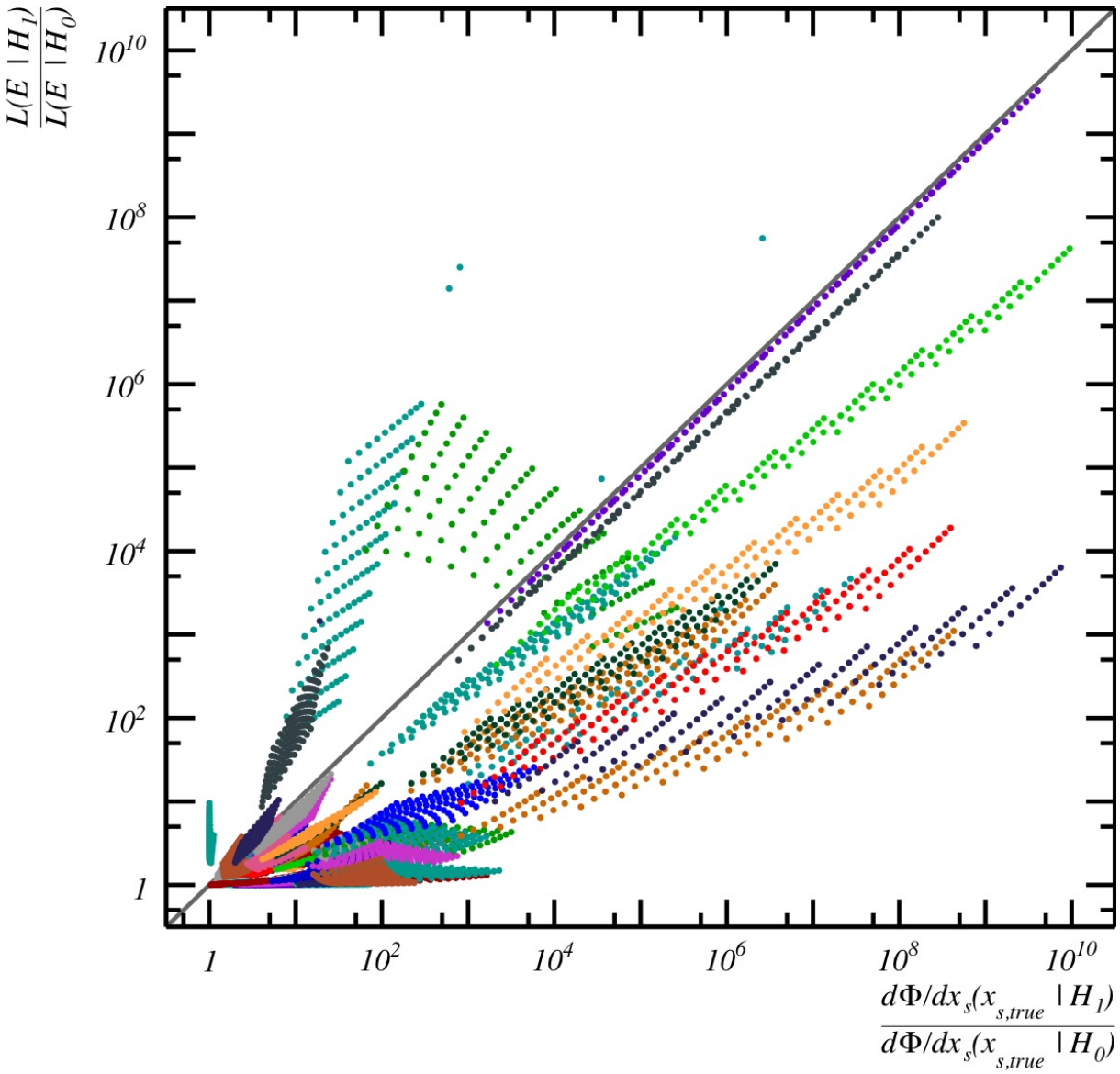


Figure 13.7: *The likelihood ratio between the test statistic integrals* $\frac{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)}{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)}$ *for the* $\mathcal{H}_1$ *hypotheses and the* $\mathcal{H}_0$ *hypothesis for each of the benchmark events (see text and table 14.5) as a function of the ratio between the corresponding expected event signature fluxes* $\frac{d\Phi/d\boldsymbol{x}_s(\boldsymbol{x}_{s,true}|\mathcal{H}_1)}{d\Phi/d\boldsymbol{x}_s(\boldsymbol{x}_{s,true}|\mathcal{H}_0)}$ *at the true event signature parameters* $\boldsymbol{x}_{s,true}$*. Each point corresponds to a single event and a single benchmark* $\mathcal{H}_1$ *hypothesis (see section 11.1.1). Points corresponding to the same event (but different* $\mathcal{H}_1$ *hypotheses) have the same colour. In the hypothetical case that the detector has a perfect resolution, the true event signature parameters can be determined with infinitesimal resolution, and the event likelihood function is given by a delta function at* $x_{s,\,\mathrm{true}}$*. Such a relation is indicated by the grey diagonal line.*

In contrast to the previous point, however, the true event signature parameters are '$\mathcal{H}_1$-like' ($\Psi_{\mathcal{H}_1/\mathcal{H}_0}(\boldsymbol{x}_{s,\,\mathrm{true}}|\mathcal{H}_1) > 1$). In conventional neutrino source searches, these

events are likely to be considered as 'mis-reconstructed' events (with an underestimated energy), and are typically excluded from the analysis.

- If $\zeta(\mathcal{E}_i|\mathcal{H}_1) \approx 1$ for an event, the 'best estimate' event signature energies $E^{\text{vis.}}_{S,\,\text{max}}$ and $E_{\mu,\,\text{max}}$ are approximately equal to the true event signature energies $E^{\text{vis.}}_{S,\,\text{true}}$ and $E_{\mu,\,\text{true}}$. If $\zeta(\mathcal{E}_i|\mathcal{H}_1) > 1$, then $E^{\text{vis.}}_{S,\,\text{max}}$ and $E_{\mu,\,\text{max}}$ are higher than $E^{\text{vis.}}_{S,\,\text{true}}$ and $E_{\mu,\,\text{true}}$.

- One event (in the lower left corner of the figure) has particular high values of $\zeta(\mathcal{E}_i|\mathcal{H}_1)$. This event is shower-like ($E^{\text{vis.}}_S = 10^{5.3}\text{GeV}$) but, in contrast to the other shower-like events, has a nonzero track energy ($E_\mu \approx 1\text{GeV}$). The corresponding muon is produced in the hadronic shower. The event signature flux $d\Phi/d\boldsymbol{x}_s(\boldsymbol{x}_{s,true}|\mathcal{H})$ at $E_\mu \approx 1\text{GeV}$ is significantly lower than the flux at $E_\mu = 0$, thus resulting in high $\zeta(\mathcal{E}_i|\mathcal{H}_1)$. This feature is discussed further detail in section 14.2.

Weighted with the event weights for either the $\mathcal{H}_0$ or $\mathcal{H}_1$ flux, the likelihood ratio $\frac{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)}{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)}$ forms the main input to the test statistic λ (equation 10.17). The larger the spread in $\frac{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)}{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)}$ for the different $\mathcal{H}_1$ hypotheses, the more the test statistic distributions of the different $\mathcal{H}_1$ hypotheses differ, and the better the different hypotheses can be distinguished from each other. Consequently, the relative contribution of a single event to the total sensitivity is related to the range of the covered $\frac{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_1)}{\mathcal{L}(\mathcal{E}_i|\mathcal{H}_0)}$ points. If this range is large, the different hypotheses are well distinguishable from each other. From the figure, it can be concluded that, as expected, mainly high-energy events contribute to the sensitivity.

Chapter 14

Conclusions, Discussion and Outlook

A novel method to identify and study the source(s) of cosmic neutrinos using data from the KM3NeT/ARCA detector has been developed. This method is designed to approach the best possible sensitivity (see chapter 10), such that the properties of these source(s) can be determined as precisely as possible.

The test statistic integral (equation 10.16) forms the backbone of the new method. Maximising the information included in each of the functions (the expected event signature flux, the detection efficiency and the event likelihood) in the integrand maximises the sensitivity. The new method focuses in particular on maximising the included information of each detected event, contained in its 'likelihood landscape'. In this chapter, the novel method is discussed, conclusions are drawn and suggestions for the application of the developed method are given.

14.1 Conclusions

The main conclusion of this work is that the novel approach to neutrino source searches works, thus providing a flexible framework in which any hypothesis (i.e. neutrino source) can be tested, all events can be used (regardless of the detector geometry), and almost all information of each event is used.

14.1.1 Event Likelihood

One the unique features of the event likelihood landscapes is that the resolution of the detector, limited by stochastic processes such as the propagation of light through the water and the light detection of the detector, is automatically taken into account in the most precise way possible. The event likelihood landscape is relatively flat in case little information can be extracted, while it is characterised by a clear maximum in case the event provides much information about the interacting neutrino. The event likelihood landscape describes the resolution of the 'reconstruction' in the vicinity of the global maximum.

The event likelihood function used in this work (section 12.1) includes almost all information provided by each detected (i.e. triggered) event. The information of both hit PMTs and non-hit PMTs is used in the event likelihood. The additional information

from the non-hit PMTs provides a better sensitivity to the energy of interacting neutrinos, and will also help in distinguishing neutrino events from (atmospheric muon) background events.

14.1.2 Event Selection

In conventional methods, event selection criteria are used to reject background events (mainly caused by atmospheric muons) and mis-reconstructed events. In the example study discussed in section 5.5, a significant fraction of the events is rejected from the analysis (see figure 5.3). The rejected events include high-energy events, which could potentially provide valuable information about their (cosmic) source(s).

In the novel method, there is no need for any event selection, since the probability that an event is 'mis-reconstructed' is inferred from the presence of one or more local maxima besides the global maximum. Even though not necessary from a principal point of view, event selection criteria can be implemented in a straightforward way to decrease computing time spent on, for example, (atmospheric muon) background events.

14.1.3 Event Signature Characterisation

The better an event can be characterised, the more information can be extracted from that event, thus resulting in a better sensitivity. Conventionally, each event is reconstructed twice under the assumption of a shower or track, with a subsequent 'flavour identification' routine in which the event is classified as a shower or track by a routine (see chapter 5). In the novel method, showers and tracks are treated equivalently, and the 'flavour identification' step is redundant.

In this work, a new event signature parameterisation has been introduced, in which each detected neutrino interaction is characterised by eight parameters; the interaction position and time, the direction which the shower and muon are assumed to propagate, the energy of the muon and/or visible shower energy. This relatively simple characterisation covers the main event signatures; electron and muon neutrino charged current interactions and all neutral current interactions, and can be extended to describe muon tracks with a single high-energy bremsstrahlung shower along its path.

The event signature parameterisation can be extended to include any event signature, as long as the probability density function which gives the expected number of photo-electrons on a PMT for such a signature can be calculated. A few examples of possible extensions are discussed in the following paragraphs.

Tau Leptons The tau lepton produced in tau neutrino interactions decays to one or more pions, to an electron or to a muon. The resulting event signatures (see section 3.3) can be characterised with one additional energy parameter; the energy, E_τ, of the tau. In figure 14.1, this characterisation is sketched.

The visible shower energy E_S^{vis} relates to the visible energy of the shower(s) produced in the primary interaction, the tau energy E_τ gives a measure for the range of the tau before decay (given by equation 3.17), and thus the position where a second shower (in case the tau decays to an electron or hadrons) or a muon is produced.

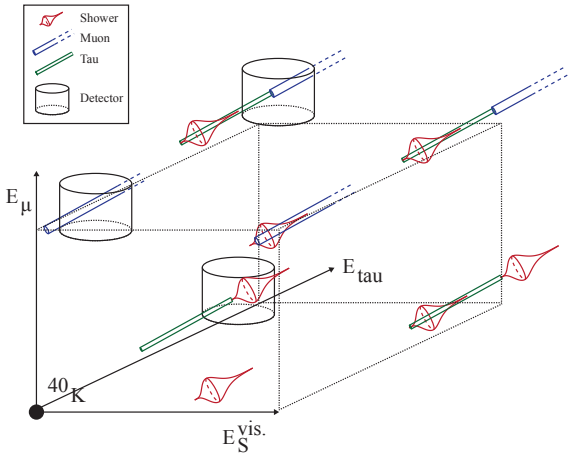


Figure 14.1: *A schematic overview of the characterisation of the different event signatures by three energies: The visible shower energy E_S^{vis} of the shower(s) at the neutrino interaction, the energy E_τ of the emerging tau lepton ($E_\tau = 0$ in case no tau is produced) and the energy E_μ of a muon produced in the neutrino interaction ($E_\tau = 0$) or in the decay of the tau ($E_\tau \geq 0$). In case a tau decays to an electron, $E_\tau \geq 0$ and $E_\mu = 0$, which means the visible shower energy of the second shower equals E_τ.*

Including tau events is of particular interest since the number of high-energy tau-like events in the atmospheric neutrino background is negligible. A typical hadronic shower produced in charged-current ν_τ interaction has a length of a few meters, while the energy of the tau produced in these interactions has to be at least 100 TeV in order to propagate 4.9 m (or more) before decaying to a muon or second shower. Each detected tau event with a clear separation between the two showers is therefore (almost) certainly caused by a neutrino of cosmic origin.

Stochastic Energy Losses of Muons A high-energy muon looses a significant fraction of its energy due to stochastic energy losses, such as Bremsstrahlung, e^+e^- pair-production and photonuclear interactions (see section 3.2.5). In figures 14.2, the total stochastic energy loss of a muon and the maximum energy loss in a single interaction are given for various initial muon energies.

For initial muon energies higher than approximately 10^4 GeV, a muon looses almost all of its energy due to these stochastic energy losses. At these energies, about 50% of the muons have at least one shower with more than 20% of the initial muon energy. It is very likely that this shower (along the muon trajectory) is more energetic than the primary shower produced at the neutrino interaction.

In its current form, the event signature characterisation is based on the assumption that each muon has a single shower at its origin. The light emitted by stochastic energy losses are included in the event likelihood function by assuming an 'average' stochastic energy loss (see appendix D). A possible way to improve this characterisation is by 'reconstructing' the total muon energy by an iterative procedure in which the number of showers along the muon trajectory is increased in every step. In the first step, the position along the trajectory and the visible energy of a single shower is estimated. This position is used as a seed for the second step, in which a the position and visible energy of two showers are estimated. This procedure is repeated (adding a single shower in each step), until the total energy of all showers converges. This value can then be taken as total muon energy.

Muon Bundles Muon bundles are produced in cosmic ray air showers. The muons in such a bundle with an energy high enough to reach KM3NeT/ARCA move approximately

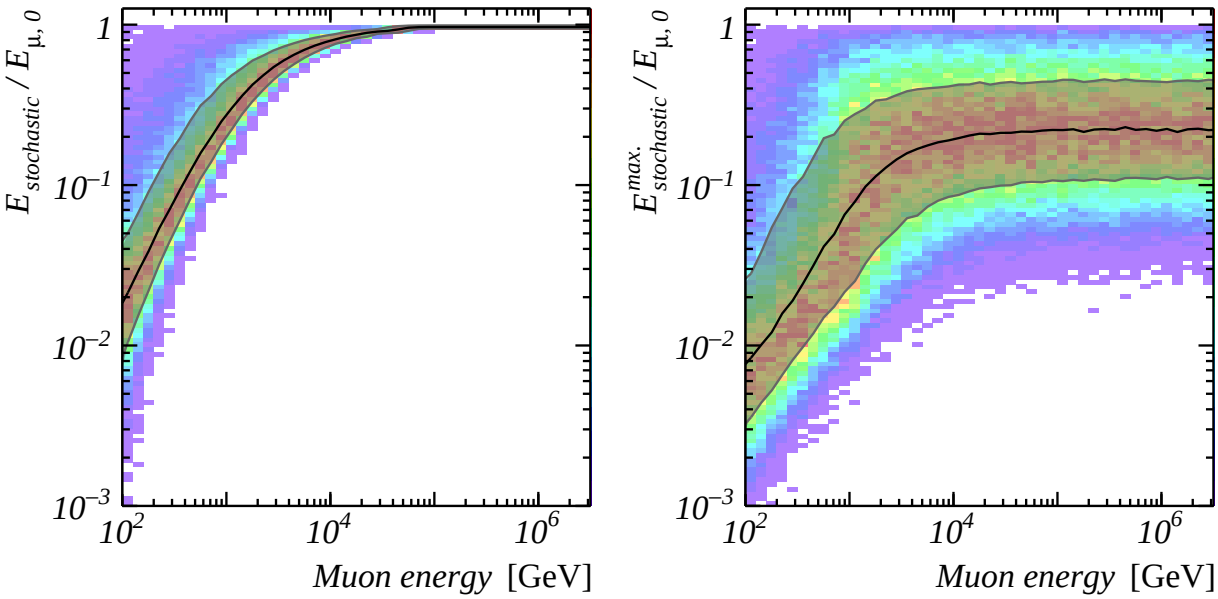


Figure 14.2: *The summed energies of all showers (**left** figure) or energy of the most energetic shower (**right** figure) of all showers produced by stochastic energy losses of a muon. These energies are given with respect to the initial energy of the muon $E_{\mu,0}$. The distributions are obtained from a large set of simulated muons, obtained using the JSirene simulation software (see section 4.4.1). The grey band indicates the 16-84% percentiles of the distribution, the black line the median.*

parallel, typically several meters apart[219].

The first step in dealing with atmospheric muon bundles is to reject them as much as possible. The main difficulty in this rejection is that the current reconstruction methods are not tuned to reconstruct multi-muon events. Such a specialised reconstruction, however, can be of great help in identifying these events, as well as characterising the corresponding flux for neutrino source searches. Additionally, such a reconstruction could be a valuable starting point for studies aimed at cosmic rays and CR air showers.

The used event signature parameterisation does not describe events from atmospheric muon bundles. It could be extended to include these event, by extending the characterisation with (multiple) additional parameters describing the number and spatial distribution of the muons in the bundle.

14.1.4 Detection Efficiency

The expected flux and detection efficiency are generally parameterised with relatively simple functions, depending only on the zenith angle, energy, flavour and interaction channel of the neutrino. The additional dimensions of the detection efficiency tables used in this work can provide additional information.

The detection efficiency used in this work describes a stationary KM3NeT/ARCA building block with nominal PMTs. In the novel neutrino source search, it is very straightforward to obtain and use detection efficiency tables for any detector geometry and/or realistic PMT efficiencies.

14.1.5 Time-Dependent and Other Cosmic Neutrino Source(s)

The method can be used to test the existence of any cosmic neutrino source; a single point source, diffuse flux, collection of objects, time-varying source etc. The corresponding flux should be provided when evaluating the test statistic integral (chapter 13). This procedure will be illustrated by the following example.

Time-Varying Source Location The benchmark $\mathcal{H}_0$ and $\mathcal{H}_1$ event signature fluxes used in this work are constant in time. For a neutrino flux from (a) cosmic point source(s) this is not the case, since the location of a point source (with respect to the orientation of the detector) varies with time. The event signature flux of such a source used in the test statistic integral for a detected event is therefore calculated using the detection time of the event. The location of the source at this time can be determined using a positional astronomy library such as slalib[220]. The test statistic integral over all possible event signature directions reduces to a single direction. For a collection of point sources, a sum over the corresponding source directions can be used.

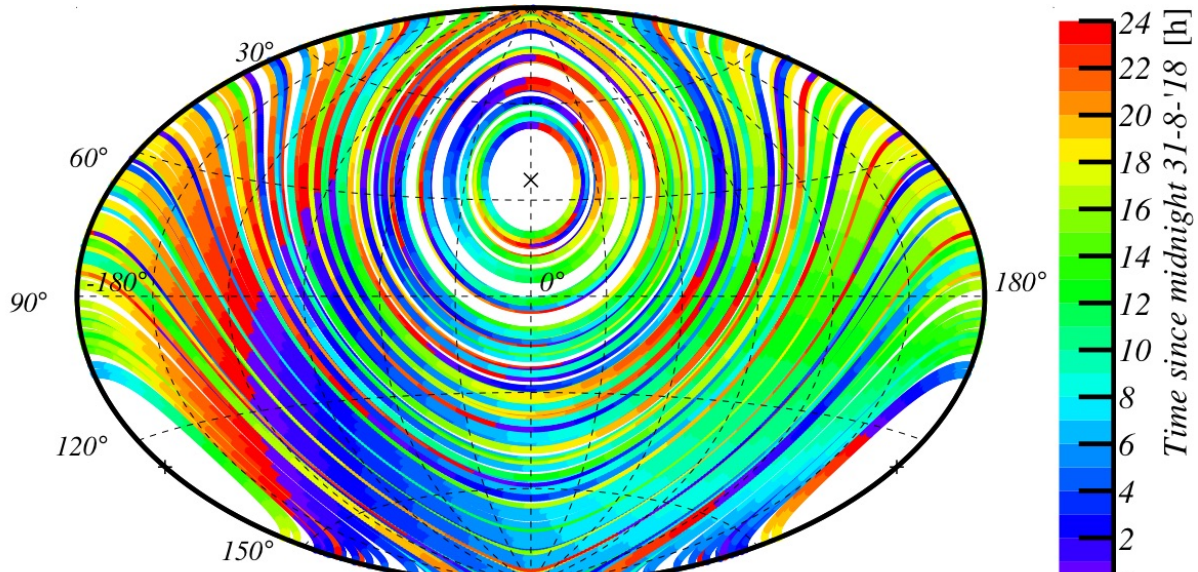


Figure 14.3: *The trajectories of the 167 ≥1 TeV γ-ray sources closest to the Earth. These trajectories are shown in local coordinates (given by the zenith and azimuth angle) with respect to the location of KM3NeT/ARCA on 31-8-2018. The 'x' and '+' indicate the locations of the North and South poles, respectively. The colour represents the time during the day. The locations of these sources are taken from the TeVCat catalogue[72], the trajectories are calculated using the Astropy software package[221, 222]*

14.1.6 Feasibility of the Novel Approach

Since KM3NeT/ARCA is not (yet) fully operational, simulations have been used to develop and investigate the feasibility of the new method. The main results have been

presented in the previous chapter. The fact that the test statistic integrals of these events can be calculated forms an important milestone, since it proves that the novel method works.

The main downside of the novel method is the significant computing time required, in particular to obtain the tables with event likelihood landscapes. This was evident from the beginning, as a logical consequence of the design choice to minimise the number of used simplifications and assumptions.

Obtaining the event likelihood landscapes for the 60 benchmark events takes a total of about 1500 computing hours (in this case distributed over many computers). Some suggestions to improve the method and reduce the required computing time will be discussed in the next section. Fortunately, the method does not have to be applied to the full sample of all detected events in order to prove its usability, as will be discussed in section 14.3.

14.2 Discussion and Possibilities for Improvement

Limitations in time (from both the author and the used computers) have left some possibilities of improvement open. An overview of these points will be given in this section.

14.2.1 Treatment of Backgrounds

In its current form, the expected event signature flux used in the novel neutrino source search method accounts only for the atmospheric neutrino background. The backgrounds of atmospheric muons, bioluminescence and ^{40}K decays has not been included. The first step in dealing with these non-neutrino backgrounds is rejecting the events caused by them from the data sample. Ways to do this have been discussed in section 5.4. Unfortunately, such an event selection is rarely 100% successful in rejecting background events. The remaining flux should therefore be included in the test statistic integral (equation 10.16), for example in the following form:

$$\begin{aligned}
\mathcal{L}(\mathcal{E}|\mathcal{H}) = & \int \mathcal{L}(\mathcal{E}_i|\boldsymbol{x}_s) \cdot P^{det}(\boldsymbol{x}_s) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) d\boldsymbol{x}_s && \text{(neutrinos)} \\
& + \mathcal{L}(\mathcal{E}|E_S^{vis.} = E_\mu = 0) \cdot R_{40K}^{det.} && (^{40}\text{K decays}) \\
& + \int \mathcal{L}(\mathcal{E}|\boldsymbol{x}_{\text{bl}}) \cdot \frac{d\Phi_{\text{bl}}^{det.}}{d\boldsymbol{x}_{\text{bl}}}(\boldsymbol{x}_{\text{bl}}) d\boldsymbol{x}_{\text{bl}} && \text{(bioluminescence)} \\
& + \int \mathcal{L}(\mathcal{E}|\boldsymbol{x}_\mu) \cdot \frac{d\Phi_{\text{single atm. } \mu}^{det.}}{d\boldsymbol{x}_\mu}(\boldsymbol{x}_\mu) d\boldsymbol{x}_\mu && \text{(single atm. muon events)}
\end{aligned}$$

with

- $R_{40K}^{det.}$ the expected rate of triggered events caused by ^{40}K decays. This rate can be obtained from simulations[181].
- $\mathcal{L}(\mathcal{E}|E_S^{vis.} = E_\mu = 0)$ the event likelihood at $E_S^{vis} = E_\mu = 0$.
- $\boldsymbol{x}_{\text{bl}}$ the position and visible shower energy of a bioluminescence burst.

- $\frac{d\Phi^{det.}_{\text{bl}}}{d\boldsymbol{x}_{\text{bl}}}(\boldsymbol{x}_{\text{bl}})$ the expected rate of bioluminescence bursts.
- $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_{\text{bl}})$ the event likelihood at $\boldsymbol{x}_{\text{bl}}$. This event likelihood relates to the event likelihood of neutrino events as $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_{\text{bl}}) = \iint \mathcal{L}(\mathcal{E}|\boldsymbol{x}_s) d\hat{d}_s$ with $E_\mu = 0$.
- $\boldsymbol{x}_\mu$ the position, time, direction and energy of a single atmospheric muon.
- $\frac{d\Phi^{det.}_{\text{single atm. }\mu}}{d\boldsymbol{x}_\mu}(\boldsymbol{x}_\mu)$ the expected flux of single atmospheric muons after event selection criteria.
- $\mathcal{L}(\mathcal{E}|\boldsymbol{x}_\mu)$ the event likelihood at $E_S^{vis} = 0$.

