Bergische Universität Wuppertal
Fakultät 4 - Mathematik und Naturwissenschaften

Search for Low Relativistic Magnetic Monopoles at the IceCube Neutrino Observatory

Inauguraldissertation zur
Erlangung der Doktorwürden (Dr. rer. nat.)

Der Fachgruppe Physik vorgelegt von

M. Sc. Frederik Hermann Lauber

im Juli 2021 in Wuppertal

BERGISCHE
UNIVERSITÄT
WUPPERTAL

The PhD thesis can be quoted as follows:

urn:nbn:de:hbz:468-urn:nbn:de:hbz:468-20211202-100837-4
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20211202-100837-4]

DOI: 10.25926/gph1-vf51
[https://doi.org/10.25926/gph1-vf51]

Gutachter:

Prof. Dr. Klaus Helbing

Bergische Universität Wuppertal

Prof. Dr. Gisela Anton

Friedrich-Alexander-Universität Erlangen-Nürnberg

Abstract

Magnetic monopoles, carriers of magnetic charge, are particles theorized to exist for nearly a century. No detection of such particles has been substantiated as of yet. One explanation for this could be a high rest mass of such particles. In this case, creation of magnetic monopoles could only have happened in early epochs of the Universe. Thus, their density would be diluted by the consequent expansion of the Universe. As magnetic monopoles carry magnetic charge, they undergo acceleration by magnetic fields. Due to the cosmic magnetic fields, their expected velocity at the position of the Earth ranges from 10^{-3} c to close to c. The IceCube Neutrino Observatory (**IceCube**) utilizes a cubic kilometer of natural Antarctic ice as a detection medium for neutrinos. Designed to measure Cherenkov light emitted by secondary particles of neutrino-ice interactions, it is also suited to detect magnetic monopoles. Previous searches at **IceCube** focused on Cherenkov light either produced directly or indirectly by magnetic monopoles or theorized light production channels like catalysis of nucleon-decay. This left the low relativistic regime between 0.1 c to 0.55 c untested. Luminescence light is induced by the conversion of kinetic energy of a magnetic monopole passing through a target material to atomic or molecular excitations of said target material. This thesis establishes luminescence light as a new light production channel to detect magnetic monopoles at **IceCube**. An analysis is presented probing the low relativistic regime at **IceCube** for the first time. 2524.6 days of **IceCube** data taken between the seasons of 2011/2012 and 2017/2018 are investigated for magnetic monopole signatures. No excess of monopole candidate was detected. The possible flux of magnetic monopoles in the low relativistic regime is restrained down to 9.6×10^{-19} cm^{-2} s^{-1} sr^{-2}, a two order of magnitude improvement over previous best flux limits.

Zusammenfassung

Magnetische Monopole sind Senken und Quellen des magnetischen Feldes. Klassisch können magnetische Monopole als magnetisches Pendant zu den elektrisch geladenen Leptonen und somit als Teilchen angesehen werden.

Mathematische Lösungen, die funktional identisch mit Senken und Quellen des magnetischen Feldes und somit mit magnetischen Monopolen sind, können in vielen Theorien, die das Standardmodell der Teilchen erweitern, gefunden werden. Die Existenz von magnetischen Monopolen konnte bisher nicht nachgewiesen werden.

Eine Theorie dies zu erklären besteht in einer prohibitiv hohen Ruhemasse eines magnetischen Monopols. Dadurch wäre die Erzeugung durch heutige, sowohl anthropogene als auch natürliche, Teilchenbeschleuniger unmöglich. Nur magnetische Monopole, die als Relikte aus der Entstehung des Universums zurückgeblieben sind, könnten detektiert werden. Da das Universum im Laufe seiner Entwicklung expandierte, wäre die heutige Dichte von magnetischen Monopolen im Universum sehr gering.

In diesem Szenario sind magnetische Monopole mit vernachlässigbar kleiner kinetischer Energie im Bezugssystem der Erde erzeugt worden. Sie können durch kosmische magnetische Felder beschleunigt werden. Verschiedene Beschleunigungsszenarien, basierend auf der kohärenten Feldlänge und Feldstärke, existieren. Je nach Szenario, Ruhemasse und magnetischer Ladung eines magnetischen Monopols sind Geschwindigkeiten zwischen $10^{-3}\ c$ bis nahezu Lichtgeschwindigkeit an der Position der Erde zu erwarten.

Das IceCube Neutrino Observatory (**IceCube**) ist ein Wassereis basierter Neutrinodetektor. Ein Kubikkilometer des natürlichen Eises der Antarktis wurde mit Lichtsensoren ausgestattet, um die Interaktionsprodukte von Neutrinos mit dem Eis zu detektieren. Die Interaktionsprodukte induzieren dabei Tscherenkow-Strahlung. Neben Neutrinos ist **IceCube** auch fähig andere Teilchen, die das Detektionsmedium passieren, zu detektieren. Voraussetzung dafür sind genügend hohe Lichtemissionen des passierenden Teilchens.

Magnetische Monopole sind eben solche Teilchen. Sie induzieren sowohl selbst Tscherenkow-Strahlung als auch durch sekundäre Teilchen. Zusätzlich können sie theoretisch Nukleonenzerfall induzieren, was ebenfalls Lichtemissionen zur Folge hätte. Diese Lichtproduktionskanäle wurden in der Vergangenheit zur Suche nach magnetischen Monopolen in den Daten von **IceCube** genutzt. Dadurch konnte das nieder-relativistische Regime zwischen 0.1 c und 0.55 c nicht untersucht werden, da nur vernachlässigbare Lichtmengen durch die zuvor genannten Produktionskanälen in diesem Bereich erzeugt werden.

Lumineszenzlicht wird durch die Umwandlung von kinetischer Energie eines magnetischen Monopols in atomare oder molekulare Anregungen innerhalb eines passierten Materials induziert. Magnetische Monopole verlieren aufgrund ihrer hoch ionisierenden Wirkung kinetische Energie in der Ordnung von 10 MeV cm^{-1} bis 1000 MeV cm^{-1} in wasserbasierten Eis, das wiederum zu Lumineszenzlicht führt. Durch ihre hohe Ruhemasse hat dieser Energieverlust auf einem Kilometer innerhalb von Eis keinen nennenswerten Geschwindigkeitsverlust zur Folge.

In dieser Arbeit wird Lumineszenz als Lichtproduktionskanal zur Detektion von magnetischen Monopolen verwendet. Zum ersten Mal wurden die Daten von **IceCube** auf Signaturen von magnetischen Monopolen im nieder-relativistische Geschwindigkeitsbereich untersucht. Dazu wurde Lumineszenz als Lichtproduktionsmechanismus in die Simulationssoftware von **IceCube** implementiert und validiert. Die Möglichkeit der Detektion von magnetischen Monopolen mit der vorhandenen Datenerfassung von **IceCube** wird demonstriert.

Im