Note that the expected flux of ^{40}K -decay events, bioluminescence, atmospheric muons and neutrino events should be parameterised including the event selection criteria used to reject background events. Depending on the complexity of the flux and these criteria, the flux can either be calculated or estimated using simulations.

14.2.2 Expected Flux

The expected event signature flux $\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H})$ is calculated from the expected flux of neutrinos arriving at the Earths' surface, taking into account the effects of the propagation of the neutrino through the Earth and neutrino flavour-specific interaction kinematics (see e.g. section 11.1). However, two factors are not taken into account in the calculated flux.

Neutrino Propagation through the Earth The treatment of the Earths' influence on the neutrino flux is relatively simple (see e.g. section 11.1.2). In particular the effects of tau neutrino regeneration and neutral current interactions are not taken into account satisfactorily. This could be done by solving the transport equation of neutrinos propagating through the Earth, for example using the methods discussed by Vincent et al.[223]. This method can also be used to account for the effects of neutrino oscillations, which is expected to be prominent at neutrino energies below 100 GeV.

Muons produced in hadronic showers Muons can be produced in the hadronic showers produced in charged- and neutral-current interactions as well as $\bar{\nu}_e + e^- \rightarrow$ hadrons. Currently, the corresponding event signature fluxes are not accounting for these muons (section 11.1.5). An example of this inaccuracy is given in section 13.3.

Even though the effect of this correction to the evaluated test statistic integrals is expected to be minimal, muons produced in hadronic showers can be included by adding an additional dimension (the summed energy of all muons that can induce Cherenkov radiation produced in the hadronic shower) to the probability density function $\frac{\partial P}{\partial F_S^{\text{vis.}}}(F_S^{\text{vis.}}|E_S)$ which, in its current form, describes the visible shower energy (section 11.1.5).

14.2.3 Detection Efficiency

The accuracy of the estimated detection efficiency is limited, due to the finite number of simulated events per abscissa. A (rather obvious) way to improve the accuracy of the

detection efficiency function is to include more sampling abscissae and use a larger number of simulated events per abscissa. The inaccuracy of $\tilde{P}^{det}$ could potentially be an issue if $\tilde{P}^{det} = 0$ in a certain region, while the true detection efficiency is small but non-zero. In case the event likelihood of a signal event is large in such a region, its test statistic integral will be zero. Therefore, it will be attributed to the background-only hypothesis, and is 'missed' in the analysis. In order to solve this issue, it may be better to set $\tilde{P}^{det}$ to a (tiny) value in case $\tilde{P}^{det} = 0$.

The detection efficiency has to account of any event selection criteria used to reject background events (see 'treatment of backgrounds' in this chapter). This can be done by implementing the selection criteria as last step in the procedure outlined in section 11.2.

14.2.4 Event Likelihood Function

Even though the event likelihood function includes a significant fraction of the information provided by each event, it can easily be replaced by an event likelihood which includes even more information from the event or requires less time to compute. The latter improves the accuracy of the event likelihood tables, since more points can be sampled.

A few possibilities to improve the event likelihood function will be discussed in this section. Improvement in the way the event likelihood tables are obtained will be discussed in the subsequent section.

First Hit Selection Approximately 353 out of the 64170 PMTs in one KM3NeT/ARCA building block are 'hit' only because of ^{40}K -decays in the currently used time window of 1100 ns (see section 12.1.1). This time window is relatively large, and little time has been spent on optimising it. By reducing the length of this time window, less PMTs are hit, which speeds up the evaluation of the event likelihood. A potential improvement of the time window can be achieved by using a time window based on the best estimate track or shower parameters from event reconstruction algorithms*.

Time-over-Threshold Information The current event likelihood function distinguishes only between 'hit' and 'non-hit' PMTs. The information contained in the number of detected L0s and the time over threshold of each L0 is not used. This information is expected to improve, in particular, the accuracy of the estimated event signature energies E_S^{vis} and E_μ.

A way to include time over threshold information in the event likelihood has been investigated in an earlier study[224]. Inspired by this study, a new model has been derived, which is discussed in appendix I. Unfortunately, this model was not derived in time to be included in this work.

Non-hit PMTs About 50% of the computing time required to calculate the event likelihood is spent on evaluating the likelihood of non-hit PMTs. A (fraction of) these non-hit PMTs contain valuable information, since the event likelihood would not converge to zero

*One should be cautious to avoid any correlations between the used hit selection and the values of $\boldsymbol{x}_s$, in order to avoid biases and discontinuities in the event likelihood landscapes.

for large event signature energies in case the non-hit PMTs would not be used. However, most of the non-hit PMTs are located relatively far away from the event signature and are not expected to detect an L0, even if the event signature energies are significantly higher. Consequently, a significant fraction of the computing time is spent on non-hit PMTs adding only little information to the event likelihood function.

A promising way to speed up the evaluation of the non-hit event likelihood without loosing a lot of information would be to introduce a new probability density function describing the probability that not a single PMT on a DOM is hit. Using the approximation that the orientations of the PMTs within a DOM are isotropically distributed, this probability depends only on two parameters, defining the position of the DOM relative to the event signature. This probability can then be used to calculate the non-hit likelihood of all DOMs with none of its PMTs hit, while evaluating the non-hit likelihood of each individual PMT on DOMs with one or more PMTs hit are evaluated.

Another option to minimise the computing time spent on low-information non-hit PMTs would be to include only a fraction of the non-hit PMTs in the evaluation of the event likelihood. The used criterion should aim to select the non-hit PMTs providing most information to the event likelihood function, namely those with a probability of approximately 50% to be hit. The corresponding expected number of photoelectrons is $\dot{n}_i^{\text{p.e.}}(t, \boldsymbol{x}_s) = -\log(0.5) \approx 0.69$. In order to avoid biases and discontinuities in the event likelihood landscapes, the used selection criterion should be such that there is no correlation between the used hit selection and the values of $\boldsymbol{x}_s$. A few suggestions for such a non-hit PMT selection are:

- Use non-hit PMTs only in those DOMs where at least one of the other PMTs did record a hit.

- Use the most likely shower position and/or muon track parameters to select only non-hit PMTs within a certain distance to the shower/muon.

- Selecting only non-hit PMTs with $\dot{n}_i^{\text{p.e.}}(t, \tilde{\boldsymbol{x}}_s)$ larger than, say, 0.1 for the most likely shower or track parameters (provided by conventional reconstruction algorithms) and $E_S^{\text{vis.}}$ and $E_\mu = 100\text{TeV}$.

Taylor Expansion of $\log(1-\exp(-x))$. The term $\log(1-\exp(-x))$ in equation 12.15 is evaluated for each hit PMT, for each abscissa in $\boldsymbol{x}_s$. The computing time required to evaluate this term has been reduced by approximately 50% by using the c++ standard library methods `std::log1p(-std::expm1(-x)-1)`, which is (for small x) more accurate and faster to evaluate than `std::log(1-std::exp(-x))`.

Even more computing time may be saved by using a custom Taylor expansion of this function. The number of terms of the Taylor expansion can be chosen to optimise the balance between the accuracy of the result and the computing time required to evaluate the function.

14.2.5 Event Likelihood Tables

Some suggestions to improve the event likelihood tables are discussed in this section. These suggestions focus mainly on minimising the required computing time to obtain these.

Time-Integrated Event Likelihood The abscissae used to evaluate the time-integrated event likelihood (section 12.3) are chosen using the procedure described in appendix G, which takes the desired relative accuracy of the integral as input parameter. In this work, this accuracy is set to 10^{-3}, leading to an average of about 200 points in the evaluation of the integral. This number can be reduced significantly by requiring a lower accuracy.

Directional Sampling Currently, a relative course grid of directions are used to sample the full solid angle. Consequently, the region around the maximum is sampled with only a few points, while a significant amount of computing time is spent on regions in the sky that are not relevant for the test statistic integral.

A possible way to solve both drawbacks is to scan the full sky with a fine(r) grid of directions, using only a single sampled event signature position per direction. This scan provides an initial estimate of the event likelihood landscape, based on which the sampled directions for the full scan (with many sampled positions) can be chosen.

Starting Value for Perpendicular Position Sampling The scan over the perpendicular position (with respect to the sampled direction) starts from two positions: the best fit track position for the given direction and the best fit shower position (see appendix G). With these starting values, it takes on average 60 points to find the maximum of the event likelihood landscape, and 25 points to characterise the region around the found maximum. The number of points needed to find the maximum can be reduced by using the position of the found maximum for other directions close the sampled direction. Currently, this information is not used, since each sampled direction is sampled independently of the other directions to share the computational load over multiple computers.

Variable Abscissae The algorithm used to determine the abscissae of the event likelihood tables is relatively simple. Differential information of the event likelihood landscape is used to calculate the time-integrated event likelihood, but not used in the sampling of the other event signature parameters. Including this information can lead to a more efficient sampling of the event likelihood landscape, in particular the sampling of the event signature direction and position.

A second similar way to improve the sampling is to parameterise the event likelihood landscape 'on-the-fly'. Currently this is done afterwards as a cross-check of the evaluated points and to calculate the test statistic integral. This option has been considered, but the decision was made to first check what a reasonable parameterisation would be. A multivariate normal distribution seems to describe the position-dependent landscape well. For the energy landscape, no such parameterisation has been found.

14.3 Outlook

Applying the novel method to KM3NeT data would signify an important next milestone. In this section, two suggestions to achieve this milestone will be discussed.

14.3.1 Hybrid Neutrino Source Search

A promising idea to use the novel source search to identify cosmic neutrino sources would be in a hybrid approach, combining the speed of conventional neutrino source searches with the sensitivity of the novel method. In this hybrid method, conventional source search methods are used for 'standard' events (a single shower or track) described well (i.e. with good fit quality) by the event reconstruction algorithms discussed in chapter 5, while using the novel approach for particularly interesting events and events which are not used in conventional methods:

- Events with a large number of hits, but rejected by the event selection criteria. These events are potentially caused by a high-energy neutrino, but the resulting event signature is not well described by conventional reconstruction algorithms.
- Events with the highest (reconstructed) neutrino energies. These events are likely caused by neutrinos originating from the cosmos. The novel method can be used to 'squeeze' all information out of these events.
- Events with a reconstructed neutrino direction close a potential source location in the sky. Such a source could, for example, be the location of TXS or other known sources of high-energy γ-rays (see section 2.3), or a potential source indicated by the conventional neutrino source. Instead of testing a single source, one could test all sources of a particular type (for example all close-by SNRs or AGNs) simultaneously. In this case, the signal flux consists of the combined flux from all these sources.

14.3.2 Application to KM3NeT/ORCA

Identifying (and studying) cosmic neutrino sources with KM3NeT/ARCA is only one of the two main objectives of the KM3NeT calibration. The KM3NeT/ORCA detector is under construction to determine the neutrino mass ordering and the properties of neutrinos (see chapter 3).

With a few modifications, the novel method can also be used for KM3NeT/ORCA. The effects of neutrino oscillations should be included, and one additional direction should be included in the event signature parameterisation, since the assumption that the shower and track produced in charged-current muon neutrino interactions are moving in parallel is not valid at the energy range KM3NeT/ORCA is sensitive to.

A feature of the novel method is that events detected in any detector geometry can be combined in a straightforward way. Using the data from KM3NeT/ORCA with two, six, etc. deployed DUs could potentially lead to the determination of the neutrino mass hierarchy ($\mathcal{H}_0$ being normal ordering while $\mathcal{H}_1$ represents inverted order) while the full detector (115 DUs) is still under construction.

Part IV

Appendices

Abstract

In this part, additional detailed information is provided for some of the steps used in the previous parts. It also contains a list of the used terms and acronyms.

A\. Neutrino Interaction Kinematics

When a neutrino interacts, a fraction of its energy and momentum is transferred to the target. For this work, the transferred momentum is of particular interest, since it may cause a detectable shower.

The kinematics of a neutrino interacting with matter is determined by the four-momenta of the incoming (anti)neutrino k, outgoing (anti)neutrino/lepton k' and target particle p. The four-momentum transferred to the target is given by $q \equiv k' - k$. In figure 14.4, these variables are given schematically:

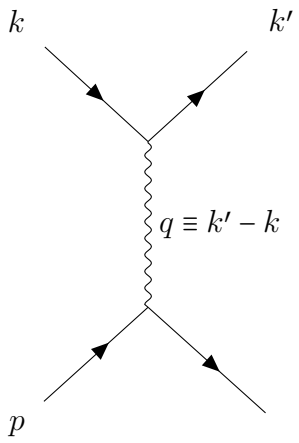


Figure 14.4

The momentum transfer onto the target can be described by two Lorentz-invariant quantities called *Bjorken-y* (y_b) and *Bjorken-x* (x_b):

$$y_b \equiv \frac{p \cdot q}{p \cdot k}, \tag{14.1}$$

$$x_b \equiv \frac{Q^2}{2p \cdot q} = \frac{-q \cdot q}{2p \cdot q}. \tag{14.2}$$

For neutrinos with energies above approximately 10 GeV, the neutrino-nucleus interaction cross-section is dominated by *deep inelastic scattering* (DIS). The leading order DIS cross-section is given by:

$$\frac{d^2\sigma}{dx_b dy_b} = \frac{1}{\pi} \cdot 2G_F^2 M_N E_\nu \cdot \left(\frac{M_{W/Z}^2}{Q^2 + M_{W/Z}^2} \right)^2 \left(x_b \cdot q(x_b, Q^2) + (1 - y_b)^2 \cdot x_b \cdot \bar{q}(x_b, Q^2) \right), \tag{14.3}$$

where G_F is the Fermi constant, M_N is the mass of the nucleon, E_ν is the neutrino energy and $M_{W/Z}$ denotes the mass of the exchanged W or Z-boson. The functions $q(x, Q^2)$ and $\bar{q}(x, Q^2)$ are called the parton distribution functions. They account for the fact that the momentum of the nucleon is shared by its three quarks, and relate to the

probability to find a quark within carrying a fraction x_b of the target nucleon for a given value of Q.

The parton distribution functions used in this work are taken from a fit[144, 145] to a series of measurements of various collaborations[138–143]. In figure 14.5, an example of the resulting double differential cross section is given.

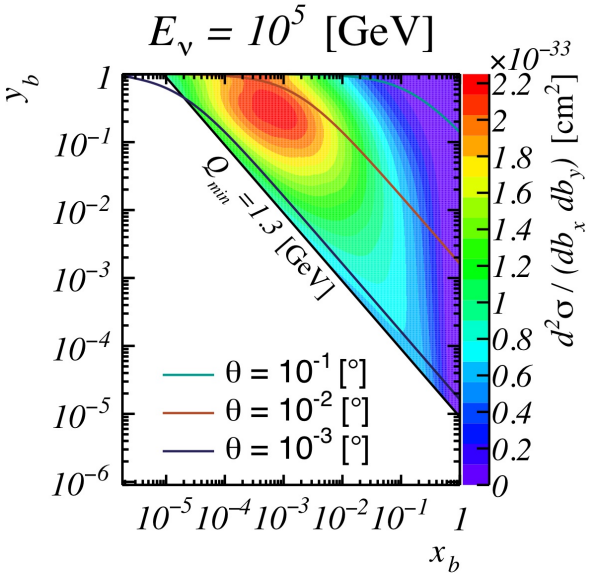


Figure 14.5: *The double differential cross-section of the interaction $\nu_\mu + n \to \mu^- + X$ for a neutrino energy of 100 TeV. The lines in the foreground correspond to different angles between the incoming neutrino and the outgoing muon in the rest frame of the target nucleon. The double differential cross-section has been calculated using the CTEQ6M parton distribution functions[145]. The lower left corner is not covered by these, and therefore left blank.*

Besides the integrated cross-section, two quantities are of interest in neutrino detectors; the fraction of neutrino energy transferred to a nucleus at rest, and the angle between the incoming neutrino and the outgoing lepton/neutrino θ in the rest frame of the target nucleus ($p = \left(m_N, \vec{0}\right)$):

$$y_b = 1 - \frac{E_l}{E_\nu}, \tag{14.4}$$

$$x_b = \frac{Q^2}{2 \cdot m_N \cdot E_\nu}, \tag{14.5}$$

$$Q^2 \equiv -q \cdot q, \tag{14.6}$$

$$= 2 \cdot E_\nu \left(E_l - |\vec{p}_l| \cdot \cos(\theta)\right) - m_l^2. \tag{14.7}$$

where natural units are used and E_l, m_l and $\vec{p}_l$ denote the energy, rest mass and momentum of the outgoing lepton/neutrino.

In figure 14.6, the differential neutrino interaction cross-section $d\sigma/dy_b$ (integrated over bjorken-x) is shown for the interaction channels at several neutrino energies. The outgoing lepton/neutrino is likely to carry a significant fraction of the energy of the interacting neutrino, which reflects in a relatively small scattering angle θ, especially at high energies.

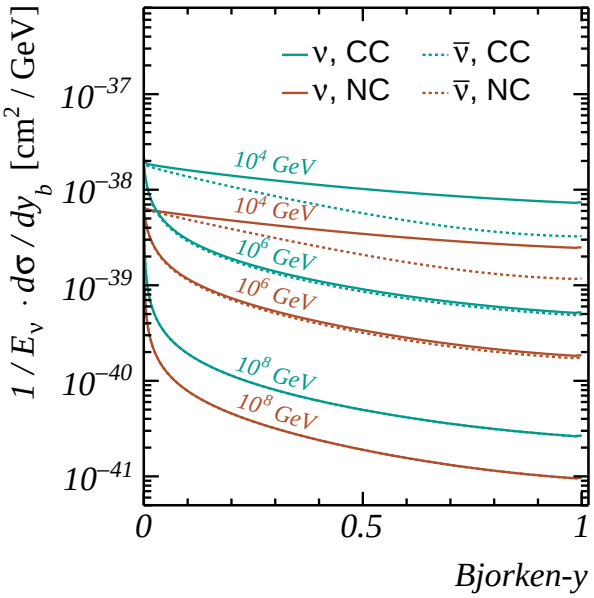


Figure 14.6: *The differential neutrino-nucleus interaction cross section as function of the fraction of neutrino energy transferred to the nucleus (Bjorken-y, or y_b) for various energies of the interacting neutrino. Each distribution is divided by this energy.*

𝔹. Singles Rate from ^{40}K Decays in Seawater

Following the argumentation given in the note by Brunner[169], an expression for the expected singles rate of photons from ^{40}K decays will be given.

The expected rate of detected single photons (the singles rate) from ^{40}K decays is given by:

$$R_{\gamma}^{40K} = r_{40K} \cdot \iint \frac{1}{4\pi \cdot L^2} \cdot 4\pi L^2 \cdot QE(\lambda) \cdot \int_{\Omega} A_{eff.}(\theta) d\Omega \cdot \frac{dN_{\gamma}}{d\lambda} \cdot exp\left[-\frac{L}{L_{abs}(\lambda)}\right] dL d\lambda, \tag{14.8}$$

$$= r_{40K} \cdot \int_{\Omega} A_{eff.}(\theta) d\Omega \cdot \int QE(\lambda) \cdot \frac{dN_{\gamma}}{d\lambda} \cdot \int_{0}^{\infty} exp\left[-\frac{L}{L_{abs}(\lambda)}\right] dL d\lambda, \tag{14.9}$$

$$= r_{40K} \cdot \int_{\Omega} A_{eff.}(\theta) d\Omega \cdot \int QE(\lambda) \cdot \frac{dN_{\gamma}}{d\lambda} \cdot L_{abs}(\lambda) d\lambda, \tag{14.10}$$

where L is the length of path before absorption, $L_{abs}(\lambda)$ is the (wavelength-dependent) absorption length of light in seawater, $A_{eff.}(\theta)$ and $QE(\lambda)$ denote the PMT effective area and quantum efficiency, respectively and r_{40K} is the expected number of ^{40}K decays per second per cubic meter seawater (13860 Bq/m^3, as obtained from measured in-situ[225].

The PMT effective area for a photon coming from a random direction has been calculated in section 3.5.1:

$$\int_{\Omega} A_{eff.}(\theta) d\Omega = 14.82 cm^2, \tag{14.11}$$

The function $\frac{dN_{\gamma}}{d\lambda}$ corresponds to the (average) number of photons emitted per unit wavelength. The spectrum of both the beta decay channel and the electron capture decay mode follow the Cherenkov spectrum. Consequently, $\frac{dN_{\gamma}}{d\lambda}$ is approximately proportional to λ^{-2}. The proportionality constant is obtained from simulations[169]:

$$\frac{dN_{\gamma}}{d\lambda} \approx 24600 \cdot \frac{1}{\lambda^2} [nm^{-1}]. \tag{14.12}$$

Evaluating the integral over the wavelength where $QE(\lambda) > 0$ yields a rate of detected photons of approximately 4.8 kHz. The total number of detected hits is about 3 % lower, as some analogue pulses do not exceed the threshold, and will not be detected.

ℂ. Event Weights

A key quantity in almost all (high-energy) physics analyses is the expected rate of events $\hat{R}$ satisfying a certain condition $\delta(\mathcal{E})$ (for example "the reconstructed energy of an event exceeds a certain threshold"). The relation between the observed parameters (denoted by $\mathcal{E}$) and the parameters x describing the interaction leading to $\mathcal{E}$ can be written as a probability density function

$$P(\mathcal{E}|x) \tag{14.13}$$

describing all relevant stochastic processes (for example the detector response). With this probability density function, $\hat{R}$ can be calculated as:

$$\hat{R} = \int \delta(\mathcal{E}) \cdot P(\mathcal{E}|x) \cdot \frac{d\Phi(x)}{dx} dx, \tag{14.14}$$

where $\frac{d\Phi(x)}{dx}$ is the expected flux of interactions.

Monte Carlo integration When working with complex detectors (like KM3NeT), it is almost impossible to derive an analytic expression for the probability $P(\mathcal{E}|x)$. Therefore, the expected rate of events $\hat{R}$ is generally calculated by means of *Monte Carlo integration* [195]:

$$\hat{R} \approx \int \frac{d\Phi(x)}{dx} dx \cdot \frac{1}{N_{gen.}} \cdot \sum_{i=1}^{N_{gen.}} \left[\delta(\mathcal{E}_i)\right], \tag{14.15}$$

where $\mathcal{E}_i$ are randomly simulated events, with interaction parameters x_i sampled from $\frac{d\Phi(x)}{dx}$. The expected event rate $\hat{R}$ is thus approximated by the simulated event rate times the fraction of the simulated events that pass the criterion.

Importance sampling In order to optimise the (computing) time spend on events that are relevant for the analysis, one often wants to simulate events from a generation flux $\frac{d\Phi^{gen.}}{dx}$ other than the (true) expected flux $\frac{d\Phi(x)}{dx}$. This technique is called *importance sampling*. In order to calculate $\hat{R}$, equation 14.15 can be written as:

$$\begin{aligned}
\hat{R} &= \int \delta(\mathcal{E}) \cdot P(\mathcal{E}|x) \cdot \frac{d\Phi}{dx}(x) dx, && (14.16)\\
&= \int \delta(\mathcal{E}) \cdot P(\mathcal{E}|x) \cdot \frac{d\Phi}{dx}(x) \cdot \frac{d\Phi^{gen.}}{dx}(x) \Big/ \frac{d\Phi^{gen.}}{dx}(x) dx\ , && (14.17)\\
&\approx \frac{1}{N_{gen.}} \cdot \sum_{i=1}^{N_{gen.}} \left[\delta(\mathcal{E}_i) \cdot \frac{d\Phi^{int.}}{dx}(x_i) \Big/ \frac{d\Phi^{gen.}}{dx}(x_i)\right] \cdot \int \frac{d\Phi^{gen.}}{dx}(x) dx, && (14.18)\\
&\approx \frac{1}{N_{gen.}} \cdot \sum_{i=1}^{N_{gen.}} \left[\delta(\mathcal{E}_i) \cdot w_i\right], && (14.19)
\end{aligned}$$

with w_i an event-specific quantity called the *event weight*:

$$w_i \equiv \frac{\frac{d\Phi^{int.}(x_i)}{dx}}{\frac{d\Phi^{gen.}}{dx}(x_i)} \cdot \int \frac{d\Phi^{gen.}}{dx}(x) dx. \tag{14.20}$$

Importance sampling in Genhen Importance sampling is used in Genhen (see section 4.3.3) to simulate more events with high neutrino energy, using a generation spectrum:

$$\frac{d^6\Phi^{gen.}}{dE_\nu d^2\Omega d^3\vec{x}} \propto E_\nu^{-\alpha}, \tag{14.21}$$

with a spectral index α generally set to -1.4 and normalisation

$$\int_{E_{min.}}^{E_{max.}} \frac{d^6\Phi^{gen.}}{dEd^2\Omega d^3\vec{x}} dE_\nu d^2\Omega d^3\vec{x} = V_{gen.} \cdot 4\pi \cdot \frac{(E_{max.})^{1-\alpha} - (E_{min.})^{1-\alpha}}{1-\alpha}, \tag{14.22}$$

where $E_{min.}$ and $E_{max.}$ give the bounds of the energy range used in the generation flux.

The event weight w_i of an arbitrary flux $\frac{d^3\Phi^{arb.}}{dEd^2\Omega}$ of neutrinos at the Earth's atmosphere can be calculated using equation 3.1:

$$w_i \equiv w_i' \cdot \frac{d^3\Phi^{arb.}}{dEd^2\Omega}, \tag{14.23}$$

$$\begin{aligned} w_i' = V_{gen.} \cdot 4\pi \cdot \frac{(E_{max.})^{1-\alpha} - (E_{min.})^{1-\alpha}}{1-\alpha} \cdot (E_{\nu,i})^\alpha \\ \times \sigma(E_{\nu,i}) \cdot N_A \cdot \rho(\vec{x}_i) \cdot P_\oplus(E_{\nu,i}, \hat{d}_{\nu,i}). \end{aligned} \tag{14.24}$$

For each event, the value of w_i' is stored, providing a convenient way to calculate $\hat{R}$ for $\frac{d^3\Phi^{arb.}}{dEd^2\Omega}$:

$$\hat{R} \approx \frac{1}{N_{gen.}} \cdot \sum_{i=1}^{N_{gen.}} \left[\delta(ev_i) \cdot w_i' \cdot \frac{d^3\Phi^{arb.}}{dEd^2\Omega}(E, \hat{d}) \right]. \tag{14.25}$$

𝔻. Time-Dependent Number of Photoelectrons

The evaluation of the time-dependent number of photoelectrons on a PMT as expected from an event signature is discussed in this appendix. These functions are used in simulations (chapter 4) and in event reconstruction (chapters 5 and 11).

Definition An incident photon on the photocathode of a PMT can cause the emission of a photoelectron through the photoelectric effect. The corresponding probability $\mathrm{QE}(\lambda)$, called the quantum efficiency, depends on the wavelength λ of the photon. The time-dependent expected number of photoelectrons on a PMT i located at $\vec{x}_i$ and pointing towards $\hat{d}_i$ is given by:

$$\begin{aligned}\dot{n}_i(t, \boldsymbol{x}_s) &\equiv \frac{\partial n^{\text{p.e.}}(\vec{x}_i, \hat{d}_i, t)}{\partial t} \\ &= \iiint_{4\pi} \phi^\gamma(\lambda, \vec{x}, \hat{d}, t) \cdot \mathrm{QE}(\lambda) \cdot A_{eff.}(\hat{d} \cdot \hat{d}_i) d^2\hat{d} d\lambda,\end{aligned} \tag{14.26}$$

where $\boldsymbol{x}_s$ denote the parameters that describe the source of the photons and A_{eff} is the effective area of the PMT. The effective area depends on the angle between the photon direction $\hat{d}$ and the direction of the PMT $\hat{d}_i$. The quantum efficiency and direction-dependent effective area of a nominal PMT are given in figures 3.7 (right) and 3.7 (left).

The expected photon flux $\phi^\gamma(\lambda, \vec{x}, \hat{d}, t)$ is defined as the number of photons per unit time, wavelength and solid angle passing through a surface (located at $\vec{x}$, perpendicular to $\hat{d}$) with unit area at time t:

$$\phi^\gamma(\lambda, \vec{x}, \hat{d}, t) \equiv \frac{\partial^6 N^\gamma(\lambda, \vec{x}, \hat{d}, t)}{\partial\lambda \partial t \partial^2 A \partial^2 \hat{d}}. \tag{14.27}$$

In the remainder of this appendix, the arguments of ϕ^γ will not be denoted explicitly.

The Transport Equation The photon flux ϕ^γ at any $\vec{x}$, $\hat{d}$ and t depends on the photon flux of the source $\phi^{\gamma,\text{source}}(\lambda, \vec{x}, \hat{d}, t|\boldsymbol{x}_s)$ via the transport equation:

$$\begin{aligned}&\frac{n_g(\lambda)}{c} \cdot \phi^\gamma + \hat{d} \cdot \nabla\phi^\gamma + \left(\frac{1}{l_{\text{abs.}}(\lambda)} + \frac{1}{l_{\text{scat.}}(\lambda)}\right) \cdot \phi^\gamma \\ &\quad = \frac{1}{l_{\text{scat.}}(\lambda)} \iint_{4\pi} \phi^\gamma(\hat{d}') \cdot P_s(\hat{d}' \cdot \hat{d}) d\hat{d}' + \frac{n_g(\lambda)}{c} \cdot \phi^{\gamma,\text{source}}(\boldsymbol{x}_s),\end{aligned} \tag{14.28}$$

where c is the speed of light, n_g is the group refractive index of light in seawater, $l_{\text{abs.}}$ is the absorption length, $l_{\text{scat.}}$ the scattering length, $P_s(\hat{d}' \cdot \hat{d})$ denotes the scattering angle probability density function.

Light Emission The charged particles produced in the interactions of neutrinos induce light through Cherenkov radiation. The spectrum of the emitted light $\frac{d^2N^\gamma}{dzd\lambda}$ is given by equation 3.3. The light emission profiles of neutrino-induced event signatures can be factorised in the angular emission profile $\frac{d^2P}{d\cos(\theta_{\text{emit}})d\phi_{\text{emit}}}$ and the longitudinal emission

profile, parameterised by the deposited energy $\frac{dE}{dz}(z)$ per unit length along the signature direction $\hat{d}_s$ and an equivalent muon track length L_{eff} per unit deposited energy:

$$\begin{aligned}\phi^{\gamma,\text{source}}(\boldsymbol{x}_s) = \frac{d^2N^\gamma}{dzd\lambda} \cdot L_{eff} \cdot \left|\frac{dE}{dz}(z)\right| \cdot \frac{d^2P}{d\cos(\theta_{\text{emit}})\, d\phi_{\text{emit}}} \\ \times \delta\big(t_{\text{emit}} = \frac{|\vec{x}_{\text{emit}} - \vec{x}_s|}{c}\big) \cdot \delta\left((\vec{x}_{\text{emit}} - \vec{x}_s) \cdot \hat{d}_s = |\vec{x}_{\text{emit}} - \vec{x}_s|\right),\end{aligned} \tag{14.29}$$

Where the δ's denote that all light emitting particles travel with the speed of light along the event signature axis.

For each of the event signatures discussed in section 3.2.5, the functions that characterise the light emission profile are summarised in table 14.1. For a muon, $L_{eff} \cdot \left|\frac{dE}{dz}(z)\right| \equiv 1$ (by definition). Showers are described by the energy of the initial particle (called the shower energy E_S) and the fraction F_S^{vis} of the shower energy that leads to the production of light. For electromagnetic showers, $F_S^{vis} = 1$.

Process		Photon flux characteristics at source	
Muon, Cherenkov	μ, C	L_{eff}	1 [m/GeV]
		$\left\|\frac{dE}{dz}\right\|$	1 [GeV/m]
		$\frac{d^2P}{d\cos\theta_{\text{emit}} d\phi_{\text{emit}}}$	$\frac{\delta(\cos(\theta_{\text{emit}})=\cos(\theta_C))}{2\pi\cdot\sin(\theta_C)}$
Muon, δ-rays	μ, δ	L_{eff}	4.0 [m/GeV]
		$\left\|\frac{dE}{dz}(z)\right\|$	$\frac{dE_\delta}{dz}(E_\mu(z))$ [GeV/m]
		$\frac{d^2P}{d\cos\theta_{\text{emit}} d\phi_{\text{emit}}}$	$\frac{1}{4\pi}$
Single EM or hadronic shower	S	L_{eff}	$F_S^{vis} \cdot 4.0$ [m/GeV]
		$\left\|\frac{dE}{dz}(z)\right\|$	$E_S \cdot z^{a-1} \cdot e^{-z/b} \cdot b^{-a}/\Gamma(a)$ [GeV/m]
			$a = 1.85 + 0.62 \cdot \ln\left(\frac{E_S}{[\text{GeV}]}\right)$, $b = 0.54$ [m]
		$\frac{d^2P}{d\cos\theta_{\text{emit}} d\phi_{\text{emit}}}$	$\frac{\exp\left(-5.40\times\left\|\cos(\theta_{\text{emit}})-\cos(\theta_C)\right\|^{0.35}\right)}{2\pi\int_{-1}^{1}\exp\left(-5.40\times\left\|\cos(\theta_{\text{emit}})-\cos(\theta_C)\right\|^{0.35}\right)d\cos(\theta_{\text{emit}})}$
EM showers along muon	μ, B	L_{eff}	4.0 [m/GeV]
		$\left\|\frac{dE}{dz}(z)\right\|$	$B \cdot E_\mu(z)$ [GeV/m], $B = 3.53 \cdot 10^{-8}$ [m^{-1}]
		$\frac{d^2P}{d\cos\theta_{\text{emit}} d\phi_{\text{emit}}}$	$\frac{\exp\left(-5.40\times\left\|\cos(\theta_{\text{emit}})-\cos(\theta_C)\right\|^{0.35}\right)}{2\pi\int_{-1}^{1}\exp\left(-5.40\times\left\|\cos(\theta_{\text{emit}})-\cos(\theta_C)\right\|^{0.35}\right)d\cos(\theta_{\text{emit}})}$

Table 14.1: *An overview of the characteristic light emission profiles of muons (Cherenkov radiation and δ-rays) and showers. The emission profile given in the bottom row does not represent a physical source, but is used in event reconstruction (see text).*

Coordinate System The time-dependent number of expected photoelectrons depends on the orientation and the position of the PMT with respect to the position $\vec{x}_s$ and direction $\hat{d}_s$ of a (segment of a) particle. A convenient set of coordinates to parameterise the photon

flux is given by (see figure 14.7):

$$\begin{aligned}
\Delta t &\equiv t - t_s, \\
\Delta z_i &\equiv (\vec{x}_i - \vec{x}_s) \cdot \hat{d}_s, \\
\rho_i &\equiv \sqrt{(\vec{x}_i - \vec{x}_s)^2 - \Delta z_i^2}, \\
\theta_i &\equiv \arccos\left(\hat{d}_i \cdot \hat{d}_s\right), \\
\phi_i &\equiv \arccos\left(\frac{\left(\vec{x}_i - \vec{x}_s - \Delta z_i \cdot \hat{d}_s\right)}{\rho_i} \cdot \left(\hat{d}_i - \cos(\theta_i) \cdot \hat{d}_s\right)\right),
\end{aligned} \tag{14.30}$$

where $\vec{x}_ß$ and $\hat{d}_i$ denote the position and pointing direction of PMT i. An additional useful parameter gives the point along the trajectory from where an unscattered photon is emitted in order to be detected by PMT i:

$$\Delta z_{i,C} \equiv \Delta z_i - \frac{\rho_i}{\tan(\theta_C)}, \tag{14.31}$$

with θ_C the Cherenkov angle.

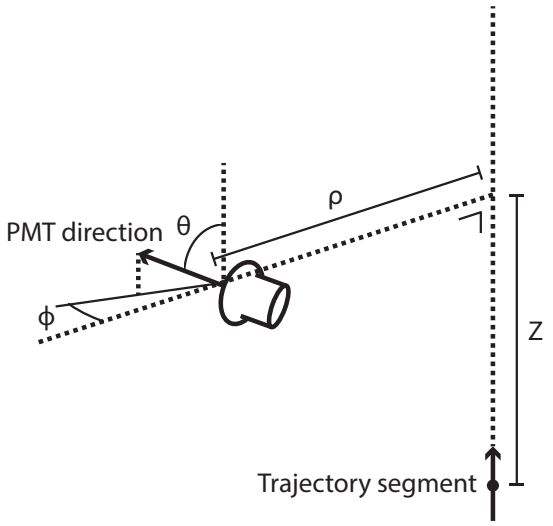


Figure 14.7: *The coordinate system used in KM3NeT light simulation software to describe the position and direction of a PMT with respect to a track segment. The dashed lines lie in one plane, and are parallel and perpendicular to each other.*

Semi-analytic solution An analytic solution of the transport equation is known only for some (simple) sources. For the relevant light emitting processes in the KM3NeT detectors, no such solution is known. Nevertheless, a semi-analytic expression for the time-dependent expected number of photoelectrons $\dot{n}_i(t)$ can be calculated in case light does not scatter or scatters only once between the source and the PMT [192]. The use of these semi-analytic solutions is supported by preliminary studies, which indicate that the contribution of multiple scattered photons to the total photon flux is relatively small [226].

In figure 14.8 (left), the semi-analytic solutions of the time-dependent number of expected photoelectrons for Cherenkov light from a muon is given for a PMT at several

distances and orientations. The main peak of the time distribution widens with increasing distance to the muon track, which can be explained by the increased contribution of scattered light. PMTs pointing away from the muon detect scattered light only (figure 14.8 (right)), resulting in fewer expected photons with respect to a PMT pointing towards the muon.

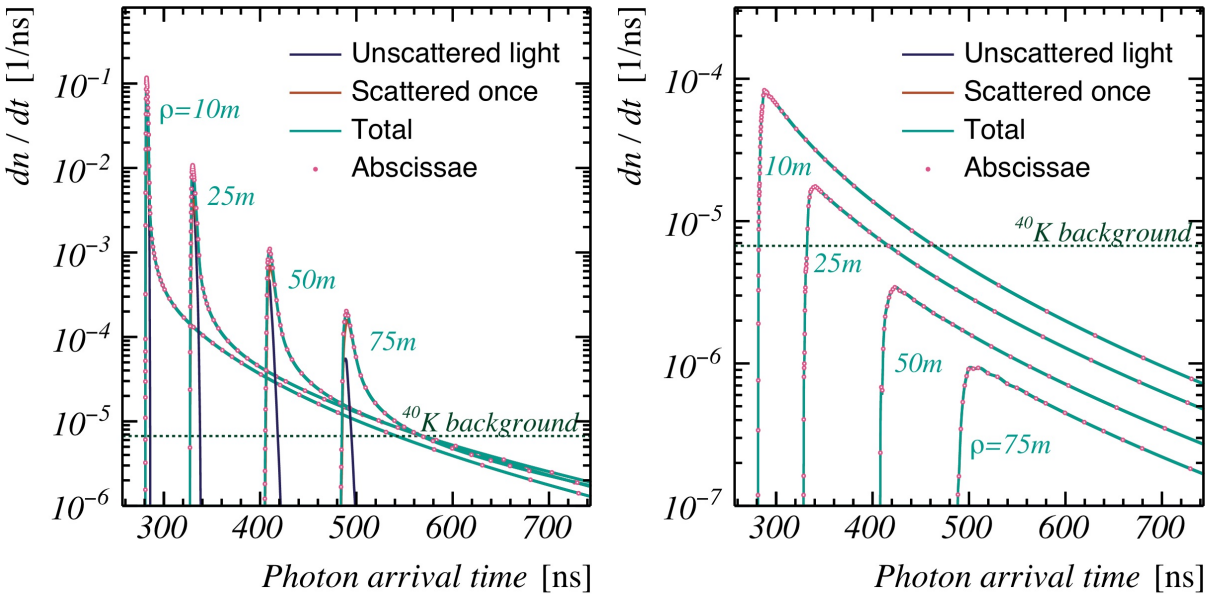


Figure 14.8: ***Left:*** *The expected number of photoelectrons from Cherenkov radiation for a PMT at several distances from a muon with infinite length. The PMT is oriented straight at the muon track ($\theta = 90°$, $\phi = 180°$ as in figure 14.7). Given are the contribution from unscattered light and light that scattered once. For comparison, the predicted number of photoelectrons due to the background of ^{40}K decays is given as well. The abscissae used for the evaluation of the function (see text) are indicated by the points.* ***Right:*** *Same as left figure, but for a PMT pointed away from the muon ($\theta = 90°$, $\phi = 0°$).*

Tables

Evaluating the semi-analytic solution for the time-dependent expected number of photoelectrons requires too much computing time for direct use in simulation and reconstruction algorithms. Precalculated tables are used instead, which contain $\dot{n}_i(t, \boldsymbol{x}_s)$ at a large set of abscissae in Δt, ρ, Δz, θ and ϕ. Two tables are stored for each of the emission profiles given in table 14.1, one for unscattered light and one for scattered light, differing mainly in the sampling of the time difference Δt.

In order to account for the spread in the transit times of the PMTs, a Gauss-Legendre integral is used to convolute each $\dot{n}_{S,i}(t|\boldsymbol{x}_s)$ distribution with a Normal distribution. This 'smearing' is done at the start of the reconstruction algorithms. In this step, the $\dot{n}_{S,i}(t|\boldsymbol{x}_s)$ distributions of scattered and unscattered light are added. The simulations include a detailed model of the PMT transit time distribution, and no 'smearing' is required. Scattered and unscattered light is treated separately in simulations.

Several simplifications make the tables faster to evaluate. In the following, these will be discussed.

Cherenkov radiation from muons A muon propagating through seawater induces Cherenkov radiation (with constant dE/dz) as long as its energy $E_\mu(\Delta z)$ exceeds the Cherenkov threshold (158 MeV). Muons of interest for KM3NeT/ARCA typically have a high energy, and therefore a long range (up to tens of kilometres). By assuming that the muon has an infinite range, two fewer parameters (the initial muon energy E_μ and position of the PMT along the muon track Δz) are needed to parameterise $\dot{n}_i$. With this assumption, $\dot{n}_i$ depends only on the time t, the orientation of the PMT (θ_i and ϕ_i) and the distance from the PMT to the muon (ρ_i).

The finite range of the muon is implemented by requiring that a PMT detects Cherenkov radiation only in case the muon energy $E_\mu(\Delta z_{i,C}) \geq 158$ MeV at the point $\Delta z_{i,C}$ (see equation 14.31). With this condition, the expected number of photoelectrons from Cherenkov radiation of muons can be written as:

$$\dot{n}_{\mu,i}^{\text{Cher.}}(t|\boldsymbol{x}_s) = \delta(E_\mu(\Delta z_{i,C}) \geq 158\text{MeV}) \cdot \dot{n}_\mu^{\text{Cher.}}(t - t_s, \rho_i, \theta_i, \phi_i), \tag{14.32}$$

where $\dot{n}_\mu^{\text{Cher.}}(t - t_s, \rho_i, \theta_i, \phi_i)$ is stored in the tables, and the delta function δ denotes the condition on the muon energy.

δ-Rays from muons Photoelectrons originating from photons induced by δ-rays along a muon trajectory are treated in the same way as Cherenkov radiation. The muon is assumed to have infinite range, with constant emission of δ-rays along its trajectory. A parameterisation is used to calculate the energy loss of the muon due to δ-rays at the point $\Delta z_{i,C}$. This parameterisation is given in equation 3.11, and is combined with the tables to give the expected number of photoelectrons from δ-rays:

$$\dot{n}_{\mu,i}^{\delta\text{-rays}}(t|\boldsymbol{x}_s) = \frac{dE}{dz}_{\delta}(E_\mu(\Delta z_{i,C})) \cdot \dot{n}_\mu^{\delta\text{-rays}}(t - t_s, \rho_i, \theta_i, \phi_i), \tag{14.33}$$

Single electromagnetic and hadronic showers The tables used to simulate light emission by a shower correspond to the expected number of photoelectrons from an electromagnetic shower ($F_S^{\text{vis.}} = 1$) with a shower energy of 1 GeV, assuming that all light is emitted from a single point at $\vec{x}_s$. The tabulated function ($\dot{n}_S^{1\,\text{GeV}}$) depends on five parameters (t, ρ, Δz, θ and ϕ) and can be used to calculate the expected number of photoelectrons from the shower:

$$\dot{n}_{S,i}(t|\boldsymbol{x}_s) = F_S^{vis} \cdot \int_0^\infty \left|\frac{dE}{dz}(z)\right| \cdot \dot{n}_S^{1\,\text{GeV}}(t - t_s, \rho_i, \Delta z_i(z), \theta_i, \phi_i)dz, \tag{14.34}$$

where $\frac{dE}{dz}(z)$ is the longitudinal energy loss distribution of a shower with shower energy E_S (given in table 14.1). Monte Carlo sampling is used in simulations to sample the emission points z and detection times t_i of detected photoelectrons from this integral.

In order to speed up the reconstruction algorithms (significantly), the light induced by both electromagnetic and hadronic showers is assumed to be emitted from a single point. This is a reasonable assumption, since the typical spacing between the DOMs is

about 36 m, while the longitudinal emission profile of a shower is a few metres. Under this assumption, the integral (equation 14.34) simplifies to:

$$\dot{n}_{S,i}(t|\boldsymbol{x}_s) = E_S^{\text{vis.}} \cdot \dot{n}_S^{1\text{ GeV}}(t - t_s, \rho_i, \Delta z_i, \theta_i, \phi_i). \tag{14.35}$$

Stochastic electromagnetic showers along muon tracks A muon induces electromagnetic showers along its trajectory due to bremsstrahlung, e^+e^- pair-production and photonuclear reactions. Each shower can have a significant fraction of the muon energy. In simulations, each of these showers is treated individually, according to the procedure described in the previous paragraph.

In (current) reconstruction algorithms, the properties of each of the stochastic electromagnetic showers are not estimated. Instead, a separate table is used, which describes the expected number of photoelectrons from a high-energy muon (the B-term in equation 3.13) per unit energy $\dot{n}_{\mu,\text{bs,pp\&pn}}^{1\text{ GeV}}$. Similar to the treatment of δ-rays, the muon is assumed to have an infinite range. This table, which depends on t, ρ, θ and ϕ, is used to calculate the expected number of photoelectrons according to:

$$\dot{n}_{\mu,\text{bs,pp\&pn},i} = E_\mu(\Delta z_{i,C}) \cdot \dot{n}_{\mu,\text{bs,pp\&pn}}^{1\text{ GeV}}(t - t_s, \rho_i, \theta_i, \phi_i). \tag{14.36}$$

Interpolation Multi-dimensional interpolation techniques (implemented in the JPP software package by De Jong[193] based on the algorithms discussed in the book by Press et al.[217]) are used for fast evaluation of the following quantities in one function call:

$$\begin{cases} n_i(t) & \equiv \int_{-\infty}^{t} \dot{n}_i(t)dt, \\ \dot{n}_i(t) & \equiv \frac{dn_i}{dt}(t), \\ \ddot{n}_i(t) & \equiv \frac{d^2n_i}{dt^2}(t), \\ N_i & \equiv n_i(\infty) = \int_{-\infty}^{\infty} \dot{n}_i(t)dt. \end{cases} \tag{14.37}$$

𝔼. Conceptual Proof of the Neyman-Pearson Lemma

The fact that the Neyman-Pearson test is the most powerful (i.e. has the lowest probability β of a type II error) at a given significance level α, can be proven *reductio ad impossible*[195]. A simplified version of this proof will be given in this appendix.

1) The acceptance region of any alternative hypothesis test (A) must include λ-values in one (or more) infinitesimal bin $\Delta\lambda_A$ centred around a value $\lambda_A < \lambda_{crit.,\mathrm{NP}}$, where $\lambda_{crit.,\mathrm{NP}}$ defines the acceptance region of the Neyman-Pearson lemma. Note that λ_A has to be smaller than $\lambda_{crit.,\mathrm{NP}}$, because otherwise it would share the same acceptance region as the Neyman-Pearson lemma, and would therefore not be an alternative.

2) Since the significance level of the alternative criterion has to be the same as the Neyman-Pearson criterion, including $\lambda \in \Delta\lambda_A$ means that some region $\Delta\lambda_{NP}$ centred around $\lambda_{NP} > \lambda_{crit.,\mathrm{NP}}$ is not included in the acceptance region of the alternative test. The 'width' of this region follows from equation 10.3:

$$\Delta\lambda_\mathrm{A} \cdot \left.\frac{dP(\lambda|\mathcal{H}_0)}{d\lambda}\right|_{\lambda_\mathrm{A}} = \Delta\lambda_\mathrm{NP} \cdot \left.\frac{dP(\lambda|\mathcal{H}_0)}{d\lambda}\right|_{\lambda_\mathrm{NP}}, \tag{14.38}$$

3) The probability β_A to make a type II error using the alternative test differs from β_NP under the Neyman-Pearson lemma by:

$$\beta_A - \beta_\mathrm{NP} = \left.\frac{dP(\lambda|\mathcal{H}_1)}{d\lambda}\right|_{\lambda_\mathrm{NP}} \cdot \Delta\lambda_\mathrm{NP} - \left.\frac{dP(\lambda|\mathcal{H}_1)}{d\lambda}\right|_{\lambda_\mathrm{A}} \cdot \Delta\lambda_\mathrm{A}. \tag{14.39}$$

Using the relation obtained in the previous point:

$$\frac{\beta_A - \beta_\mathrm{NP}}{\Delta\lambda_\mathrm{NP}} = \left.\frac{dP(\lambda|\mathcal{H}_1)}{d\lambda}\right|_{\lambda_\mathrm{NP}} - \left.\frac{dP(\lambda|\mathcal{H}_1)/d\lambda}{dP(\lambda|\mathcal{H}_0)/d\lambda}\right|_{\lambda_\mathrm{A}} \cdot \left.\frac{dP(\lambda|\mathcal{H}_0)}{d\lambda}\right|_{\lambda_\mathrm{NP}}, \tag{14.40}$$

$$= \left.\frac{dP(\lambda|\mathcal{H}_1)/d\lambda}{dP(\lambda|\mathcal{H}_0)/d\lambda}\right|_{\lambda_\mathrm{NP}} \cdot \left.\frac{dP(\lambda|\mathcal{H}_0)}{d\lambda}\right|_{\lambda_\mathrm{NP}} - \left.\frac{dP(\lambda|\mathcal{H}_1)/d\lambda}{dP(\lambda|\mathcal{H}_0)/d\lambda}\right|_{\lambda_\mathrm{A}} \cdot \left.\frac{dP(\lambda|\mathcal{H}_0)}{d\lambda}\right|_{\lambda_\mathrm{NP}}. \tag{14.41}$$

4) An important step in this proof is to realise that $\frac{dP(\lambda|\mathcal{H}_1)/d\lambda}{dP(\lambda|\mathcal{H}_0)/d\lambda}$ is proportional to the test statistic value λ at which it is evaluated. This follows from the definition of the likelihood ratio test statistic:

$$\lambda(\mathcal{D}) \equiv \frac{\mathcal{L}(\mathcal{D}|\mathcal{H}_1 = \mathrm{true})}{\mathcal{L}(\mathcal{D}|\mathcal{H}_0 = \mathrm{true})} \propto \frac{P(\mathcal{D}|\mathcal{H}_1 = \mathrm{true})}{P(\mathcal{D}|\mathcal{H}_0 = \mathrm{true})} \propto \frac{dP(\lambda|\mathcal{H}_1)/d\lambda}{dP(\lambda|\mathcal{H}_0)/d\lambda}. \tag{14.42}$$

5) Combining steps 3) and 4) gives:

$$\frac{\beta_A - \beta_\mathrm{NP}}{\Delta\lambda_\mathrm{NP}} = (\lambda_\mathrm{NP} - \lambda_\mathrm{A}) \cdot \left.\frac{dP(\lambda|\mathcal{H}_0)}{d\lambda}\right|_{\lambda_\mathrm{NP}}. \tag{14.43}$$

6) For any alternative hypothesis test, $\lambda_A < \lambda_{crit.,\mathrm{NP}} < \lambda_B$, which gives (since $\Delta\lambda_\mathrm{A}$ and $\frac{dP(\lambda|\mathcal{H}_0)}{d\lambda}$ are positive) $\beta_A > \beta_\mathrm{NP}$. In other words, there is no acceptance region with a better power $(1-\beta)$ than the Neyman-Pearson likelihood ratio test.

𝔽. The Preliminary Reference Earth Model (PREM)

The preliminary reference earth model is obtained from seismic data and is based on the assumption that the Earth consists of a set of spherical regions. The density profile of each region is parameterised as a third order polynomial[227]:

$$\rho_{\oplus,i}(R) = a_i \cdot \left(\frac{R}{R_\oplus}\right)^3 + b_i \cdot \left(\frac{R}{R_\oplus}\right)^2 + c_i \cdot \frac{R}{R_\oplus} + d_i, \qquad R_i^{min.} \leq R < R_i^{max.} \tag{14.44}$$

where R denotes the distance to the centre of the Earth, $R_i^{min.}$ and $R_i^{max.}$ are the inner and outer radius of region i and $R_\oplus$ is the radius of the Earth (6378 km). The characteristic parameters of the PREM are given in table 14.2.

Region	Depth [km]	$R_i^{min.}$ - $R_i^{max.}$ [km]	a	b	c	d
			[g/cm³]			
Seawater	0 - 3.5	6375.5 - $R_\oplus$	0	0	0	1.038
Rock	3.5 - 12	6356 - 6375.5	0	0	0	2.65
Upper mantle	12 - 31.4	6346.6 - 6356	0	0	0	2.90
	31.4 - 227	6151 - 6346.6	0	0	0.6924	2.6910
	227 - 407	5971 - 6151	0	0	−3.8045	7.1089
Transition zone	407 - 607	5771 - 5971	0	0	−8.0298	11.2494
	607 - 677	5701 - 5771	0	0	−1.4836	5.3197
Lower mantle	677 - 2898	3480 - 5701	3.0807	5.5283	−6.4761	7.9565
Outer core	2898 - 5156.5	1221.5 - 3480	−5.5281	−3.6426	−1.2638	12.5815
Inner core	5156.5 - 6378	0 - 1221.5	0	−8.8381	0	13.0885

Table 14.2: *Parameterisation of the Earths' density profile as given by the PREM[227]. Each row corresponds to a spherical region in the Earths' interior (see text).*

It is straightforward to calculate the total column density $\rho_{\oplus,tot.}(\theta_\nu)$ for a neutrino traversing the Earth under a zenith angle θ_ν. Since the PREM consists of several regions, the total column density is given by the sum of the column density of each of the regions j the neutrino passes through (see figure 14.9 (left)):

$$\rho_{\oplus,tot.}(\theta_\nu) = \sum_i \left[\rho_{\oplus,tot.,i}(\theta_\nu)\right], \tag{14.45}$$

$$\rho_{\oplus,tot.,i}(\theta_\nu) = 2 \cdot \int_{R_i^A}^{R_i^{max}} \rho_\oplus(R) \frac{dz}{dR} dR, \tag{14.46}$$

with R_i^{max} the outer radius of shell i and

$$z(R) \equiv \sqrt{R^2 - R_\oplus^2 \cdot \sin^2(\theta_\nu)}, \tag{14.47}$$

$$R_i^A \equiv \begin{cases} R_i^{min} & \text{if } R_i^{min} > R_\oplus \cdot \sin(\theta_\nu) \\ R_\oplus \cdot \sin(\theta_\nu) & \text{if } R_i^{min} \leq R_\oplus \cdot \sin(\theta_\nu) < R_i^{max} \\ R_i^{max} & \text{if } R_\oplus \cdot \sin(\theta_\nu) > R_i^{max} \end{cases}. \tag{14.48}$$

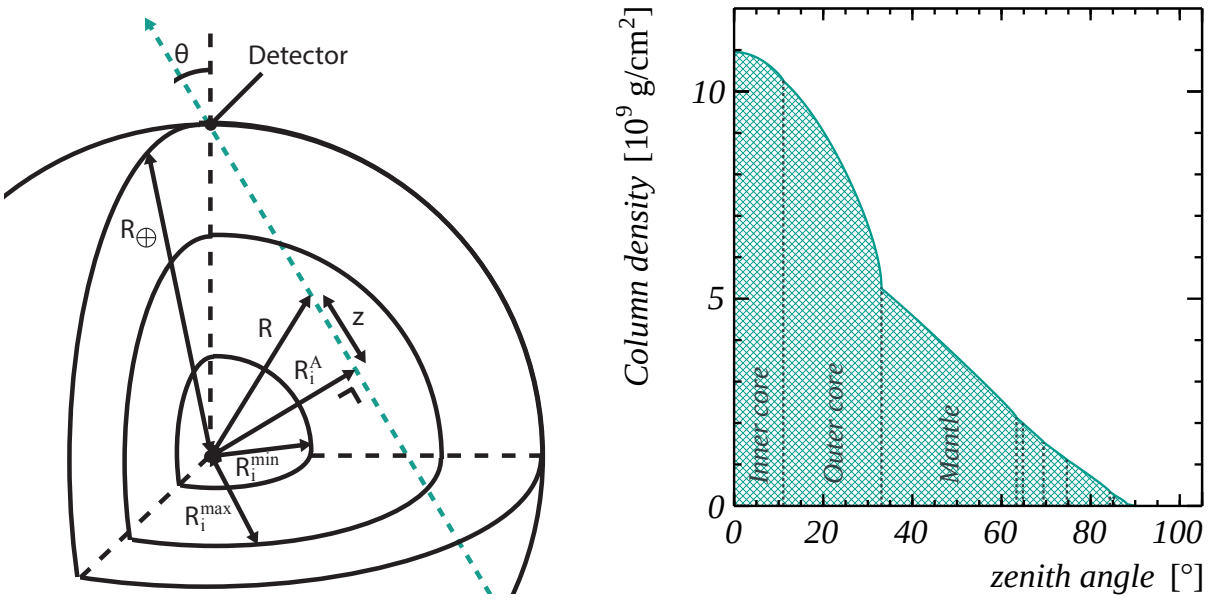


Figure 14.9: ***Left:*** *An illustration of the coordinates used in the calculation of the traversed matter density of the Earth. Each circle corresponds to a certain region in the Earth, the dashed arrow indicates the path of a neutrino passing through the detector with a zenith angle θ.* ***Right:*** *The traversed matter density as given by the PREM[227]. Dashed vertical lines correspond to the edges of the regions. A zenith angle of 0 degrees corresponds to a neutrino going straight up. In this figure, the water overburden and the density profile in the vicinity of the detector are not taken into account.*

The integral can be solved analytically:

$$\begin{aligned}\rho_{\oplus,tot.,i}(\theta_\nu) = \Bigg[& z(R)\cdot\left(2d_i + c_i\left(\frac{R}{R_\oplus}\right) + \frac{2}{3}b_i\left(\frac{R}{R_\oplus}\right)^2 + \frac{1}{2}a_i\left(\frac{R}{R_\oplus}\right)^3\right) \\ & + z(R)\cdot\sin^2(\theta_\nu)\cdot\left(\frac{4}{3}b_i + \frac{3}{4}a_i\left(\frac{R}{R_\oplus}\right)\right) \\ & + R_\oplus\cdot\sin^2(\theta_\nu)\cdot\left(c_i + \frac{3}{4}a_i\sin^2(\theta_\nu)\right)\cdot\log\left(R + z(R)\right)\Bigg]\Bigg|_{R_i^A}^{R_i^{max}} ,\end{aligned} \tag{14.49}$$

𝔊. Sampling Event Likelihood Landscapes

Order of Scanned Parameters

The main factor influencing the choice of the abscissae is the total computing time needed to sample the event likelihood landscape. The time per evaluated $\boldsymbol{x}_s$ abscissae is reduced with several orders of magnitude by sampling the parameters of $\boldsymbol{x}_s$ in a particular order, making use of the fact that some quantities do not depend on all parameters of $\boldsymbol{x}_s$:

1) For a given event signature position and direction, the evaluation of the event likelihood consists of two main steps (see section 12.1). First, for each hit PMT, the time-dependent expected number of photoelectrons $\dot{n}_{S,1\text{ GeV}}$, $\dot{n}_{\mu,C}$, $\dot{n}_{\mu,B}$, $\dot{n}_{\mu,\delta}$, and for each non-hit PMT the time-integrated expected number of photoelectrons $N_{S,1\text{ GeV}}$, $N_{\mu,C}$, $N_{\mu,B}$ and $N_{\mu,\delta}$ are evaluated. These numbers are used in equations 12.7 and 12.15 to calculate the event likelihood for a given muon energy E_μ and visible shower energy E_S^{vis}. The first step is (by far) the most time-consuming, in particular the identification of the right ordinate values in the stored tables ($2^3 = 8$ abscissae for each three-dimensional $N_{\mu,\dots}$, $2^4 = 16$ for $N_{S,1\text{ GeV}}$ and $\dot{n}_{\mu,\dots}$ and $2^5 = 32$ abscissae for $\dot{n}_{S,1\text{ GeV}}$).

2) The abscissae in the time domain $t - t_s$ that define $\dot{n}_{\dots}(t_s)$ for each hit PMT, are copied to the RAM. These abscissae describe the one-dimensional time-dependent expected number of photoelectrons, which can are used to calculate $\dot{n}_{S,1\text{ GeV}}$, $\dot{n}_{\mu,C}$, $\dot{n}_{\mu,B}$, $\dot{n}_{\mu,\delta}$ requiring the identification of only 2 abscissae each. Even though copying the abscissae to the RAM is a relatively CPU-intensive step, this procedure is still relatively fast, since the events signature time t_s is typically sampled with tens of abscissae for each $\vec{x}_s$ and $\hat{d}_s$.

3) The evaluation of the time-dependent number of photoelectrons $\dot{n}_{\dots}$ at a certain time t_s requires approximately the same computing time as needed for the calculation of the event likelihood at a certain E_μ and E_S^{vis}. Therefore, the calculated $\dot{n}_{\dots}$ for a given $t_{s,i}$ are used to calculate l_i and $\dot{l}_i$ for a set of energies E_μ and E_S^{vis}. The scan over the signature energies E_S^{vis} and E_μ is split into a scan over E_μ, in which $\dot{n}_{\mu,i}(t|\boldsymbol{x}_s)$ and $N_{\mu,i}(t|\boldsymbol{x}_s)$ are calculated, followed by a scan over E_S^{vis} in which the remaining parameters leading to $l(\mathcal{E}|\boldsymbol{x}_s)$ are calculated.

4) The fact that $\dot{n}_{\mu,C}$, $\dot{n}_{\mu,B}$ and $\dot{n}_{\mu,\delta}$ do not depend on the interaction position along the signature direction Z has been used to split the scan over the signature position $\vec{x}_s$ over a position perpendicular to the direction of the signature $\hat{d}_s$ and the coordinate Z along $\hat{d}_s$. Similar to section 12.2, the coordinates Δx, Δy and Δz (as given in equations 12.16 and 12.17) are used to parameterise the event signature position. However, rather than using the (unknown) true event signature position $\vec{x}_s^{\text{true}}$, the reference position (where $\Delta x = \Delta y = \Delta z = 0$) is given by the point at the seabed ($z_s = 0$) where $R_s = 0$, with R_s given in equation 11.21.

In table 14.3, the order in which the parameters of $\boldsymbol{x}_s$ is given, as well as the calculated quantities in each step and an indication of the required time to calculate these.

Abscissae

Ideally, the event likelihood landscape is scanned over the full phase-space of $\boldsymbol{x}$. Computationally, however, this is not feasible. Therefore, the abscissae in $\boldsymbol{x}$ at which the event

	Step		Calculated parameters: event	all PMTs	hit PMTs	$\frac{\text{time}}{\text{call}}$ [ms]
1.	Load event		$t_{\text{1st trig.}}$, $l(\mathcal{E}\|^{40}K)$	$\frac{dn_{^{40}K}}{dt}$	t_i	≈ 0
2.	Direction	$\hat{d}_s$		θ_i		≈ 0
3.	Perpendicular position	Δx, Δy		$\rho_i, \phi_i, N_{\mu,..,i}$	$\dot{n}_{\mu,..,i}$ abscissae	69
4.	Longitudinal position	Δz, $\vec{x}_s$	$l_{\text{hit/n.h.}}(\mathcal{E}\|\boldsymbol{x}_s)$	$Z_i, N_{S,1\text{ GeV},i}$	$\dot{n}_{S,1\text{ GeV},i}$ abscissae	77
5.	Time	t_s			$\dot{n}_{\mu,..,i}$ at t_s $\dot{n}_{S,1\text{ GeV},i}$ at t_s	2.37
6.	Muon energy	E_μ	L_μ	$E_\mu(Z_i)$ $E_\delta(Z_i), N_{\mu,i}$	$\dot{n}_{\mu,i}(t\|\boldsymbol{x}_s)$	2.79
7.	Vis. shower energy	$E_S^{\text{vis.}}$	$l(\mathcal{E}\|\boldsymbol{x}_s)$	$N_{S,i}$	$\dot{n}_{S,i}(t\|\boldsymbol{x}_s)$	0.74

Table 14.3: *The order in which each of the parameters $\boldsymbol{x}_s$ of the event signature are sampled to obtain event likelihood tables. The parameters related to the event likelihood that are calculated (and used in all succeeding stages) are summarised in the middle columns (see text for a description of each). The subscript $\dot{n}_{\mu,\cdots,i}$ is used to denote the time-dependent number of photoelectrons on PMT i for each of the three components (Cherenkov light, delta-rays, bremsstrahlung pair-production and photonuclear reactions) contributing to muon light. The time needed (on an AMD, 7551P, 32-core processor with a 2 GHz clock) to compute these parameters for a typical event (1622 hit PMTs, 62548 non-hit PMTs) is listed in the last column.*

likelihood is evaluated are chosen such that the event likelihood landscape is sampled accurately in the regions of $\boldsymbol{x}_s$ close to the maximum of the integrand of the test statistic integral:

$$\mathcal{L}(\mathcal{E}_i|\mathcal{H}) \equiv \int \mathcal{L}(\mathcal{E}_i|\boldsymbol{x}_s) \cdot P^{det}(\boldsymbol{x}_s) \cdot \frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s|\mathcal{H}) d\boldsymbol{x}_s. \tag{14.50}$$

The expected flux and detection efficiency depend mainly on the event signature direction $\hat{d}_s$ and energies E_S^{vis} and E_μ, and less on the interaction position $\vec{x}_s$ and time t_s. Therefore, $\hat{d}_s$, E_S^{vis} and E_μ are sampled over a broad domain, while the abscissae in $\vec{x}_s$ and t_s are chosen such that the integrated event likelihood $\iiiint \mathcal{L} d\vec{x}_s dt_s$ is evaluated with a relative accuracy of about 10^{-3} for each abscissa in $\hat{d}_s$, E_S^{vis} and E_μ.

In order to give an impression of the typical number of abscissae used, the routine described in this section has been applied to a set of 100 simulated events. These events will be referred to as the 'benchmark' events.

Event Signature Direction

The sampling of the event signature direction $\hat{d}_s$ consists of two steps: First, the full sky is sampled with 211 directions, distributed (approximately) isotropic using the sampling given in table 11.2, using a grid angle $\Delta\hat{d}$ of 10°. This relatively course sampling provides a first impression of the event likelihood landscape, which is foreseen to be used in a second step, in which directions are sampled close to the found maximum, close to the estimated neutrino direction using the reconstruction algorithms discussed in chapter 5, or close to directions matching the locations of potential neutrino sources.

Energies

The abscissae used for the event signature energies E_S^{vis} and E_μ lie on a two-dimensional grid, given by $E \in \{0, 10^{0.25}, 10^{0.75}, 10^{1.25}, 10^{1.75}, 10^{2.25}, \cdots, 10^{9.25}, 10^{9.75}\}$GeV for both $E_S^{vis.}$ and E_μ. However, in order to minimise the required computing time, a fraction of these abscissae are used for each given $\hat{d}_s$ and $\vec{x}_s$. An abscissa is not selected if the time-integrated event likelihood $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ for that energy is unlikely to be within 10^{-30} of the maximum $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ for all energy abscissae at that $\vec{x}_s$ and $\hat{d}_s$. The value of 10^{-30} is based on the domain of the expected neutrino flux of typical sources (see chapter 11.1).

In order to select the energy abscissae, an estimate of the time-integrated event likelihood is used, called the hit/non-hit likelihood:

$$l_{\text{hit/n.h.}}(\mathcal{E}|\boldsymbol{x}_s) \equiv l^{\text{hits}}_{\text{hit/n.h.}}(\mathcal{E}|\boldsymbol{x}_s) + l^{\text{n.h. PMTs}}(\mathcal{E}|\boldsymbol{x}_s), \tag{14.51}$$

with $l^{\text{n.h. PMTs}}(\mathcal{E}|\boldsymbol{x}_s)$ as defined in equation 12.15, and

$$l^{\text{hits}}_{\text{hit/n.h.}}(\mathcal{E}|\boldsymbol{x}_s) \equiv \sum_{\text{hit PMTs}} \left[\log\left(1 - \exp\left(N_i(\boldsymbol{x}_s)\right)\right)\right], \tag{14.52}$$

For each given direction and position, only the energy abscissa with the maximum hit/non-hit likelihood ($l^{max.}_{\text{hit/n.h.}}(\mathcal{E}|\boldsymbol{x}_s)$) and all abscissae with $l_{\text{hit/n.h.}}(\mathcal{E}|\boldsymbol{x}_s) \geq l^{max.}_{\text{hit/n.h.}}(\mathcal{E}|\boldsymbol{x}_s) - 30$ are used.

Time

The time-integrated event likelihood $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ is calculated for each selected energy abscissa (see previous paragraph). In order to make optimal use of the evaluated $\dot{n}_{\ldots}$ and $N_{\ldots}$, these integrals share the same abscissae in the time $\{t_{s,1}, t_{s,2}, \cdots, t_{s,m}\}$. The time abscissae are chosen such that the integral $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ (equation 12.21) is calculated with a relative accuracy of about 10^{-3} for each of the considered energy abscissae.

A straightforward algorithm is used to select the time abscissae. The following abbre-

viations will be used to describe this:

$$
\begin{aligned}
x &\Leftrightarrow t_{sig}, \\
\{x_1, x_2, \cdots, x_m\} &\Leftrightarrow \{t_{s,1}, t_{s,2}, \cdots, t_{s,m}\}, \\
f_j(x) &\Leftrightarrow \max\left[\mathcal{L}(\mathcal{E}|\boldsymbol{x}_s) - \mathcal{L}_{\text{late } t_s}, 0\right] && \text{at energy abscissa } j, \\
f_j^{max} & && \text{Maximum of } f_j(x), \\
f^{max} &\equiv \max_{j,x}\left[f_j(x)\right] && \text{Maximum of any } f_j(x)., \\
x^{max} & && \text{Time corresponding to } f^{max},
\end{aligned}
$$

In order to avoid evaluating the same t_s multiple times, and to limit the total number of abscissae, a new abscissa at x_{N+1} is added to the existing set of N abscissae only in case three conditions are met:

$$
\begin{cases}
|x_{N+1} - t_{\text{1st trig.}}| & \leq dx_{max}, \\
|x_{N+1} - x_i| & \geq dx_{min.} \quad \forall x_i \in \{x_1, x_2, \cdots, x_N\}, \\
N & \leq N_{max},
\end{cases}
\tag{14.53}
$$

where the constants $dx_{min.}$ =0.1 ns, $dx_{max.}$ =10^5 ns and N_{max} = 1000 are used.

The algorithm used to determine the abscissae consists of four steps:

1. **Start values**. The first abscissae are chosen such that at least one of these abscissae is relatively close to the maximum of each $f_j(x)$. For the first sampled longitudinal position at a given perpendicular position, multiple set of start values are used. Each start value corresponds to the event signature time t_s for which the maxima of either $\dot{n}_{S,1\text{ GeV}}(t-t_s, \rho_i, Z_i, \theta_i, \phi_i)$ or $\dot{n}_{\mu,\text{Cher.}}(t-t_s, \rho_i, \theta_i, \phi_i)$ (equations 14.32) coincides with any of the detected hit times.

2. **Extension**. For each evaluated abscissa x_i, the points $x_i - dx_1$ and $x_i + dx_1$ are evaluated in case $f_j(x_i) \geq f_j^{max}$ for any j. This step is repeated until this condition is met for all j, or the points to add do not meet the conditions given in equation 14.53. The step size dx_1 is chosen to be 1 ns for the first iteration (k = 1), and is increased by $dx_{k+1} = F_{dx} \cdot dx_k$, with F_{dx} = 2 for each subsequent iteration.

3. **Refinement**. With the previous step finished, each f_j^{max} is enclosed by at least one abscissae at smaller x and one abscissa at larger x. In this step, points are added in between each pair of subsequent abscissae x_i, x_{i+1} in case the integral $\mathcal{I}_i(\mathcal{E}|\boldsymbol{x}_s)$ in this time domain is higher than one percent of the total integral $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ at any energy abscissa j. This step is repeated until this condition is met for each pair of abscissae, or no points can be added under the conditions outlined in equation 14.53.

4. **Start value for next point**. The coordinate x^{max} at the evaluated Δz is used as a start value for the sampling of the next time profile. In case the time-dependent event likelihood landscape $f_j(x)$ has been characterised at multiple longitudinal positions at the same perpendicular position, a linear inter- or extrapolation of the corresponding x^{max} values is used as start value. Similarly, the full width half maximum of the evaluated time profile(s) is used to define the step size dx_1 of the next evaluated longitudinal position.

Event Signature Position

The algorithm to sample the longitudinal (Δz) and perpendicular (Δx and Δy) positions is almost identical as the used to sample t_s. It consists of three steps; the start value(s), an extension phase and a refinement stage.

1. Start values For each given direction, the event signature position is sampled in the vicinity of two start values: the reconstructed shower position by the shower reconstruction algorithm (see section 5.2) and the reconstructed position by the first step in the muon track reconstruction (the prefit, see section 5.3) for the given $\hat{d}_s$.

	Perpendicular position	Long. pos.	Time
Variables x	Δx, Δy	Δz	t_s
Function $f(x)$	maximum $\mathcal{I}(\mathcal{E}\|\boldsymbol{x}_s)$ over all Δz	$\mathcal{I}(\mathcal{E}\|\boldsymbol{x}_s)$	$\max\left[\mathcal{L}(\mathcal{E}\|\boldsymbol{x}_s) - \mathcal{L}_{\text{late } t_s}, 0\right]$
x_1	Reconstructed $\vec{x}_s$		Projected hits
x-range	Instr. vol. + 800 m		$t_{\text{1st trig.}} \pm 10^5$ ns
dx_{min}	0.1 m		0.1 ns
N_{max}	200	100	1000
dx_1	4 m	1 m	1 ns
F_{dx}	2		

Table 14.4: *An overview of the variables and parameters used in the sampling of the event signature position $\vec{x}_s$ and time t_s discussed in the text.*

2&3. Extension and refinement The algorithm for the sampling of Δx, Δy and Δz is identical to the one used for t_s, but the quantities used are different. These quantities and the settings of the routine are given in table 14.4. The settings are such that the time-integrated event likelihood $\mathcal{I}(\mathcal{E}|\boldsymbol{x}_s)$ is sampled with at least five points Δx, Δy and Δz in the domain defined by the maximum value and 10^{-3} times the maximum value. In order to sample the two-dimensional parameter space of the perpendicular position, each abscissa is given by two coordinates $[\Delta x_i; \Delta y_i]$, and the four 'neighbouring' abscissae $[\Delta x_i - dx_k; \Delta y_i]$, $[\Delta x_i + dx_k; \Delta y_i]$, $[\Delta x_i; \Delta y_i - dx_k]$ and $[\Delta x_i; \Delta y_i + dx_k]$ at the given dx_k are added. A hexagonal sampling of the perpendicular position, in which the six points $[\Delta x_i - dx_k; \Delta y_i]$, $[\Delta x_i + dx_k; \Delta y_i]$, $[\Delta x_i - \frac{1}{2}dx_k; \Delta y_i - \sqrt{\frac{3}{4}}dx_k]$, $[\Delta x_i - \frac{1}{2}dx_k; \Delta y_i + \sqrt{\frac{3}{4}}dx_k]$, $[\Delta x_i + \frac{1}{2}dx_k; \Delta y_i - \sqrt{\frac{3}{4}}dx_k]$ and $[\Delta x_i + \frac{1}{2}dx_k; \Delta y_i + \sqrt{\frac{3}{4}}dx_k]$ are added in the second step, has been tried as well. This option leads typically to more abscissae without improvement of the accuracy of the obtained event likelihood 'landscape'.

The range of the sampled longitudinal and perpendicular position is bounded by a maximum distance to the instrumented volume of at most 800 m. The extra 800 m is given by the range of the tables used to evaluate the time-dependent number of photoelectrons on each PMT.

ℍ. Simulated Benchmark Events for Chapter 13

The main characteristics of the 60 simulated benchmark events, used to obtain the results discussed in section 13.3, are given in table 14.5.

event nr.	Int. type	True interaction parameters E_ν	$E_S^{vis.}$ [GeV]	E_μ	θ [°]	ϕ [°]	x [m]	y [m]	z [m]	Event $N_{\text{hit PMTs}}$	$N_{\text{n.h. PMTs}}$	At true parameters $\mathcal{L}(\mathcal{E}\|\boldsymbol{x}_s)$	$P^{det}(\boldsymbol{x}_s)$	$\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s\|\mathcal{H}_0)$ [$1/(GeV^2 \cdot m^3 \cdot s \cdot sr)$]	$\frac{d\Phi}{d\boldsymbol{x}_s}(\boldsymbol{x}_s\|\mathcal{H}_1)$ [$1/(GeV^2 \cdot m^3 \cdot s \cdot sr)$]	$\mathcal{L}(\mathcal{E}\|\mathcal{H}_0)$	$\mathcal{L}(\mathcal{E}\|\mathcal{H}_1)$	$\zeta(\mathcal{E}_i\|\mathcal{H}_1)$
2173	ν_e, CC	$10^{5.36}$	$10^{5.30}$	$10^{0.08}$	21.9	−61.4	462.8	426.5	392.6	1622	62548	$10^{-20296.8}$	0.80	$10^{-51.46}$	$10^{-51.46}$	$10^{-20310.34}$	$10^{-20309.82}$	$10^{12795.38}$
2175	ν_e, CC	$10^{5.44}$	$10^{5.42}$	0	105.8	139.4	10.0	−651.8	484.0	1087	63083	$10^{-12933.8}$	1.00	$10^{-52.51}$	$10^{-51.40}$	$10^{-12946.13}$	$10^{-12945.79}$	$10^{0.31}$
2178	ν_e, CC	$10^{5.26}$	$10^{5.25}$	0	103.7	176.7	−20.6	−449.4	261.4	3932	60238	$10^{-47182.6}$	1.00	$10^{-50.61}$	$10^{-49.89}$	$10^{-46627.34}$	$10^{-46626.93}$	$10^{0.56}$
2180	ν_e, CC	$10^{5.04}$	$10^{5.03}$	0	106.1	−105.7	−456.7	−159.1	715.2	1970	62200	$10^{-24354.7}$	1.00	$10^{-49.55}$	$10^{-49.07}$	$10^{-24237.68}$	$10^{-24237.41}$	$10^{0.57}$
2190	ν_e, CC	$10^{5.36}$	$10^{5.35}$	0	99.6	−146.3	500.2	359.2	352.8	2179	61991	$10^{-25874.3}$	1.00	$10^{-51.59}$	$10^{-50.79}$	$10^{-25892.97}$	$10^{-25892.62}$	$10^{0.43}$
2192	ν_e, CC	$10^{5.39}$	$10^{5.38}$	0	88.6	176.9	111.8	297.2	37.6	2572	61598	$10^{-29448.1}$	1.00	$10^{-51.21}$	$10^{-50.59}$	$10^{-29616.56}$	$10^{-29616.40}$	$10^{0.27}$
2193	ν_e, CC	$10^{5.27}$	$10^{5.26}$	0	63.3	149.6	−339.6	−407.3	402.4	3183	60987	$10^{-36751.6}$	1.00	$10^{-51.04}$	$10^{-50.03}$	$10^{-37126.22}$	$10^{-37125.93}$	$10^{0.29}$
2195	ν_e, CC	$10^{5.24}$	$10^{5.23}$	0	148.2	−126.7	110.1	72.7	657.1	2997	61173	$10^{-33020.4}$	1.00	$10^{-52.16}$	$10^{-50.97}$	$10^{-33956.19}$	$10^{-33956.19}$	$10^{0.00}$
2197	ν_e, CC	$10^{5.03}$	$10^{4.83}$	0	90.6	−0.0	45.2	363.3	174.4	2660	61510	$10^{-30265.4}$	1.00	$10^{-47.00}$	$10^{-46.87}$	$10^{-29604.52}$	$10^{-29604.42}$	$10^{0.73}$
2200	ν_e, CC	$10^{5.38}$	$10^{5.38}$	0	146.9	−131.8	148.3	563.3	273.3	2412	61758	$10^{-27934.0}$	1.00	$10^{-53.69}$	$10^{-52.23}$	$10^{-27543.52}$	$10^{-27543.41}$	$10^{0.07}$
2201	ν_e, CC	$10^{5.24}$	$10^{5.19}$	0	17.1	89.4	84.2	659.0	466.2	523	63647	$10^{-6582.4}$	0.70	$10^{-50.64}$	$10^{-49.45}$	$10^{-6618.01}$	$10^{-6617.64}$	$10^{0.31}$
2203	ν_e, CC	$10^{5.02}$	$10^{5.01}$	0	104.0	−116.9	−407.7	−238.2	578.3	2275	61895	$10^{-26012.8}$	1.00	$10^{-49.24}$	$10^{-48.82}$	$10^{-26411.57}$	$10^{-26411.37}$	$10^{0.49}$
2205	ν_e, CC	$10^{5.01}$	$10^{4.96}$	0	138.1	94.4	−546.7	122.6	328.5	2540	61630	$10^{-29948.7}$	1.00	$10^{-50.15}$	$10^{-49.45}$	$10^{-30184.48}$	$10^{-30184.07}$	$10^{0.58}$
2207	ν_e, CC	$10^{5.22}$	$10^{5.22}$	0	155.9	68.2	136.9	566.4	531.7	1774	62396	$10^{-21791.5}$	1.00	$10^{-52.36}$	$10^{-51.17}$	$10^{-21657.53}$	$10^{-21657.02}$	$10^{0.42}$
2210	ν_e, CC	$10^{5.16}$	$10^{5.15}$	0	106.9	−79.6	33.1	690.0	378.6	1366	62804	$10^{-16733.7}$	0.18	$10^{-50.08}$	$10^{-49.43}$	$10^{-16725.06}$	$10^{-16724.69}$	$10^{0.56}$
2211	ν_e, CC	$10^{5.16}$	$10^{5.03}$	0	24.3	146.6	−359.3	−516.5	186.8	1591	62579	$10^{-19573.5}$	0.96	$10^{-50.21}$	$10^{-49.30}$	$10^{-19556.64}$	$10^{-19555.93}$	$10^{0.78}$
2212	ν_e, CC	$10^{5.42}$	$10^{5.33}$	0	151.1	−16.8	−331.8	−378.1	624.8	2833	61337	$10^{-31372.9}$	1.00	$10^{-53.72}$	$10^{-52.34}$	$10^{-31123.19}$	$10^{-31122.83}$	$10^{0.26}$
2214	ν_e, CC	$10^{5.07}$	$10^{5.06}$	0	109.1	−5.6	433.8	301.4	644.7	769	63401	$10^{-9321.2}$	0.92	$10^{-49.90}$	$10^{-49.34}$	$10^{-9106.57}$	$10^{-9106.35}$	$10^{0.40}$
2215	ν_e, CC	$10^{5.02}$	$10^{5.01}$	0	77.3	164.9	−193.0	−383.1	62.1	2367	61803	$10^{-28237.7}$	1.00	$10^{-49.04}$	$10^{-48.64}$	$10^{-27649.87}$	$10^{-27649.76}$	$10^{0.28}$
2219	ν_e, CC	$10^{5.44}$	$10^{5.44}$	0	139.3	54.2	−491.7	−13.7	739.4	2468	61702	$10^{-28695.0}$	1.00	$10^{-54.04}$	$10^{-52.48}$	$10^{-29052.32}$	$10^{-29048.45}$	$10^{2.47}$
2221	ν_e, CC	$10^{5.01}$	$10^{5.00}$	0	85.1	−139.0	−441.7	−178.9	57.3	2231	61939	$10^{-26972.3}$	1.00	$10^{-48.41}$	$10^{-48.17}$	$10^{-27001.84}$	$10^{-27001.68}$	$10^{0.68}$
2222	ν_e, CC	$10^{5.35}$	$10^{5.02}$	0	73.8	−151.8	−298.5	259.1	779.7	797	63373	$10^{-9895.5}$	0.43	$10^{-49.25}$	$10^{-48.78}$	$10^{-9873.62}$	$10^{-9873.09}$	$10^{1.14}$
2223	ν_e, CC	$10^{5.07}$	$10^{5.07}$	0	117.2	90.6	−127.7	−234.4	765.1	1170	63000	$10^{-14057.0}$	1.00	$10^{-50.33}$	$10^{-49.64}$	$10^{-14052.27}$	$10^{-14051.97}$	$10^{0.43}$
2224	ν_e, CC	$10^{5.29}$	$10^{5.23}$	0	3.4	−167.6	320.3	550.8	636.3	649	63521	$10^{-8124.4}$	0.61	$10^{-51.14}$	$10^{-49.86}$	$10^{-8146.88}$	$10^{-8146.21}$	$10^{0.52}$
2225	ν_e, CC	$10^{5.04}$	$10^{5.01}$	0	24.2	26.8	−92.2	615.6	442.4	860	63310	$10^{-9953.3}$	0.78	$10^{-50.12}$	$10^{-49.23}$	$10^{-9891.80}$	$10^{-9890.07}$	$10^{1.95}$
2230	ν_e, CC	$10^{5.32}$	$10^{5.31}$	0	61.2	−110.1	369.2	−123.8	748.8	699	63471	$10^{-8494.7}$	0.72	$10^{-51.85}$	$10^{-50.69}$	$10^{-8458.11}$	$10^{-8457.51}$	$10^{0.52}$
2236	ν_e, CC	$10^{5.25}$	$10^{5.07}$	0	83.9	−112.0	−193.5	−169.7	567.5	2732	61438	$10^{-28728.9}$	1.00	$10^{-49.04}$	$10^{-48.71}$	$10^{-30757.42}$	$10^{-30757.42}$	$10^{0.00}$
2237	ν_e, CC	$10^{5.14}$	$10^{5.14}$	0	126.6	30.9	443.3	301.9	340.2	846	63324	$10^{-10055.4}$	1.00	$10^{-50.92}$	$10^{-50.03}$	$10^{-9915.01}$	$10^{-9914.26}$	$10^{0.84}$
3769	ν_μ, CC	$10^{6.75}$	0	$10^{6.04}$	88.8	−139.0	1373.9	1705.6	329.6	5024	59146	$10^{-64353.2}$	0.00	$10^{-63.80}$	$10^{-60.07}$	$10^{-56705.87}$	$10^{-56701.87}$	$10^{1.07}$
3770	ν_μ, CC	$10^{6.82}$	0	$10^{5.48}$	86.4	134.0	966.0	−1195.9	56.1	4967	59203	$10^{-64246.4}$	0.00	$10^{-60.84}$	$10^{-58.68}$	$10^{-58578.09}$	$10^{-58578.01}$	$10^{0.04}$
3771	ν_μ, CC	$10^{6.61}$	0	$10^{6.37}$	75.2	132.2	664.8	−1146.4	506.3	1102	63068	$10^{-14484.0}$	0.00	$10^{-64.37}$	$10^{-59.03}$	$10^{-13515.35}$	$10^{-13512.96}$	$10^{0.45}$
3775	ν_μ, CC	$10^{6.76}$	0	$10^{4.39}$	94.8	−64.1	−7161.1	14089.8	1642.4	2186	61984	$10^{-28434.3}$	0.00	$10^{-55.21}$	$10^{-54.62}$	$10^{-27147.67}$	$10^{-27147.67}$	$10^{0.00}$
3778	ν_μ, CC	$10^{6.60}$	0	$10^{6.40}$	155.2	−96.9	−115.8	635.6	1876.8	6027	58143	$10^{-77082.5}$	0.00	$10^{-72.13}$	$10^{-67.48}$	$10^{-69135.07}$	$10^{-69132.72}$	$10^{0.51}$
3781	ν_μ, CC	$10^{6.75}$	0	$10^{5.04}$	111.3	110.6	2103.8	−4452.3	2253.1	3300	60870	$10^{-42877.9}$	0.00	$10^{-59.05}$	$10^{-57.46}$	$10^{-41011.70}$	$10^{-41011.43}$	$10^{0.17}$
3786	ν_μ, CC	$10^{6.57}$	0	$10^{4.07}$	27.6	−27.4	−279.6	−294.6	0.0	1200	62970	$10^{-15961.9}$	0.76	$10^{-54.67}$	$10^{-53.84}$	$10^{-15329.02}$	$10^{-15328.99}$	$10^{0.04}$
3788	ν_μ, CC	$10^{6.57}$	0	$10^{4.79}$	64.4	−35.9	−813.9	220.7	0.0	2348	61822	$10^{-30897.1}$	0.00	$10^{-57.72}$	$10^{-56.39}$	$10^{-29387.95}$	$10^{-29387.91}$	$10^{0.03}$
3789	ν_μ, CC	$10^{6.59}$	0	$10^{5.04}$	96.6	130.3	3152.3	−3306.9	1012.3	2407	61763	$10^{-31430.1}$	0.00	$10^{-58.72}$	$10^{-57.28}$	$10^{-29369.17}$	$10^{-29369.14}$	$10^{0.02}$
3790	ν_μ, CC	$10^{6.88}$	0	$10^{4.39}$	116.5	155.4	4271.5	−1812.7	2622.2	2016	62154	$10^{-26463.2}$	0.00	$10^{-55.97}$	$10^{-55.03}$	$10^{-25553.29}$	$10^{-25553.19}$	$10^{0.11}$
3792	ν_μ, CC	$10^{6.94}$	0	$10^{6.09}$	146.5	121.2	1027.3	−989.5	2938.0	4633	59537	$10^{-59653.5}$	0.00	$10^{-66.84}$	$10^{-63.09}$	$10^{-54586.50}$	$10^{-54583.66}$	$10^{0.76}$
3794	ν_μ, CC	$10^{7.01}$	0	$10^{4.96}$	100.8	133.0	7984.4	−8962.9	2635.0	2653	61517	$10^{-34307.9}$	0.00	$10^{-58.46}$	$10^{-57.06}$	$10^{-31957.36}$	$10^{-31957.33}$	$10^{0.02}$
3799	ν_μ, CC	$10^{7.02}$	0	$10^{4.78}$	51.3	−135.4	419.2	−236.9	0.0	1797	62373	$10^{-23948.7}$	1.00	$10^{-57.74}$	$10^{-56.36}$	$10^{-23141.95}$	$10^{-23141.92}$	$10^{0.02}$
3801	ν_μ, CC	$10^{7.42}$	0	$10^{5.31}$	164.5	7.9	−933.3	−494.1	3302.6	2885	61285	$10^{-37588.5}$	0.00	$10^{-62.06}$	$10^{-60.08}$	$10^{-34839.81}$	$10^{-34839.37}$	$10^{0.22}$
3802	ν_μ, CC	$10^{7.06}$	0	$10^{6.38}$	135.6	34.1	−2174.3	−1926.3	2822.7	4117	60053	$10^{-53458.9}$	0.00	$10^{-68.85}$	$10^{-64.19}$	$10^{-49378.37}$	$10^{-49375.83}$	$10^{0.55}$
3804	ν_μ, CC	$10^{7.47}$	0	$10^{6.85}$	131.4	136.6	1446.8	−1143.0	2577.3	5992	58178	$10^{-75759.6}$	0.00	$10^{-75.32}$	$10^{-69.30}$	$10^{-67598.46}$	$10^{-67592.79}$	$10^{0.94}$
3805	ν_μ, CC	$10^{7.38}$	0	$10^{7.05}$	171.2	137.9	606.6	−606.8	2578.1	4793	59377	$10^{-62007.1}$	0.00	$10^{-86.27}$	$10^{-80.04}$	$10^{-55478.81}$	$10^{-55474.99}$	$10^{0.61}$
3808	ν_μ, CC	$10^{7.16}$	0	$10^{4.31}$	20.4	103.1	142.3	−183.8	0.0	1379	62791	$10^{-18090.6}$	1.00	$10^{-55.78}$	$10^{-54.73}$	$10^{-17034.61}$	$10^{-16449.88}$	$10^{561.53}$
3809	ν_μ, CC	$10^{7.49}$	0	$10^{5.03}$	13.4	−155.8	124.2	328.4	0.0	2534	61636	$10^{-33354.4}$	1.00	$10^{-58.90}$	$10^{-57.23}$	$10^{-30717.11}$	$10^{-30716.68}$	$10^{0.26}$
3811	ν_μ, CC	$10^{7.12}$	0	$10^{5.57}$	91.7	−155.8	4187.9	1637.1	241.9	4562	59608	$10^{-58643.4}$	0.00	$10^{-61.31}$	$10^{-58.97}$	$10^{-55155.24}$	$10^{-55154.80}$	$10^{0.19}$
3812	ν_μ, CC	$10^{7.18}$	0	$10^{3.01}$	23.3	161.0	78.2	14.6	0.0	900	63270	$10^{-11873.7}$	0.60	$10^{-49.05}$	$10^{-48.91}$	$10^{-11390.97}$	$10^{-11390.97}$	$10^{0.06}$
3813	ν_μ, CC	$10^{7.32}$	0	$10^{5.42}$	26.4	102.7	411.0	272.7	0.0	1968	62202	$10^{-26156.1}$	1.00	$10^{-60.71}$	$10^{-58.51}$	$10^{-24791.22}$	$10^{-24790.79}$	$10^{0.20}$
3814	ν_μ, CC	$10^{7.34}$	0	$10^{7.30}$	99.4	−56.0	−630.4	1336.6	416.8	11629	52541	$10^{-145470.9}$	0.00	$10^{-74.87}$	$10^{-68.02}$	$10^{-136979.81}$	$10^{-136973.05}$	$10^{0.99}$
3815	ν_μ, CC	$10^{7.46}$	0	$10^{7.02}$	89.1	153.2	1991.3	−1021.8	690.2	5379	58791	$10^{-69623.1}$	0.00	$10^{-70.01}$	$10^{-63.87}$	$10^{-63335.34}$	$10^{-63333.63}$	$10^{0.28}$
3818	ν_μ, CC	$10^{7.14}$	0	$10^{4.49}$	15.9	−111.0	263.9	−341.9	0.0	1108	63062	$10^{-14653.4}$	1.00	$10^{-56.57}$	$10^{-55.38}$	$10^{-13533.54}$	$10^{-13533.41}$	$10^{0.11}$
3824	ν_μ, CC	$10^{7.46}$	0	$10^{5.31}$	118.0	83.9	−546.0	−2855.8	2042.1	4563	59607	$10^{-58986.8}$	0.00	$10^{-60.62}$	$10^{-58.64}$	$10^{-54714.20}$	$10^{-54713.95}$	$10^{0.13}$
3828	ν_μ, CC	$10^{7.46}$	0	$10^{7.43}$	134.7	18.1	−26.3	81.7	1235.4	4409	59761	$10^{-56448.1}$	0.00	$10^{-82.74}$	$10^{-75.62}$	$10^{-52946.52}$	$10^{-52941.09}$	$10^{0.76}$
3830	ν_μ, CC	$10^{7.02}$	0	$10^{5.62}$	121.6	−45.2	−2319.1	2057.6	2699.3	2547	61623	$10^{-32943.6}$	0.00	$10^{-62.50}$	$10^{-59.94}$	$10^{-30380.04}$	$10^{-30379.18}$	$10^{0.33}$
3837	ν_μ, CC	$10^{7.26}$	0	$10^{6.98}$	149.0	−118.3	63.8	1076.7	2833.5	4221	59949	$10^{-54861.6}$	0.00	$10^{-79.58}$	$10^{-73.46}$	$10^{-50799.81}$	$10^{-50797.11}$	$10^{0.44}$
3840	ν_μ, CC	$10^{7.48}$	0	$10^{5.50}$	104.6	−45.0	−6133.7	5870.2	2912.9	2651	61519	$10^{-34264.0}$	0.00	$10^{-61.30}$	$10^{-59.03}$	$10^{-32198.84}$	$10^{-32198.72}$	$10^{0.05}$
3844	ν_μ, CC	$10^{7.04}$	0	$10^{5.07}$	30.7	62.8	390.3	−411.4	0.0	2106	62064	$10^{-27812.1}$	0.37	$10^{-59.09}$	$10^{-57.37}$	$10^{-27417.00}$	$10^{-27416.96}$	$10^{0.02}$
3847	ν_μ, CC	$10^{7.46}$	0	$10^{7.39}$	129.4	−161.7	1122.3	−243.8	998.6	3924	60246	$10^{-51676.4}$	0.00	$10^{-81.12}$	$10^{-74.07}$	$10^{-48907.79}$	$10^{-48905.46}$	$10^{0.33}$

Table 14.5: *A summary of the characteristic properties of the benchmark events (see text). Given are the interaction channel (second column). The true energies of the neutrino, visible shower and track (columns three-five). The direction of the neutrino (col. 6-7) and position of the direction (col 8-10) in the coordinate frame discussed in section 11.2.1. The number of 'hit' and 'non-hit' PMTs (columns 11 and 12, see section efsec:eventlikelihooddefinition for the definition of a hit PMT). The event likelihood (col. 13), detection efficiency (col. 14) and event signature fluxes under the $\mathcal{H}_0$ (col. 15) and an arbitrary benchmark $\mathcal{H}_1$ hypothesis, with $\gamma = 2.4$ and $\Phi^{100\ TeV}_{\nu+\bar{\nu},e+\mu+\tau} = 4.5$ $(GeV \cdot m^2 \cdot sr \cdot s)^{-1}$ (col. 16). The calculated test statistic integrals for these two hypotheses are given in the last two columns.*

Ⅱ. Including Time-over-Threshold Information in the Event Likelihood

In this appendix, a model will be derived which can be used to include ToTs information in reconstruction algorithms or the event likelihood function described in chapter 12. Because this model has been realised relatively recently, the results presented in this work are obtained without using it.

The ToTs values of all L0 on a PMT can be used to obtain an estimate $\tilde{n}_\gamma$ of the total number of detected photons by the PMT:

$$\tilde{n}_\gamma \equiv \sum_{\text{L0s } j} \left[\text{ToT}^{inv}\left(\text{ToT}_j\right)\right], \tag{14.54}$$

where $\text{ToT}^{inv}\left(\text{ToT}\right)$ is the inverse of $\text{ToT}(n_\gamma)$, given in equation 4.2.

The probability density $\frac{\partial P}{\partial \tilde{n}_\gamma}(\tilde{n}_\gamma|n_\gamma)$ to measure $\tilde{n}_\gamma$ for a given number n_γ of detected photons has been parameterised using simulations. In figure 14.10, the distributions of $\frac{\partial P}{\partial \tilde{n}_\gamma}(\tilde{n}_\gamma|n_\gamma)$ are shown for several n_γ. These distributions are reasonably described by a gamma distribution

$$\frac{\partial P}{\partial \tilde{n}_\gamma}(\tilde{n}_\gamma|n_\gamma) \approx \frac{e^{-(\tilde{n}_\gamma - x_0)/\theta} \cdot \left(\tilde{n}_\gamma - x_0\right)^{k-1}}{\Gamma\left(k\right) \cdot \left(\theta\right)^k}, \tag{14.55}$$

where the values of k, θ and x_0 depend on n_γ. This dependence is parameterised using a fit to the distributions in which the mean, variance and offset of the distributions are varied:

$$\begin{aligned} x_0 &\approx \begin{cases} -29.2 & \text{if } n_\gamma = 1 \\ 0 & \text{else} \end{cases}, \\ \theta = \frac{\text{Var}\left[\frac{\partial P}{\partial \tilde{n}_\gamma}\right]}{\text{Mean}\left[\frac{\partial P}{\partial \tilde{n}_\gamma}\right] - x_0} &\approx \frac{0.228 - 0.202 \cdot n_\gamma + 0.129 \cdot {n_\gamma}^2}{1.169 \cdot n_\gamma - x_0}, \\ k = \frac{\left(\text{Mean}\left[\frac{\partial P}{\partial \tilde{n}_\gamma}\right] - x_0\right)^2}{\text{Var}\left[\frac{\partial P}{\partial \tilde{n}_\gamma}\right]} &\approx \frac{\left(1.169 \cdot n_\gamma - x_0\right)^2}{0.228 - 0.202 \cdot n_\gamma + 0.129 \cdot {n_\gamma}^2}. \end{aligned} \tag{14.56}$$

Combining $\frac{\partial P}{\partial \tilde{n}_\gamma}(\tilde{n}_\gamma|n_\gamma)$ with the Poisson probability $P(n_\gamma|\hat{n}_\gamma)$ to detect n_γ photons in case $\hat{n}_\gamma$ are expected

$$P(n_\gamma|\hat{n}_\gamma) = \frac{e^{-\hat{n}_\gamma} \cdot \left(\hat{n}_\gamma\right)^{n_\gamma}}{n_\gamma!}, \tag{14.57}$$

the probability density to observe $\tilde{n}_\gamma$ for a given number of expected detected photons $\hat{n}_\gamma$ is then given by:

$$\frac{\partial P}{\partial \tilde{n}_\gamma}(\tilde{n}_\gamma|\hat{n}_\gamma) \equiv \sum_{n_\gamma=0}^{\infty} \left[\frac{\partial P}{\partial \tilde{n}_\gamma}(\tilde{n}_\gamma|n_\gamma) \cdot P(n_\gamma|\hat{n}_\gamma)\right]. \tag{14.58}$$

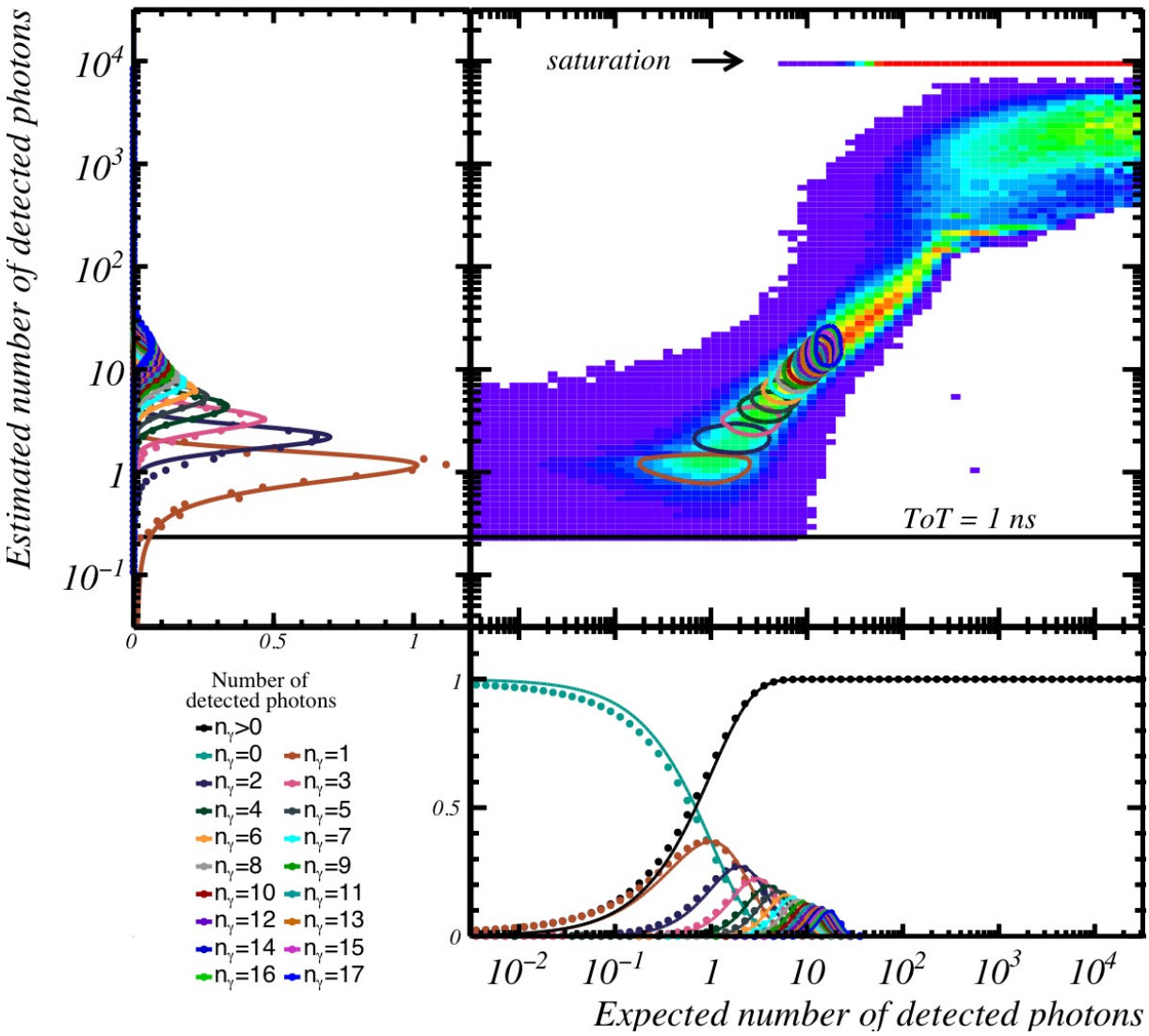


Figure 14.10: ***Top left**: The distributions of the probability density of the estimated number of detected photons $\tilde{n}$ for $n_\gamma \in \{1, 2, \cdots\}$ detected photons on a PMT, obtained from simulations (dots) and as given by the model discussed in the text. Each distribution is scaled such that $\int \frac{\partial P}{\partial \tilde{n}_\gamma}(\tilde{n}_\gamma|n_\gamma)d\tilde{n}_\gamma = 1$ for each n_γ. **Bottom right**: The number of detected photons n_γ as a function of the expected number of detected photons $\hat{n}_\gamma$ on a PMT. In the **top right** panel, the probability density function $\frac{\partial P}{\partial \tilde{n}_\gamma}(\tilde{n}_\gamma|\hat{n}_\gamma)$ is shown (obtained from simulations), with normalisation $\int \frac{\partial P}{\partial \tilde{n}_\gamma}(\tilde{n}_\gamma|\hat{n}_\gamma)d\tilde{n}_\gamma = 1$. This distribution is obtained from a large sample of simulated events (ν_e and ν_μ, CC and NC neutrino interactions) with the JSirene simulation software (see section 4.4.1), using a nominal PMT model.The coloured contours give half-maximum contours of the model discussed in the text (for $n_\gamma \in \{1, 2, \cdots\}$ detected number of photons).*

Implementing this model in the event likelihood (equation 12.1) is straightforward by

updating the definition of $\mathcal{L}_i^{\text{hit}}(\boldsymbol{x}_s)$ given in equation 12.5 to:

$$\mathcal{L}_i^{\text{hit}}(\boldsymbol{x}_s) \approx e^{-n_i(t_i|\boldsymbol{x}_s)+\frac{1}{2}\cdot\dot{n}_i(t_i|\boldsymbol{x}_s)} \cdot \left(1 - e^{-\dot{n}_i(t_i|\boldsymbol{x}_s)}\right) \cdot \frac{\partial P}{\partial \tilde{n}_{\gamma,i}}(\tilde{n}_{\gamma,i}|\hat{n}_{\gamma,i}(\boldsymbol{x}_s)). \tag{14.59}$$

with the parameters $\boldsymbol{x}_s$, $\dot{n}_i$, n_i and $\hat{n}_{\gamma,i}$ defined in section 12.1.

An even more detailed event likelihood could include the probability to observe the time and ToT of each detected L0s in the event for any given event signature $\boldsymbol{x}_s$. This probability can be derived, but is difficult to evaluate, since it depends on the conditional probability that a hit is not followed by another hit within its ToT on the same PMT.

Bibliography

[1] A. Brouwer Ancher and J. Breen. "De Doleantie van een deel der burgerij van Amsterdam tegen den magistraat dier stad in 1564 en 1565". *Bydragen en Mededelingen van het Historisch Genootschap te Utrecht* 24 (1903) (cit. on p. ii).

[2] C. Van der Sterr. "Amsterdamsch Peil". *Tijdschrift voor Kadaster en Landmeetkunde* (1934) (cit. on p. ii).

[3] N. Aghanim et al. (The Planck Collaboration). "Planck 2018 results - VI. Cosmological parameters". *A&A* 641 (2020) (cit. on pp. 1, 15).

[4] P. Zyla et al. (Particle Data Group). "Review of Particle Physics". *Progress of Theoretical and Experimental Physics* 2020 (8) (Aug. 2020) (cit. on pp. 1, 15, 30, 64, 65, 74).

[5] K. Hoffman. "High energy neutrino telescopes". *New Journal of Physics* 11 (5) (May 2009) (cit. on pp. 2, 30).

[6] Y. Akrami et al. (The Planck Collaboration). "Planck 2018 results - IV. Diffuse component separation". *A&A* 641 (2020) (cit. on pp. 2, 16).

[7] J. Thomson. "XL. Cathode Rays". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 44 (269) (1897) (cit. on p. 2).

[8] A. Einstein. "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt". *Annalen der Physik* 322 (6) (1905) (cit. on p. 2).

[9] A. Goldstein. "Über eine noch nicht untersuchte Strahlungsform an der Kathodeinducirter Entladungen". *Berlin Akd. Monatsber.* 2 (2) (1886) (cit. on p. 2).

[10] E. Rutherford. "LIV. Collision of alpha particles with light atoms. IV. An anomalous effect in nitrogen". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 37 (222) (1919) (cit. on p. 2).

[11] J. Chadwick. "The existence of a neutron". *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 136 (830) (1932) (cit. on p. 2).

[12] C. Anderson. "The Positive Electron". *Phys. Rev.* 43 (6 Mar. 1933) (cit. on p. 2).

[13] S. Neddermeyer and C. Anderson. "Note on the Nature of Cosmic-Ray Particles". *Phys. Rev.* 51 (10 May 1937) (cit. on p. 2).

[14] D. J. Bird et al. (The Fly's eye Collaboration). "Detection of a cosmic ray with measured energy well beyond the expected spectral cutoff due to cosmic microwave radiation". *Astrophys. J.* 441 (1995) (cit. on pp. 3, 16, 31).

[15] M. Ackermann et al. (The Fermi-LAT Collaboration). "Detection of the Characteristic Pion-Decay Signature in Supernova Remnants". *Science* 339 (6121) (2013) (cit. on pp. 3, 37).

[16] T. Gaisser, R. Engel, and E. Resconi. *Cosmic Rays and Particle Physics*. 2nd ed. Cambridge University Press, 2016. ISBN: 978-0-521-01646-9 (cit. on pp. 3, 32, 40, 41).

[17] M. A. Markov. "On high energy neutrino physics". *10th International Conference on High Energy Physics*. 1960 (cit. on p. 4).

[18] K. Hirata et al. (The Kamiokande Collaboration). "Observation of a neutrino burst from the supernova SN1987A". *Phys. Rev. Lett.* 58 (14 Apr. 1987) (cit. on pp. 5, 50, 51).

[19] R. Bionta et al. "Observation of a Neutrino Burst in Coincidence with Supernova 1987A in the Large Magellanic Cloud". *Phys. Rev. Lett.* 58 (14 Apr. 1987) (cit. on pp. 5, 50, 51).

[20] E. N. Alekseev et al. "Possible detection of a neutrino signal on 23 February 1987 at the Baksan underground scintillation telescope of the Institute of Nuclear Research". *Pis'ma Zh. Eksp. Teor. Fiz.* 45 (10 Apr. 1987) (cit. on pp. 5, 50, 51).

[21] W. Kunkel et al. "Supernova 1987A in the Large Magellanic Cloud". *International Astronomical Union Circular* 4316 (Feb. 1987) (cit. on pp. 5, 50).

[22] *IceCube-170922A*. GCN/Amon Notice, `https://gcn.gsfc.nasa.gov/notices_amon/50579430_130033.amon` (cit. on pp. 5, 50, 52).

[23] Y. Tanaka et al. (The Fermi-LAT Collaboration). *Fermi-LAT detection of increased gamma-ray activity of TXS 0506+056, located inside the IceCube-170922A error region.* ATel #10791, `http://www.astronomerstelegram.org/?read=10791` (cit. on pp. 5, 50).

[24] M. G. Aartsen et al. (The IceCube, Fermi-LAT, MAGIC, AGILE, ASAS-SN, HAWC, H.E.S.S., INTEGRAL, Kanata, Kiso, Kapteyn, Liverpool Telescope, Subaru, Swift NuSTAR, VERITAS and VLA/17B-403 Collabrations). "Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A". *Science* 361 (6398) (2018) (cit. on pp. 5, 50).

[25] P. Padovani et al. "Dissecting the region around IceCube-170922A: the blazar TXS 0506+056 as the first cosmic neutrino source". *Mon. Not. Roy. Astron. Soc.* 480 (1) (2018) (cit. on pp. 5, 51).

[26] J. Stettner (for the IceCube Collaboration). "Measurement of the Diffuse Astrophysical Muon-Neutrino Spectrum with Ten Years of IceCube Data". *Proceedings of Science (ICRC2019)*. Vol. 1017. 2019 (cit. on pp. 5, 18, 54, 55).

[27] H. M. Niederhausen and Y. Xu (for the IceCube Collaboration). "High Energy Astrophysical Neutrino Flux Measurement Using Neutrino-induced Cascades Observed in 4 Years of IceCube Data". *Proceedings of Science (ICRC2017)*. 968. 2017 (cit. on pp. 5, 18, 54, 55).

[28] A. Schneider (for the IceCube Collaboration). “Characterization of the Astrophysical Diffuse Neutrino Flux with IceCube High-Energy Starting Events”. *Proceedings of Science (ICRC2019)*. Vol. 1004. 2019 (cit. on pp. 5, 18, 54, 55).

[29] L. A. Fusco and F. Versari (on behalf of the ANTARES Collaboration). “Study of the high-energy diffuse neutrino signal with the ANTARES telescope”. *Proceedings of Science (ICRC2019)*. Vol. 891. 2019 (cit. on pp. 5, 18, 54, 55).

[30] K. Melis (on behalf of the KM3NeT Collaboration). “In-Situ Calibration of KM3NeT”. *Proceedings of Science (ICRC2017)*. 1059. Aug. 2017 (cit. on pp. 7, 19, 111, 112).

[31] M. Jongen. *Time calibration with the nanobeacons*. May 2017. Presentation (cit. on pp. 7, 19, 77, 111, 112).

[32] E. V. Bugaev et al. “Atmospheric muon flux at sea level, underground, and underwater”. *Phys. Rev. D* 58 (5 July 1998) (cit. on pp. 8, 20, 122, 125).

[33] J. A. Aguilar Sánchez et al. (The ANTARES Collaboration). “Measurement of the atmospheric muon flux with a 4GeV threshold in the ANTARES neutrino telescope”. *Astroparticle Physics* 33 (Mar. 2010) (cit. on pp. 8, 20, 21, 122, 125).

[34] S. Aiello et al. (The KM3NeT Collaboration). “Sensitivity of the KM3NeT/ARCA neutrino telescope to point-like neutrino sources”. *Astroparticle Physics* 111 (2019) (cit. on pp. 18, 149, 150).

[35] W. Schlosser. “Astronomische Deutung der Himmelsscheibe von Nebra”. *Sterne und Weltraum* 42 (12) (Dec. 2003) (cit. on p. 29).

[36] M. Amenomori et al. (The Tibet ASγ Collaboration). “First Detection of Photons with Energy beyond 100 TeV from an Astrophysical Source”. *Phys. Rev. Lett.* 123 (July 2019) (cit. on p. 30).

[37] P. Sokolsky (for the HiRes Collabration). “Final Results from the High solution Fly’s Eye (HiRes) Experiment”. *Nuclear Physics B - Proceedings Supplements* 212-213 (Mar. 2011) (cit. on p. 30).

[38] F. Aharonian et al. (The H.E.S.S. Collaboration). “Discovery of Very High Energy Gamma-ray Emission From Centaurus A with H.E.S.S.” *The Astrophysical Journal* 695 (1) (Mar. 2009) (cit. on p. 30).

[39] V. Hess. “Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten”. *Physikalische Zeitschrift* 13 (1912) (cit. on p. 30).

[40] A. Aab et al. (The Pierre Auger Collaboration). “Combined fit of spectrum and composition data as measured by the Pierre Auger Observatory”. *Journal of Cosmology and Astroparticle Physics* 2017 (04) (Apr. 2017) (cit. on pp. 30, 37, 38, 47).

[41] A. Aab et al. (The Pierre Auger Collaboration). *Inferences on Mass Composition and Tests of Hadronic Interactions from 0.3 to 100 EeV using the water-Cherenkov Detectors of the Pierre Auger Observatory*. 2017 (cit. on pp. 30, 47).

[42] N. L. Grigorov et al. “Measurements of particle spectra on the Proton-1,2,3 satellites”. *Yadernaya Fizika* 11 (1970) (cit. on p. 31).

[43] K. Asakimori et al. *Proceedings, 23rd International Cosmic Ray Conference*. Vol. 2. July 1993 (cit. on p. 31).

[44] T. V. Danilova et al. *Proceedings, 15th International Cosmic Ray Conference*. Vol. 8. 1977 (cit. on p. 31).

[45] Y. A. Fomin et al. "Energy spectrum of cosmic rays at energies of $5 \cdot 10^{15}$ - $5 \cdot 10^{17}$ eV". *Proceedings, 22nd International Cosmic Ray Conference*. 2. Aug. 1991 (cit. on p. 31).

[46] M. Amenomori et al. (The Tibet ASγ Collaboration). "The Cosmic-Ray Energy Spectrum between $10^{14.5}$ and $10^{16.3}$ eV Covering the Knee Region". *The Astrophysical Journal* 461 (Apr. 1996) (cit. on p. 31).

[47] N. Nagano et al. "Energy spectrum of primary cosmic rays between $10^{14.5}$ and 10^{18}eV". *Journal of Physics G: Nuclear Physics* 10 (9) (Sept. 1984) (cit. on p. 31).

[48] F. Arqueros et al. (The HEGRA Collaboration). "Energy Spectrum and Chemical Composition of Cosmic Rays between 0.3 and 10 PeV determined from the Cherenkov-Light and Charged-Particle distributions in Air Showers". *Astronomy & Astrophysics* 359 (682) (2000) (cit. on p. 31).

[49] M. A. K. Glasmacher et al. "The cosmic ray energy spectrum between 10^{14} and 10^{16} eV". *Astroparticle Physics* 10 (4) (1999) (cit. on p. 31).

[50] T. Antoni et al. (The KASCADE Collaboration). "KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems". *Astroparticle Physics* 24 (1) (2005) (cit. on p. 31).

[51] M. Amenomori et al. (The Tibet ASγ Collaboration). "Zenith angle dependence of the size spectrum of air showers around the knee observed with the Tibet air shower array". *Proceedings, 30th International Cosmic Ray Conference*. Vol. 4. July 2007 (cit. on p. 31).

[52] D. Bird et al. (The Fly's Eye Collaboration). "The Cosmic-Ray Energy Spectrum Observed by the Fly's Eye". *The Astrophysical Journal* 424 (Mar. 1994) (cit. on p. 31).

[53] W. D. Apel et al. (The KASCADE-Grande Collaboration). "Kneelike Structure in the Spectrum of the Heavy Component of Cosmic Rays Observed with KASCADE-Grande". *Physical Review Letters* 107 (17 Oct. 2011) (cit. on p. 31).

[54] M. Takeda et al. "Energy determination in the Akeno Giant Air Shower Array experiment". *Astroparticle Physics* 19 (4) (2003) (cit. on p. 31).

[55] R. U. Abbasi et al. (The High Resolution Fly's Eye Collaboration). "First Observation of the Greisen-Zatsepin-Kuzmin Suppression". *Phys. Rev. Lett.* 100 (10 Mar. 2008) (cit. on p. 31).

[56] J. Abraham et al. (The Pierre Auger Collaboration). "Observation of the Suppression of the Flux of Cosmic Rays above 4×10^{19} eV". *Physical Review Letters* 101 (6 Aug. 2008) (cit. on p. 31).

[57] Y. Tsunesada et al. (Telescope Array). "Highlights from Telescope Array". *Proceedings, 32nd International Cosmic Ray Conference*. July 2011 (cit. on p. 31).

[58] P. Abreu et al. (The Pierre Auger Collaboration). "The Pierre Auger Observatory I: The Cosmic Ray Energy Spectrum and Related Measurements". *Proceedings, 32nd International Cosmic Ray Conference*. 2011 (cit. on p. 31).

[59] K. Rawlins et al. (for the IceCube Collaboration). "Cosmic ray spectrum and composition from three years of IceTop and IceCube". *Journal of Physics: Conference Series* 718 (May 2016) (cit. on p. 31).

[60] T. K. Gaisser. "Spectrum of cosmic-ray nucleons, kaon production, and the atmospheric muon charge ratio". *Astroparticle Physics* 35 (12) (July 2012) (cit. on pp. 31, 32, 45, 46, 157).

[61] O. Adriani et al. "Observations of the 2006 December 13 and 14 solar particle events in the 80 MeV/n - 3 GeV/n range from space with the PAMELA detector". *The Astrophysical Journal* 742 (2) (Nov. 2011) (cit. on p. 32).

[62] R. Alves Batista et al. "Effects of uncertainties in simulations of extragalactic UHECR propagation, using CRPropa and SimProp". *JCAP* 1510 (10) (2015) (cit. on p. 32).

[63] W. Pauli. *Offener Brief and die Gruppe der Radioaktiven bei der Gauverein-Tagung zu Tübingen*. Dec. 1930 (cit. on p. 32).

[64] C. Cowan, F. Reines, et al. "Detection of the Free Neutrino: A Confirmation". *Science* 124 (July 1956) (cit. on pp. 32, 44).

[65] G. Danby et al. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos". *Physical Review Letters* 9 (1) (July 1962) (cit. on p. 32).

[66] K. Kodama et al. (The DONUT Collaboration). "Observation of tau neutrino interactions". *Physics Letters B* 504 (3) (Apr. 2001) (cit. on p. 32).

[67] D. Buskulic et al. (The ALEPH Collaboration). "Update of electroweak parameters from Z decays". *Zeitschrift für Physik C Particles and Fields* 60 (1) (Mar. 1993) (cit. on p. 32).

[68] B. Abbott et al. (The LIGO and Virgo Collaborations). "Observation of Gravitational Waves from a Binary Black Hole Merger". *Physical Review Letters* 116 (6) (Feb. 2016) (cit. on p. 33).

[69] B. Abbott et al. (The LIGO and Virgo Collaborations). "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs". *Physical Review X* 9 (3) (Sept. 2019) (cit. on p. 33).

[70] L. Nellen, K. Mannheim, and P. L. Biermann. "Neutrino production through hadronic cascades in AGN accretion disks". *Phys. Rev. D* 47 (12 June 1993) (cit. on p. 34).

[71] J. K. Becker. "High-energy neutrinos in the context of multimessenger astrophysics". *Physics Reports* 458 (4-5) (Mar. 2008) (cit. on p. 34).

[72] S. Wakely and D. Horan. "TeVCat: An online catalog for Very High Energy Gamma-Ray Astronomy". *International Cosmic Ray Conference*. Vol. 3. International Cosmic Ray Conference. Jan. 2008 (cit. on pp. 35, 211).

[73] A. Reisenegger. *Origin and evolution of neutron star magnetic fields*. 2003 (cit. on p. 35).

[74] A. Bykov et al. "VNIIEF achievements on ultra-high magnetic fields generation". *Physica B: Condensed Matter* 294-295 (2001). Proceedings of the Sixth International Symposium on Research in High Magnetic Fields (cit. on p. 35).

[75] Véron-Cetty, M.-P. and Véron, P. "A catalogue of quasars and active nuclei: 12th edition". *A&A* 455 (2) (2006) (cit. on p. 38).

[76] *VizieR catalogue access tool, CDS, Strasbourg, France*. `https://vizier.u-strasbg.fr/viz-bin/VizieR`. Accessed: 2020-01-17 (cit. on p. 38).

[77] J. Abraham et al. (The Pierre Auger Collaboration). "Correlation of the Highest-Energy Cosmic Rays with Nearby Extragalactic Objects". *Science* 318 (5852) (2007) (cit. on p. 38).

[78] A. Aab et al. (The Pierre Auger Collaboration). "Large-scale Cosmic-Ray Anisotropies above 4 EeV Measured by the Pierre Auger Observatory". *The Astrophysical Journal* 868 (1) (Nov. 2018) (cit. on p. 37).

[79] E. Fermi. "On the Origin of the Cosmic Radiation". *Phys. Rev.* 75 (8 Apr. 1949) (cit. on p. 38).

[80] A. M. Bykov et al. "Cosmic Ray Production in Supernovae". *Space Science Reviews* 214 (1) (Jan. 2018) (cit. on p. 38).

[81] J. G. Learned and K. Mannheim. "High-Energy Neutrino Astrophysics". *Annual Review of Nuclear and Particle Science* 50 (1) (2000) (cit. on p. 40).

[82] A. M. Hillas. "The Origin of Ultra-High-Energy Cosmic Rays". *Annual Review of Astronomy and Astrophysics* 22 (1) (1984) (cit. on pp. 41, 42).

[83] F. Fraschetti. "On the acceleration of ultra-high-energy cosmic rays". *Phil. Trans. R. Soc. A* 366 (1884) (Sept. 2008) (cit. on pp. 41, 42).

[84] E. Vitagliano, I. Tamborra, and G. Raffelt. "Grand Unified Neutrino Spectrum at Earth" (2019) (cit. on pp. 43, 44).

[85] B. Follin et al. "First Detection of the Acoustic Oscillation Phase Shift Expected from the Cosmic Neutrino Background". *Physical Review Letters* 115 (9) (Aug. 2015) (cit. on p. 43).

[86] J. Bernstein, L. S. Brown, and G. Feinberg. "Cosmological helium production simplified". *Rev. Mod. Phys.* 61 (1 Jan. 1989) (cit. on p. 43).

[87] M. Betti et al. (The PTOLEMY Collaboration). "Neutrino physics with the PTOLEMY project: active neutrino properties and the light sterile case". *Journal of Cosmology and Astroparticle Physics* 2019 (07) (July 2019) (cit. on p. 43).

[88] B. T. Cleveland et al. "Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector". *The Astrophysical Journal* 496 (1) (Mar. 1998) (cit. on pp. 44, 57).

[89] E. Vitagliano, J. Redondo, and G. Raffelt. *Solar neutrinos at keV energies: thermal flux*. 2017 (cit. on p. 44).

[90] M. Honda et al. "Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data". *Phys. Rev. D* 75 (4 Feb. 2007) (cit. on pp. 45, 46, 157).

[91] R. Enberg, M. H. Reno, and I. Sarcevic. "Prompt neutrino fluxes from atmospheric charm". *Phys. Rev. D* 78 (4 Aug. 2008) (cit. on p. 46).

[92] R. Gauld et al. "Charm production in the forward region: constraints on the small-x gluon and backgrounds for neutrino astronomy". *JHEP* 11 (2015) (cit. on pp. 46, 157).

[93] M. G. Aartsen et al. (The IceCube Collaboration). "Development of a general analysis and unfolding scheme and its application to measure the energy spectrum of atmospheric neutrinos with IceCube". *The European Physical Journal C* 75 (3) (2015), p. 116 (cit. on p. 46).

[94] M. G. Aartsen et al. (The IceCube Collaboration). "Measurement of the Atmospheric ν_e Flux in IceCube". *Phys. Rev. Lett.* 110 (15 Apr. 2013) (cit. on p. 46).

[95] M. G. Aartsen et al. (The IceCube Collaboration). "Measurement of the Atmospheric ν_e Spectrum with IceCube". *Phys. Rev.* D91 (2015) (cit. on p. 46).

[96] S. Aguilar et al. (The ANTARES collaboration). "Measurement of the atmospheric $\nu_\mu$energy spectrum from 100 GeV to 200 TeV with the ANTARES telescope". *The European Physical Journal C* 73 (10) (2013), p. 2606 (cit. on p. 46).

[97] K. Daum et al. (The Fréjus Collaboration). "Determination of the atmospheric neutrino spectra with the Fréjus detector". *Zeitschrift für Physik C Particles and Fields* 66 (3) (Sept. 1995) (cit. on p. 46).

[98] E. Waxman and J. Bahcall. "High energy neutrinos from astrophysical sources: An upper bound". *Physical Review D* 59 (2) (Dec. 1998) (cit. on p. 46).

[99] J. Bahcall and E. Waxman. "High energy astrophysical neutrinos: The upper bound is robust". *Phys. Rev. D* 64 (2 June 2001) (cit. on p. 46).

[100] E. Waxman. "Astrophysical sources of high energy neutrinos". *Nuclear Physics B - Proceedings Supplements* 118 (2003). Proceedings of the XXth International Conference on Neutrino Physics and Astrophysics (cit. on pp. 47, 48).

[101] K. Murase and J. F. Beacom. "Neutrino background flux from sources of ultrahigh-energy cosmic-ray nuclei". *Phys. Rev. D* 81 (12 June 2010) (cit. on p. 47).

[102] L. Anchordoqui et al. "High energy neutrinos from astrophysical accelerators of cosmic ray nuclei". *Astroparticle Physics* 29 (1) (Feb. 2008) (cit. on p. 47).

[103] F. W. Stecker et al. "High-energy neutrinos from active galactic nuclei". *Phys. Rev. Lett.* 66 (21 May 1991) (cit. on pp. 47, 48).

[104] F. W. Stecker et al. "Erratum: "High-energy neutrinos from active galactic nuclei" [Phys. Rev. Lett. 66, 2697 (1991)]". *Phys. Rev. Lett.* 69 (18 Nov. 1992) (cit. on p. 48).

[105] F. W. Stecker. "Note on high-energy neutrinos from active galactic nuclei cores". *Phys. Rev. D* 72 (10 Nov. 2005) (cit. on p. 48).

[106] J. Becker, B. P.L., and R. Rhode. "The diffuse neutrino flux from FR-II radio galaxies and blazars: A source property based estimate". *Astroparticle Physics* 23 (4) (2005) (cit. on p. 48).

[107] M. Mandelartz and J. B. Tjus. "Prediction of the diffuse neutrino flux from cosmic ray interactions near supernova remnants". *Astroparticle Physics* 65 (2015) (cit. on p. 48).

[108] E. Waxman and J. N. Bahcall. "High-energy neutrinos from cosmological gamma-ray burst fireballs". *Phys. Rev. Lett.* 78 (1997) (cit. on p. 48).

[109] K. Murase, S. Inoue, and S. Nagataki. "Cosmic Rays above the Second Knee from Clusters of Galaxies and Associated High-Energy Neutrino Emission". *The Astrophysical Journal* 689 (2) (Nov. 2008) (cit. on p. 48).

[110] A. Loeb and E. Waxman. "The cumulative background of high energy neutrinos from starburst galaxies". *Journal of Cosmology and Astroparticle Physics* 2006 (05) (May 2006) (cit. on p. 48).

[111] A. Palladino et al. "Interpretation of the diffuse astrophysical neutrino flux in terms of the blazar sequence". *Astrophys. J.* 871 (1) (2019) (cit. on p. 48).

[112] P. B. Denton and I. Tamborra. "Exploring the Properties of Choked Gamma-ray Bursts with IceCubes High-energy Neutrinos". *Astrophys. J.* 855 (1) (2018) (cit. on p. 48).

[113] K. Møller, P. B. Denton, and I. Tamborra. "Cosmogenic Neutrinos Through the GRAND Lens Unveil the Nature of Cosmic Accelerators". *JCAP* 1905 (05) (2019) (cit. on p. 48).

[114] T. Kashti and E. Waxman. "Astrophysical Neutrinos: Flavor Ratios Depend on Energy". *Phys. Rev. Lett.* 95 (18 Oct. 2005) (cit. on p. 48).

[115] B. Pontecorvo. "Mesonium and Antimesonium". *Zh. Eksp. Teor. Fiz.* 6 (2) (Feb. 1957) (cit. on p. 49).

[116] I. Esteban et al. "Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of $\theta 23$, δCP, and the mass ordering". *Journal of High Energy Physics* 2019 (1) (Jan. 2019) (cit. on p. 49).

[117] *NuFit 4.1*. `www.nu-fit.org`. 2019 (cit. on p. 49).

[118] G. Pagliaroli et al. "Improved analysis of SN1987A antineutrino events". *Astroparticle Physics* 31 (3) (Apr. 2009) (cit. on p. 50).

[119] M. Aartsen et al. (The IceCube Collaboration). "Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert". *Science* 361 (6398) (2018) (cit. on pp. 50, 52).

[120] S. Garrappa et al. (The Fermi-LAT, ASAS-SN and IceCube Collaborations). "Investigation of two Fermi-LAT gamma-ray blazars coincident with high-energy neutrinos detected by IceCube". *Astrophys. J.* 880 (2) (2019) (cit. on pp. 51, 52).

[121] N. Sahakyan. "Origin of the multiwavelength emission of PKS 0502+049". *Astron. Astrophys.* 622 (2019) (cit. on p. 52).

[122] W. Atwood et al. (The Fermi-LAT Collaboration). "The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission". *The Astrophysical Journal* 697 (2) (June 2009) (cit. on p. 52).

[123] E. Visser. "Neutrinos From the Milky Way". PhD thesis. Leiden University, May 2015 (cit. on p. 52).

[124] M. G. Aartsen et al. (The IceCube Collaboration). "Time-integrated Neutrino Source Searches with 10 years of IceCube Data" (2019) (cit. on pp. 52, 53).

[125] P. B. Denton, D. Marfatia, and T. J. Weiler. "The Galactic Contribution to IceCube's Astrophysical Neutrino Flux". *JCAP* 1708 (08) (2017) (cit. on p. 52).

[126] J. Aublin, G. Illuminati, and S. Navas (for the Antares Collaboration). "Searches for point-like sources of cosmic neutrinos with 11 years of ANTARES data". *Proceedings of Science (ICRC2019)*. 920. 2019 (cit. on p. 53).

[127] M. G. Aartsen et al. (The IceCube Collaboration). "Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector". *Science* 342 (6161) (2013) (cit. on p. 54).

[128] M. Aartsen et al. (The IceCube Collaboration). "Differential limit on the extremely-high-energy cosmic neutrino flux in the presence of astrophysical background from nine years of IceCube data". *Phys. Rev.* D98 (6) (2018) (cit. on pp. 54, 55).

[129] A. Aab et al. (The Pierre Auger Collaboration). "Improved limit to the diffuse flux of ultrahigh energy neutrinos from the Pierre Auger Observatory". *Phys. Rev. D* 91 (9 May 2015) (cit. on p. 54).

[130] P. W. Gorham et al. "Observational constraints on the ultrahigh energy cosmic neutrino flux from the second flight of the ANITA experiment". *Phys. Rev. D* 82 (2 July 2010) (cit. on p. 54).

[131] *The highest energy neutrinos in IceCube*. `https://icecube.wisc.edu/viewer/he_neutrinos#`. Accessed: 2020-02-04 (cit. on p. 55).

[132] M. G. Aartsen et al. (The IceCube Collaboration). "A Combined Maximum-Likelihood Analysis of the High-Energy Astrophysical Neutrino Flux Measured with IceCube". *The Astrophysical Journal* 809 (1) (Aug. 2015) (cit. on pp. 55, 56).

[133] Y. Fukuda et al. (The Super-Kamiokande Collaboration). "Evidence for Oscillation of Atmospheric Neutrinos". *Physical Review Letters* 81 (8) (Aug. 1998) (cit. on p. 57).

[134] Q. R. Ahmad et al. (The SNO Collaboration). "Measurement of the Rate of $\nu_e + d\top + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory". *Physical Review Letters* 87 (7) (July 2001) (cit. on p. 57).

[135] S. Adrian-Martinez et al. (The KM3NeT Collaboration). "Letter of intent for KM3NeT 2.0". *J. Phys.* G43 (8) (2016) (cit. on pp. 57, 58, 101, 159).

[136] M. Jongen. "Particle Physics in the Mediterranean Sea". PhD thesis. Universiteit van Amsterdam, May 2018 (cit. on p. 57).

[137] A. V. Akindinov et al. (The KM3NeT Collaboration). "Letter of interest for a neutrino beam from Protvino to KM3NeT/ORCA". *The European Physical Journal C* 79 (9) (2019) (cit. on p. 58).

[138] C. Adloff et al. (The H1 Collaboration). "Deep inelastic inclusive e p scattering at low x and a determination of alpha(s)". *Eur. Phys. J.* C21 (2001) (cit. on pp. 58, 222).

[139] S. Chekanov et al. (The ZEUS Collaboration). "Measurement of the neutral current cross-section and F(2) structure function for deep inelastic e + p scattering at HERA". *Eur. Phys. J.* C21 (2001) (cit. on pp. 58, 222).

[140] A. Benvenuti et al. (The BCDMS Collaboration). "A high statistics measurement of the proton structure functions F2(x, Q2) and R from deep inelastic muon scattering at high Q2". *Physics Letters B* 223 (3) (1989) (cit. on pp. 58, 222).

[141] B. Abbott et al. (The D0 Collaboration). "High-p_T jets in $\bar{p}p$ collisions at $\sqrt{s}$ = 630 GeV and 1800 GeV". *Phys. Rev.* D64 (2001) (cit. on pp. 58, 222).

[142] W. G. Seligman et al. "Improved Determination of α_s From Neutrino-Nucleon Scattering". *Phys. Rev. Lett.* 79 (7 Aug. 1997) (cit. on pp. 58, 222).

[143] F. Abe et al. (The CDF Collaboration). "Measurement of the Lepton Charge Asymmetry in W-Boson Decays Produced in $p\overline{p}$ Collisions". *Phys. Rev. Lett.* 81 (26 Dec. 1998) (cit. on pp. 58, 222).

[144] A. Buckley, J. Ferrando, et al. "LHAPDF6: parton density access in the LHC precision era". *The European Physical Journal C* 75 (3) (Mar. 2015) (cit. on pp. 58, 222).

[145] J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". *JHEP* 07 (2002) (cit. on pp. 58, 222).

[146] R. Gandhi et al. "Ultrahigh-energy neutrino interactions". *Astroparticle Physics* 5 (2) (Aug. 1996) (cit. on pp. 58, 60).

[147] S. L. Glashow. "Resonant Scattering of Antineutrinos". *Phys. Rev.* 118 (1 Apr. 1960) (cit. on p. 59).

[148] P. Cherenkov. "Visible emission of clean liquids by action of γ radiation". *Doklady Akademii Nauk SSSR* 2 (451) (1934) (cit. on p. 61).

[149] J. D. Jackson. *Classical Electrodynamics*. 3rd ed. John Wiley & Sons, 1999 (cit. on p. 61).

[150] I. Tamm and I. Frank. "Coherent radiation of fast electrons in a medium". *Doklady Akademii Nauk SSSR* 14 (107) (1937) (cit. on p. 61).

[151] C. Kopper. "Performance Studies for the KM3NeT Neutrino Telescope". PhD thesis. Friedrich-Alexander-Universität Erlangen-Nürnberg, Mar. 2010 (cit. on pp. 62, 64).

[152] R. Mirani. "Parametrisation of EM-showers in the ANTARES detector-volume". MA thesis. Universiteit van Amsterdam, Jan. 2002 (cit. on pp. 62, 63).

[153] B. Hartmann. "Reconstruction of Neutrino-Induced Hadronic and Electromagnetic Showers with the ANTARES Experiment". PhD thesis. Friedrich-Alexander-Universität Erlangen-Nürnberg, June 2006 (cit. on p. 63).

[154] V. Niess. "Underwater acoustic detection of ultra-high energy neutrinos in ANTARES". Thesis. Université de la Méditerranée - Aix-Marseille II, Sept. 2005 (cit. on pp. 63, 64).

[155] M. De Jong and P. Kooijman. *JPP:JDeltaRays*. Computer Software (cit. on pp. 64, 66).

[156] S. Klimushin, E. Bugaev, and I. Sokalski. "Precise parametrizations of muon energy losses in water". *Proceedings of ICRC (2001)*. Vol. 1009. Aug. 2001 (cit. on p. 65).

[157] J. G. Learned and S. Pakvasa. "Detecting nu-tau oscillations at PeV energies". *Astroparticle Physics* 3 (3) (May 1995) (cit. on p. 66).

[158] S. Aguilar et al. (The ANTARES Collaboration). "Transmission of light in deep sea water at the site of the ANTARES Neutrino Telescope". *Astropart. Phys.* 23 (2005) (cit. on p. 67).

[159] W. Schuster. "Measurement of the Optical Properties of the Deep Mediterranean and the ANTARES Detector Medium". PhD thesis. University of Oxford, 2002 (cit. on p. 67).

[160] L. Henyey and J. Greenstein. "Diffuse Radiation in the Galaxy". *Astrophysical Journal* 93 (1941) (cit. on p. 67).

[161] L. F. Rayleigh. "On the electromagnetic theory of light". *Phylosophical Magazine Series* 5 (73) (Dec. 1881) (cit. on p. 67).

[162] R. Bruijn and D. Van Eijk (for the KM3NeT Collaboration). "The KM3NeT Multi-PMT Digital Optical Module". *Proceedings of Science (ICRC2015)*. 1157. Aug. 2015 (cit. on p. 67).

[163] S. Viola et al. (for the KM3NeT Collaboration). "Acoustic positioning system for KM3NeT". *Proceedings of Science (ICRC2015)*. 1169. Aug. 2015 (cit. on p. 68).

[164] S. Viola (for the KM3NeT Collaboration). "KM3NeT acoustic positioning and detection system". *EPJ Web Conf.* 216 (2019) (cit. on p. 68).

[165] S. Aiello et al. (The KM3NeT Collaboration). "Characterisation of the Hamamatsu photomultipliers for the KM3NeT Neutrino Telescope". *Journal of Instrumentation* 13 (05) (May 2018) (cit. on pp. 70–72, 75, 114).

[166] *Photomultiplier tube R12199*. `https://www.hamamatsu.com/jp/en/product/type/R12199/index.html`. Accessed: 2019-07-11 (cit. on p. 70).

[167] "Detector simulations for KM3NeT". June 2013. KM3NeT internal note (cit. on p. 71).

[168] O. Kalekin. Personal communication. Oct. 2019 (cit. on pp. 70, 71).

[169] J. Brunner. "Upgrade of ^{40}K simulation". May 2006. Antares Internal Note (cit. on pp. 72, 83, 224).

[170] E. Fermi. "Versuch einer Theorie der β-Strahlen". *Zeitschrift für Physik* 88 (3) (Mar. 1934) (cit. on p. 72).

[171] D. Samtleben. Personal communication (cit. on pp. 72, 114).

[172] M. Ageron et al. (The KM3NeT Collaboration). "Dependence of atmospheric muon flux on seawater depth measured with the first KM3NeT detection units". *Eur. Phys. J.* C80 (2) (2020) (cit. on pp. 72, 119).

[173] S. Adrian-Martinez et al. (The KM3NeT Collaboration). "Deep sea tests of a prototype of the KM3NeT digital optical module". *Eur. Phys. J.* C74 (9) (2014) (cit. on p. 73).

[174] M. Crouch. "An Improved World Survey Expression for Cosmic Ray Vertical Intensity VS. Depth in Standard Rock". *Proceedings, 20th International Cosmic Ray Conference*. Vol. 6. International Cosmic Ray Conference. Jan. 1987 (cit. on p. 74).

[175] C. Waltham et al. (The SNO Collaboration). "Through-Going Muons in the Sudbury Neutrino Observatory". *Proceedings, 27th International Cosmic Ray Conference*. Vol. 3. International Cosmic Ray Conference. 2001 (cit. on p. 74).

[176] Y. Andreyev et al. "Muon Intensity from the Baksan Underground Scintillation Telescope". *Proceedings, 20th International Cosmic Ray Conference*. Vol. 6. International Cosmic Ray Conference. 1987 (cit. on p. 74).

[177] M. Aglietta et al. (The LVD Collaboration). "Neutrino-induced and atmospheric single-muon fluxes measured over five decades of intensity by LVD at Gran Sasso Laboratory". *Astroparticle Physics* 3 (4) (1995) (cit. on p. 74).

[178] C. Berger et al. (The Fréjus Collaboration). "Experimental study of muon bundles observed in the Fréjus detector". *Phys. Rev. D* 40 (7 Oct. 1989) (cit. on p. 74).

[179] M. Ambrosio et al. (The MACRO Collaboration). "Vertical muon intensity measured with MACRO at the Gran Sasso laboratory". *Phys. Rev. D* 52 (7 Oct. 1995) (cit. on p. 74).

[180] R. Bruijn. "The Antares Neutrino Telescope: Performance Studies and Analysis of First Data". PhD thesis. Universiteit van Amsterdam, Mar. 2008 (cit. on p. 74).

[181] B. Bakker. "Trigger studies for the Antares and KM3NeT neutrino telescopes". PhD thesis. Universiteit van Amsterdam, July 2011 (cit. on pp. 76, 212).

[182] R. Bormuth. "Description of the two standard trigger algorithms in the JPP software package". Dec. 2015. KM3NeT internal note (cit. on p. 76).

[183] M. Bouwhuis (on the behalf of the KM3NeT Collaboration). *Technical Design Report KM3NeT Time Calibration*. `https://indico.cern.ch/event/570416/contributions/2307233/attachments/1340384/2018328/KM3NeT_CALIB_2015_010-TDR_TimeCalibration.pdf`. 2005 (cit. on pp. 76, 77).

[184] M. Bouwhuis. "Time synchronization and time calibration in KM3NeT". *Proceedings of Science (ICRC2015)*. 1170. 2015 (cit. on p. 77).

[185] C. Hugon (on behalf of the KM3NeT Collaboration). "GEANT4 simulation of optical modules in neutrino telescopes". *Proceedings of Science (ICRC2015)*. 1106. July 2015 (cit. on p. 83).

[186] S. Agostinelli et al. "Geant4-a simulation toolkit". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506 (3) (2003) (cit. on pp. 83, 86).

[187] D. Motta and S. Schönert. "Optical properties of bialkali photocathodes". *Nuclear Instruments and Methods in Physics Research A* 539 (Feb. 2005) (cit. on p. 83).

[188] G. Carminati et al. "Atmospheric MUons from PArametric formulas: a fast GEnerator for neutrino telescopes (MUPAGE)". *Computer Physics Communications* 179 (12) (2008) (cit. on p. 84).

[189] G. Battistoni, C. Forti, and J. Ranft. "Study of the high energy cosmic ray cascades using the dual parton model". *Astroparticle Physics* 3 (2) (1995) (cit. on p. 84).

[190] D. Bailey. "Monte Carlo tools and analysis methods for understanding the ANTARES experiment and predicting its sensitivity to dark matter". PhD thesis. Oxford U., 2002 (cit. on p. 84).

[191] P. Antonioli et al. "A three-dimensional code for muon propagation through the rock: MUSIC". *Astroparticle Physics* 7 (4) (1997) (cit. on pp. 85, 86).

[192] M. De Jong. "The probability density function of the arrival time of light". Mar. 2012. ANTARES internal note (2010-002) (cit. on pp. 85, 229).

[193] M. De Jong. "Multi-dimensional interpolations in C++". July 2019 (cit. on pp. 86, 176, 232).

[194] M. De Jong. *PMT simulation*. `https://indico.nikhef.nl/event/717/session/0/contribution/7/material/slides/`. Mar. 2017. Presentation (cit. on pp. 86, 88).

[195] W. Metzger. *Statistical Methods in Data Analysis*. HEN-343. Faculteit der Natuurwetenschappen, Wiskunde en Informatica, Katholieke Universiteit Nijmegen, Aug. 2002 (cit. on pp. 89, 90, 146, 147, 225, 233).

[196] H. Cramér. *Mathematical Methods of Statistics*. Princeton Univ. Press, 1946. ISBN: 0-691-08004-6 (cit. on p. 89).

[197] C. Rao. "Information and the accuracy attainable in the estimation of statistical parameters". *Bulletin of the Calcutta Mathematical Society.* 37 (1945) (cit. on p. 89).

[198] P. J. Huber. "Robust Estimation of a Location Parameter". *Ann. Math. Statist.* 35 (1) (Mar. 1964) (cit. on p. 91).

[199] R. Anderson. *Modern Methods for Robust Regression*. 152. SAGE Publications, 2008. ISBN: 978-1412940726 (cit. on p. 91).

[200] P. Huber and E. Ronchetti. *Robust Statistics*. 1st ed. Wiley, Mar. 1981 (cit. on p. 92).

[201] K. Clarkson and D. Woodruff. "Sketching for M -Estimators: A Unified Approach to Robust Regression". *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms* 2015 (Oct. 2015) (cit. on p. 92).

[202] A. Heijboer. "Shower Direction Reconstruction: aashowerfit". 2014. KM3NeT internal note (cit. on p. 92).

[203] F. James. *MINUIT - Function Minimization and Error Analysis*. `https://root.cern.ch/doc/master/classTMinuit.html`. 1994. CERN Program Library Long Writeup D506 (cit. on p. 93).

[204] K. Melis, A. Heijboer, and M. de Jong (on behalf of the KM3NeT Collaboration). "KM3NeT/ARCA Event Reconstruction Algorithms". *Proceedings of Science (ICRC2017)*. 950. Aug. 2017 (cit. on pp. 94, 96).

[205] K. Levenberg. "A Method for the Solution of Certain Non-Linear Problems in Least Squares". *Quarterly of Applied Mathematics* 2 (2) (1944) (cit. on p. 97).

[206] D. W. Marquardt. "An Algorithm for Least-Squares Estimation of Nonlinear Parameters". *Journal of the Society for Industrial and Applied Mathematics* 11 (2) (1963) (cit. on p. 97).

[207] T. Heid. "Characterizing the diffuse neutrino flux with the future KM3NeT/ARCA detector". PhD thesis. Friedrich-Alexander-Universität Erlangen-Nürnberg, Jan. 2019 (cit. on p. 101).

[208] A. Heijboer. Personal communication. Feb. 2020 (cit. on p. 103).

[209] S. I. Klimushin, E. V. Bugaev, and I. A. Sokalski. "Parametrization of atmospheric muon angular flux underwater". *Phys. Rev. D* 64 (1 May 2001) (cit. on p. 122).

[210] E. Berbee. Personal communication. July 2018 (cit. on p. 133).

[211] K. Melis. *KM3NeT Calibration*. Oct. 2017. Mediterranean Antarctic Neutrino Telescope Symposium (MANTS) meeting (cit. on p. 134).

[212] J. Neyman, E. S. Pearson, and K. Pearson. "On the problem of the most efficient tests of statistical hypotheses". *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 231 (694-706) (1933) (cit. on p. 146).

[213] S. Karlin and H. Rubin. "The Theory of Decision Procedures for Distributions with Monotone Likelihood Ratio". *Annals of Mathematical Statistics* 27 (2) (June 1956) (cit. on p. 147).

[214] A. Albert et al. (The ANTARES and IceCube Collaborations). "ANTARES and IceCube Combined Search for Neutrino Point-like and Extended Sources in the Southern Sky". *The Astrophysical Journal* 892 (2) (Apr. 2020) (cit. on pp. 149, 150).

[215] *Atmospheric Neutrino Flux Tables by HKKMS (2006)*. `http://www.icrr.u-tokyo.ac.jp/~mhonda/nflx2006/index.html`. (Frejus site without mountain over the detector, solar-min), Accessed: 2020-03-08 (cit. on p. 157).

[216] L. Rottoli. *PromptNuFlux: Prompt atmospheric neutrino flux calculator*. Nov. 2015 (cit. on p. 157).

[217] W. Press et al. *Numerical Recipes in C++*. 2nd ed. Cambridge University Press, 1992 (cit. on pp. 176, 189, 199, 232).

[218] B. Delauney. “Sur la sphère vide. A la mémoire de Georges Vorono”. *Bulletin de lÀcadémie des Sciences de lÙRSS* 1708 (6) (1934) (cit. on pp. 202, 203).

[219] M. Jongen. `https://indico.nikhef.nl/event/844/session/0/contribution/0/material/slides/0.pdf`. July 2017. Presentation (cit. on p. 210).

[220] P. Wallace. “The SLALIB Library”. *Astronomical Data Analysis Software and Systems III*. Ed. by D. Crabtree, R. Hanisch, and J. Barnes. Vol. 61. Astronomical Society of the Pacific Conference Series. Jan. 1994 (cit. on p. 211).

[221] T. Robitaille et al. (The Astropy Collaboration). “Astropy: A community Python package for astronomy”. *Astronomy & Astrophysics* 558, A33 (Oct. 2013), A33 (cit. on p. 211).

[222] A. Price-Whelan et al. (The Astropy Collaboration). “The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package”. *aj* 156 (3), 123 (Sept. 2018) (cit. on p. 211).

[223] A. C. Vincent, C. A. Argüelles, and A. Kheirandish. “High-energy neutrino attenuation in the Earth and its associated uncertainties”. *Journal of Cosmology and Astroparticle Physics* 2017 (11) (Nov. 2017) (cit. on p. 213).

[224] K. Melis. “Reconstruction of High-energy Neutrino-induced Particle Showers in KM3NeT”. MA thesis. Universiteit van Amsterdam, Nov. 2014 (cit. on p. 214).

[225] C. Hugon. *Physics, experiments and applications*. `https://agenda.infn.it/event/9583/contributions/80410/attachments/58143/68560/2Physicsandapplications.pdf`. Apr. 2015. Presentation (cit. on p. 224).

[226] M. De Jong. *Comparison km3-JSirene*. `https://indico.cern.ch/event/726424/contributions/2989127/attachments/1642080/2623400/JSirene_-_km3.pdf`. May 2018. Presentation (cit. on p. 229).

[227] A. M. Dziewonski and D. L. Anderson. “Preliminary reference Earth model”. *Physics of the Earth and Planetary Interiors* 25 (4) (1981) (cit. on pp. 235, 236).

[228] G. Bomans. *De brandmeester*. Van Oorschot, 2021 (cit. on p. 275).

Terms and Acronyms

AGN core region Core region of an active galactic nucleus, i.e. the black hole in its centre with (optionally) the surrounding accretion disc. See pp. 37

AGN hot spot Region of enhanced radio emission, where the matter of an AGN jet interacts with clouds of matter. See pp. 37

AGN jet Stream of ultrarelativistic particles emerging along the rotation axes of the accretion disc surrounding a black hole in a AGN core. See pp. 37

AGN lobe Region of enhanced radio emission surrounding AGN jets. See pp. 37

Cherenkov threshold Lower limit on the energy a charged particle should have to emit Cherenkov radiation. See pp. 62, 65

DOM time offset Time offset accounting for the phase difference between the clock in a DOM and the master clock, as well as the mean time delay of the 31 PMTs due to the PMT transit time and electrical wires inside the DOM. See pp. 76, 77, 127, 129–133

Glashow resonance Interaction of an electron antineutrino with an electron, creating a W-boson, which subsequently decays into an antineutrino-lepton pair or pion. See pp. 58, 59

Greisen-Zatsepin-Kuzmin limit (GZK) Upper limit on the energy of cosmic ray protons can have to reach the Earth from the nearest galaxies. Above this energy, interactions of the cosmic rays with the cosmic microwave background are too strong. See pp. 32

HTR calibration inter-DOM time calibration technique based on the distributions of hit time residuals See pp. 132, 133

Hillas criterion Upper limit on the maximum energy of cosmic rays produced in first order Fermi accel-

eration in the vicinity of a source. This limit corresponds to the energy above which the gyroradius of the cosmic rays is too large to be confined by the magnetic field of the source. See pp. 41

JSirene Software package used to simulate the emission, propagation and detection of light induced by charged particles propagating through seawater. See pp. ii, 85

KM3 Software package used to simulate the emission, propagation and detection of light induced by charged particles propagating through seawater. See pp. ii, 85, 86

L0 runs Data taking run, in which all L0 timeslices are written to disc See pp. 76

L0 timeslice Data package containing all recorded L0s in a time period of 0.1 s See pp. 75, 76

L1 runs Data taking run, in which all L1 timeslices are written to disc. See pp. 76

L1 timeslice Data package containing all L1s in a time period of 0.1 s. See pp. 76, 111, 119

L1dt calibration inter-DOM time calibration technique based on the distributions of time differences of L1 on DOMs (see section 8.1). See pp. 132–134

Monte Carlo integration Integration technique, giving an estimate of the total integral by sampling the function at random points. See pp. 225

PMT base Electronics board connected to a KM3NeT PMT to provide power to the PMT. It digitised the PMT pulse by means of a discriminator. See pp. 73

PMT time offset Time to add to the detected hit times of a PMT to obtain the calibrated hit times. It accounts for all time-delays introduced by the PMT transit time, electric cables and optical fibres between the PMT and base module. See pp. 76, 77, 109, 111, 112, 114, 115

Active galactic nucleus (AGN) Centre of an active galaxy, containing a supermassive black hole. The core region of an active galaxy outshines the outer regions by orders of magnitude. See pp. 37

All-data-to-shore Term used for the data collection strategy used in KM3NeT. All level 0 hits are sent to shore, no data reduction is applied off-shore. See pp. 75

Atmospheric muon Muon produced in the Earths' atmosphere in cosmic ray interaction particle showers. See pp. 73, 74, 77, 79, 81, 82, 84, 121, 129, 132, 135, 136, 138, 139

Atmospheric neutrino Neutrinos produced in the decays of mesons in cosmic ray air showers See pp. 45

Base module Container attached to the anchor of each detection unit. It collects the data from the DOMs of the detection unit and sends it to shore using the seafloor network. See pp. 75, 77

Bioluminescence Light emitted by organism. See pp. 73

Bjorken-x Lorentz-invariant scaling variable to describe inelastic scattering of two particles. It is related to the deflection of the incoming particle. See pp. 221, 222

Bjorken-y (y_b) Lorentz-invariant scaling variable to describe inelastic scattering of two particles. In the rest frame of the target particle, y_b corresponds to the fraction of energy of the incoming particle transferred to the target particle. See pp. 163, 164, 221, 223

Blazar AGN with one of it jets pointed towards the Earth. See pp. 37

Bremsstrahlung Energy loss of a charged particle (by emission of a photon), caused by a change in momentum due to an interaction with another charged particle. See pp. 62–64, 85, 232

Building block Set of 115 detection units. The horizontal spacing between the detection is ≈ 90 m or ≈ 20 m for KM3NeT/ARCA and KM3NeT/ORCA respectively. See pp. 68, 69, 109

Can Volume in which light emission is simulated. See pp. 82, 83

Central logic board (CLB) Electronic circuit in each DOM, used to operate the DOM, collect all data and communicate with the shore. See pp. 73, 75, 76

Charge-parity (CP) Symmetry of the combined effect of charge conjugation (particle-antiparticle exchange) and parity (mirroring of the spatial coordinates). See pp. 58

Charged current (CC) Interaction of the weak force with the exchange of a W-boson. See pp. 58, 67

Charmed meson A meson (quark and antiquark) consisting of at least one charm quark. See pp. 45

Collection efficiency Probability that a photoelectron emitted by

	the cathode of a PMT reached its first dynode. See pp. 70
Composite hypothesis	A hypothesis (used in hypothesis testing) consisting of a set of (often comparable) hypotheses. See pp. 147
Conventional atmospheric neutrino	Neutrinos produced in the decays of charged pions and kaons in cosmic ray air showers. The conventional flux is the main component of the total atmospheric neutrino flux. See pp. 45
Cosmic microwave background (CMB)	Light from the early Universe, which decoupled from matter about 379,000 years after the big bang. The wavelength of these photons is in the microwave regime. See pp. 29, 32
Cosmic neutrino background (CνB)	Neutrinos produced in interactions between neutrons, protons and electrons in the early Universe (about one second after the big bang). See pp. 43
Cosmic ray	Charged particles produced in the Universe (mainly protons). See pp. 30, 45
Cosmic ray air shower	Particle cascade caused by cosmic rays interacting in the upper atmosphere See pp. 30, 45
Cosmogenic neutrino	Flux of neutrinos caused by the interactions of high-energy cosmic rays with the cosmic microwave background. See pp. 46
Dark room calibration (DR calibration)	On shore inter-DOM time calibration using a calibrated laser signal illuminating a reference PMT on each DOM. See pp. 77, 128
Dataqueue	Software responsible for collecting and arranging the data of all DOMs See pp. 75
Deep inelastic scattering (DIS)	Deflection of a particle due to an interaction with the quarks in a hadron. See pp. 58, 221
Delayed pulse	PMT signal obtained if a secondary electron scatters of the first dynode, subsequently accelerated back to the first dynode. The resulting PMT signal is delayed. See pp. 70, 114
Delta rays	Same as ionisation, with the difference that the freed electron obtains sufficient energy to cause secondary ionisation See pp. 63, 64, 85
Detection unit (DU)	Vertical string, kept upright by a buoy, supporting 18 DOMs. The vertical spacing between the DOMs is $\approx$ 36 m or $\approx$ 9 m for KM3NeT/ARCA and KM3NeT/ORCA respectively. A dead weight holds the detection unit tight to the seabed. See pp. 68, 69, 75, 77–79, 111,

119, 133, 195, 217

Digital optical module (DOM) Pressure resistant, 43 cm diameter, glass sphere housing 31 PMTs, electronics and calibration devices. See pp. 67, 68, 72, 76, 83, 172

Discovery potential Flux needed to have a 50% probability to claim a discovery (typically at a 5σ confidence level). See pp. 52

Distance of closest approach The shortest distance between the trajectory of a (muon) track and a detection unit, both of which are described by an infinitely long straight line. See pp. 100

Effective area ($A_{eff.}$) Area of a detector, corrected for its detection efficiency, interaction probability of the interacting particles and (optionally) additional event selection criteria. When multiplied with the flux, it gives the rate of detected interactions. See pp. 70, 77

Effective track length Quantity used to describe the number of emitted photons by a shower by the length of a muon track emitting the same number of photons. An electromagnetic shower of 1 GeV has an effective track length of 4 m. See pp. 62

Electromagnetic shower (EM shower) Particle shower induced by a photon, electron or positron. All subsequent interactions are electromagnetic. See pp. 62–64, 66, 85

Estimator A rule to estimate a certain (set of) parameter(s) based on observed data. Often, an estimator based on the maximum likelihood principle is used, as it is proven to be unbiased and the most efficient estimator. See pp. 89

Event signature probability The probability $\frac{\partial^2 P}{\partial E_S^{\text{vis.}} \cdot \partial E_\mu}$ that an interacting neutrino of certain energy gives an event signature with visible shower energy $E_S^{\text{vis.}}$ and muon energy E_μ. This probability depends on the neutrino flavour and interaction channel. See pp. 163

Event signature probability table Calculated (cumulative) event signature probabilities at a large set of $E_S^{\text{vis.}}$, E_μ and E_ν for a certain interaction channel, stored in a tabulated form. See pp. 166

Event weight Quantity determined for simulations used in Monte Carlo integration techniques. It facilitates a straightforward evaluation of the integral for a flux other than the one used in the generation. See pp. 85, 225

False negative Error occurring in hypothesis testing; the hypothesis that includes a signal is rejected, while in reality it was the true hypothesis. See pp. 145

False positive Error occurring in hypothesis testing; the hypothesis that includes a signal is accepted, while in reality, it was (a statistical fluctuation of) the background. See pp. 145

Field-programmable gate array (FPGA) Integrated circuit that is fully programmable. In a KM3NeT DOM, it is used for digitisation of the data. See pp. 73, 76

Flavour identification Determination of the most likely flavour of interacting neutrinos. In general, it gives the probability that an event is caused by the (atmospheric muon) background. See pp. 101

Floor Group of DOMs at the same depth. Floor one corresponds to the lowest DOM, floor eighteen to the top DOMs. See pp. 68

Gain The (average) number of electrons hitting the anode of a PMT as the result of a single emitted photoelectron. See pp. 70

Generation volume Volume used in simulations. It defines the volume in which the primary interactions of interest are simulated See pp. 82, 83

Hadronic shower Particle shower caused by a hadron. The shower develops through hadronic as well as electromagnetic interactions. See pp. 63–66

Hadronuclear Production of neutrinos by interactions of protons with matter. See pp. 33, 34

Height of closest approach The height along a detection unit corresponding to the distance of closest approach. See pp. 100

Height of first light Term used in the reconstruction of muons, corresponding to the height along the detection unit where light from the fitted muon could first been observed. See pp. 100, 136

Hidden core Term used in neutrino astronomy to describe an cosmic ray acceleration site from which light cannot escape, such as the core region of AGN. See pp. 47, 48

High rate veto Veto applied to hit rate on a PMT exceeds 20 kHz. This veto is applied by the CLB in a DOM. See pp. 73

High voltage tuning Term used to describe the in-situ tuning of the PMT high voltages to set the PMT gains to

$3 \cdot 10^6$, or 1 p.e.. See pp. 75, 77

High-energy starting event (HESE) High-energy track- and shower-like events with the interaction position within the instrumented volume. See pp. 55

Hit time residual Time difference between the measured time of a L0 and the expectation of a reconstructed event signature. The hit time residual is positive in case the measured hit time is larger than the expectation. See pp. 131–133

Hypothesis test Statistical method used to make a judgement, based on the available data, which hypothesis out of two hypotheses represents the (unknown) true hypothesis. See pp. 145, 146

Importance sampling See pp. 225, 226

In-situ From Latin; in place. In this thesis, this term is generally used to describe analyses done using data taken offshore rather than in a laboratory environment. See pp. 67, 75, 77, 109, 111, 114, 132, 133

Inter-DOM time calibration Calibration of DOM time offsets. In combination with the inter-PMT calibration it defines the time calibration of the detector. See pp. 76, 77

Inter-DU time offset Time shift accounting for different cable lengths from each DU base to the shore station See pp. 77, 134

Inter-PMT calibration In-situ calibration of the relative time offsets, transit time spreads and relative efficiencies of the 31 PMTs in a DOM. This is done exploiting the background of ^{40}K decays. See pp. 77

Ionisation A charged particle knocks an electron out of an atom. See pp. 63, 66, 85

Late hit A hit from a photoelectron with an anomalous long drift time, for instance when it is emitted from a region of the photocathode where the electric field is relatively weak. See pp. 70, 114

Level one hit (L1) Two or more L0s recorded on a single DOM within a certain time window. For KM3NeT/ARCA, a time window of 25 ns is used. See pp. 76, 83, 99, 119–121, 127–129, 131–133, 179

Level zero hit (L0) Elementary quantity describing a recorded PMT pulse. It consist of the time at which a PMT pulse passes a certain threshold and the time it exceeds this threshold. See pp. 74–76, 80, 81, 86, 87, 89, 92, 93, 95, 97–99, 101, 109–114,

116, 117, 119, 120, 127, 131, 177–181, 214, 215, 243, 245

Likelihood ratio test statistic Ratio between the likelihood to obtain a certain data set under the assumption of $\mathcal{H}_1$ being true, and the likelihood under $\mathcal{H}_0$. See pp. 146

Maximum likelihood estimate Estimate of (a) parameter(s) given by the coordinates of the maximum of the likelihood to observe the data given the parameters to fit. See pp. 89, 90

Maximum likelihood ratio Likelihood ratio test statistic, often used for composite hypotheses, defined as the maximum likelihood ratio for all possible parameters that describe the hypotheses. See pp. 147

Mean free path Average length a particle travels through a medium before it undergoes an interaction. See pp. 61

Mis-reconstructed track Tracks with fitted parameters significantly different from the true parameters. In the case of the single line reconstruction (section 9.1), these are defined as tracks with a reconstructed zenith angle differing more than 20 ° from the true zenith angle. See pp. 135

Multi-messenger Astronomy exploiting the correlation(s) between photons, cosmic rays, neutrinos and/or gravitational waves. See pp. 50

Multiplicity Number of PMTs in a DOM registering a coincidence of L1s within a time window of 25 ns. See section 7. See pp. 120, 121, 123

Muon bundles Bundle of atmospheric muons, originating from the same cosmic ray air shower. See pp. 84

Nanobeacon calibration (NB calibration) Inter-DOM time calibration based on the arrival time of the light emitted by *nanobeacons* on surrounding DOMs. See pp. 77, 131, 132

Neutral current (NC) Interaction of the weak force with the exchange of Z-boson. See pp. 58

Neutron star Massive, very dense star consisting mainly of neutrons. Neutron stars are produced in type II supernovae. See pp. 35

Nominal PMT Model of a PMT, resembling a set of PMTs characterised in a laboratory environment. This model is used in simulations. See pp. 70, 72, 87, 110, 113, 122

Opening angle The angle between the axes of two PMTs. An angle of 0 ° corresponds to two parallel PMT looking in the same direction. See pp. 110

Outlier Data point which deviates from its expectation value by several standard deviations. These data particularly affect χ^2-fits. See pp. 91, 92, 136

P00075 Term for the model used to describe lightscattering in seawater at KM3NeT. See pp. 67, 68

Pair production Decay of a (high-energy) photon to an electron-positron pair. See pp. 62–64, 85

Parton distribution function Set of functions giving the probability that a neutrino interacts with quarks and gluons in a nucleon, as function of the Lorentz-invariant Bjorken-y and Bjorken-x variables. In the rest frame of the nucleon, these parameters relate to the transferred energy and scattering angle of the neutrino. See pp. 58

Photodisintegration Process in which a photon knocks one or more neutrons, protons or alpha-particles out of a nucleus See pp. 32

Photoelectron The name used for an electron emitted by the photoelectric effect. See pp. 70, 75

Photoelectron equivalent (p.e.) Quant ity to describe the deposited charge on a PMT anode. A nominal PMT has a gain of 1 p.e.. See pp. 75

Photohadronic Production of neutrinos by interactions of protons with photons. See pp. 33, 34, 46

Photomultiplier tube (PMT) Instrument capable of detecting single photons, using a chain of dynodes to amplify the charge induced by a photon hitting the photocathode. It is the principal detection element of the KM3NeT detectors. See pp. 67–73, 75–77, 172, 178

Photonuclear interactions Inelastic scattering of a lepton with a nucleus. See pp. 63, 64, 85

Power The probability not to make a type II error. A hypothesis test is often designed to maximise the power at a given significance level. See pp. 145, 146

Preliminary reference earth model (PREM) Model of the density profile of the Earth, obtained from seismic data. See pp. 159, 235

Prepulse A hit with a short transit time and small amplitude, caused by a photon freeing a photoelectron from one of the first dynodes rather than the photocathode. See pp. 70

Probability density function (PDF) Function describing the likelihood to observe

a continuous parameter x at a certain value. See pp. 70, 163

Prompt atmospheric neutrino Atmospheric neutrinos produced in the decays of heavy, charmed mesons in cosmic ray air showers See pp. 45

Pulsar Fast spinning neutron star, of which the magnetic fields propels two jets of electromagnetic radiation. See pp. 36

Pulsar wind nebula Supernova remnant with a pulsar in its core. See pp. 36

Quantum chromodynamic (QCD) Theory describing the strong interaction between quarks and gluons. See pp. 58

Quantum efficiency (QE) Ratio between number of emitted photoelectrons (due to the photoelectric effect) and number of incident photons. See pp. 70, 81, 224

Quasar Extremely luminous active galactic nucleus. See pp. 37, 51

Recombination First moment after the big bang when electrons and protons bound together to form hydrogen atoms, making the Universe transparent for photons. See pp. 30

Reductio ad impossible Proof by contradicting the opposite. See pp. 233

Seed Start value for a (maximum likelihood) fit. See pp. 90, 94, 97

Sensitivity Flux need to have a probability of 50% to observe a signal with a certain confidence level. Typically a 90 % confidence is used. See pp. 52, 145

Shower energy Energy of a primary particle inducing a particle shower. See pp. 62–64

Significance level The probability to make a type I error. Often, the significance level is set to a predefined level, and the acceptance region is chosen accordingly. See pp. 145, 233

Singles rate The L0s detected by a single PMT. See pp. 113

Solar atmospheric neutrinos Neutrinos produced in cosmic ray interactions in the Sun's atmosphere. See pp. 45

Spectral index Factor γ in a $dN/dE \propto E^{-\gamma}$ energy spectrum. See pp. 54

Starburst galaxy Galaxy with a extraordinary rate of star formation and supernovae. See pp. 37

Sterile neutrino Hypothetical fourth type of neutrino, not interacting through any of the standard model

	interactions (thus only gravity). See pp. 33
Supernova (SN)	Explosion caused by the collapse of a massive star. See pp. 34–37, 50
Supernova remnant (SNR)	Remnants of a supernova, consisting of an expanding shockwave and the remaining object in its core. See pp. 35, 38
Test statistic	Value assigned to each possible outcome of an experiment (the data). See pp. 146
Time over threshold (ToT)	Time period a PMT pulse exceeds a predefined threshold (see the definition of a L0). See pp. 73–75, 87, 88, 116, 214, 243, 245
Transit time	Mean time difference between a photon hitting the PMT cathode and the arrival of the electrons on the PMT anode. See pp. 70, 81
Transit time spread (TTS)	Spread in the transit time distribution of a PMT, defined as the standard deviation of the transit time distribution. See pp. 70, 71, 109
Trigger	A pattern of L0s compatible with a shower- or track-like event signature. See pp. 76, 138, 142
Triggered event	Snapshot of all recorded L0s encompassing a trigger. The time window of a triggered event is about 10 μs for KM3NeT/ARCA. See pp. 76, 99, 179
Type I error	See the definition of a false positive See pp. 145
Type II error	See definition of a false negative. See pp. 145, 233
Type II supernova	Supernova caused by the collapse of a star with more than ten solar masses See pp. 35
Type Ia supernova	Supernova caused by the collapse of white dwarf with a mass exceeding 1.4 solar masses See pp. 35
Unobstructed PMTs	A PMT unobstructed by the DOM support structure shadowing it. The unobstructed PMTs in a DOM are labelled C1,3,4&6, D1&4, E1,3,4&6 and F1&4 (see figure 3.6). See pp. 112
Vertex	A point in time and space, given by a position and a time coordinate See pp. 92
Vertical electro-optical cable (VEOC)	Pressure resistant cable housing the electrical wires and optical fibres used to power and communicate with the DOMs in a detection unit. See pp. 75
Visible shower energy	Fraction of the total electromagnetic and hadronic shower energy that is not carried away by par-

ticles not emitting light See pp. 92, 152

White dwarf Dense cosmic object, produced in the collapse of star weighing a few times the mass of the Sun. It consists of a plasma of neutrons and electrons See pp. 34

Zenith angle Angle between the (reconstructed) direction of an event signature and the zenith. An upward moving particle has a zenith angle of 0 ° See pp. 79, 80, 84, 135, 136, 138–140

Dankwoord / Acknowledgements

"Met verstandige dingen kun je halverwege ophouden, absurditeiten echter dienen voltooid te worden, omdat zij geen andere rechtvaardiging dan hun consequent doorzetten bezitten" [228]. Mensen in mijn omgeving vonden het soms absurd dat ik geruime tijd onbetaald gewerkt heb aan een nieuwe methode om het universum nét iets beter te kunnen bestuderen. Terugkijkend kan ik hen niet op alle punten ongelijk geven. Desalniettemin kan ik mijn promotietraject zeker niet onverstandig noemen. Integendeel, het is een enorm leerzame, prettige en prachtige ervaring voor mij geweest. Met blijdschap kijk ik terug op de talrijke fijne momenten die ik beleefd heb. Dit was niet mogelijk geweest zonder de hulp en het gezelschap van vele mensen. Met het risico dat niet iedereen de aandacht toekomt die hij/zij verdient wil ik bij dezen een aantal mensen in het bijzonder bedanken.

Dit proefschrift had hier niet gelegen zonder de gewaardeerde hulp van mijn promotor en copromotor. Paul, hartelijk dank voor de aangename, met anekdotes dooraderde, wijze waarop je mijn promotietraject begeleid hebt. Zonder je heldere suggesties was dit proefschrift een stuk lastiger te lezen. Ronald, ondanks je enorm drukke agenda stond de deur van je kantoor altijd open. Als ik langskwam om over een probleem te praten wist jij mij altijd te wijzen op mogelijke problemen die ik zelf nog niet voorzien had. Bovenal gunden jullie mij veel vrijheid in het ontwikkelen van mijn analyses. Ik ben hier zeer dankbaar voor.

I would like to thank the members of my reading-committee for their patience, advice and for the compliments I received. Those mean a lot to me.

Almost all paragraphs in this thesis have suffered from many iterations while trying to give a full overview of the research project as concisely as possible. Clearly, I did not succeed. It was only after other people provided feedback that I realised which parts are important (and which not) and how to structure the text. In this regard I'm particularly thankful to Paul, Ronald, Brían and Thijs, as well as Martijn, Lodewijk, Rodrigo, Jordan, Rasa, Bouke, Alfonso, Milo, Paul, Aart, Suzan, Johan and Adri.

The passionate joy I found in doing fundamental research would not have been such great if I wouldn't have felt so welcome in the Nikhef communittee. I will always remember the pleasant conversations and discussions on a wide variety of subjects. In particular I would like to thank my office-mates Martijn, Lodewijk and Brían for motivating me, sharing your and solving my problems and the nice time we had together. Maarten, Dorothea, Aart and Paul, you provided indispensable input to my analyses and taught me how to structure a research project. I envy the way you only needed a single word to realise what the problem is and the promptness with which the solutions were found. Marc, Marko and Broos, thanks for the insights into the politics of the scientific community and the mental support in the last phase of my PhD candidacy. Support of various kinds

was also given to me by Laurent, Remco, Jacco, Els, Mieke, Maria, Luuk, Maarten, Mick and Erik. Many thanks for making the past years such a pleasant time.

Fysieke en mentale afleiding is onmisbaar gebleken. Sjoerd en Sander, bedankt voor de spiegel die jullie me van tijd tot tijd voorhielden. Sander, Maarten, Nico, Arno en Thijs, de 'strafkampen' waren verre van een straf. Sjoerd, Elias, Bart, Mirjam en Frans, het is een voorrecht om met jullie de dag op de Amstel te beginnen.

Mijn familie, en met name mijn ouders, hebben mij van jongs af aan nieuwsgierigheid, een gezonde dosis doorzettingsvermogen, eigenwijsheid en een kritische blik aangeleerd. Mijn schoonfamilie heeft mij geleerd om niet altijd meteen te zeggen wat ik denk, en dat wat ik denk in wat tactischer woorden te verpakken. Ik wil jullie allen enorm bedanken voor deze eigenschappen en voor de onvoorwaardelijke steun. Ik hoop dat ik nog veel van jullie mag leren.

Het promotietraject, en zeker het schrijven van dit proefschrift, heb ik soms als een eenzaam traject ervaren. Het verminderde sociale contact tijdens de pandemie van afgelopen anderhalf jaar heeft hier niet bij geholpen. Desalniettemin heb ik deze tijd niet als onprettig ervaren. Dit komt voornamelijk door de extra tijd die ik door heb kunnen brengen met de twee meest belangrijke personen in mijn leven. Milou, reeds in de buik van je moeder heb je me in laten inzien dat ik moet werken om te leven, en niet moet leven om te werken. Met onbeschrijfbaar plezier zie ik hoe jij je elke dag ontwikkeld. Graag zou ik de verwondering waarmee jij de wereld om je heen bekijkt willen ervaren. Suzan, niemand weet zo goed als jij hoeveel tijd ik aan mijn PhD besteed heb, de dingen ik afgezegd heb, hoe chagrijnig en teneergeslagen ik kon zijn als iets niet werkte zoals ik het wilde, maar ook de grote vreugde als ik een oplossing gevonden had of een nieuw idee had. Jij hebt me altijd gesteund, en opgeven was ook voor jou nooit een optie. Suzan, jij hebt mij echt gevormd tot wie ik nu ben.

Karel Melis, september 2021

9 789464 215069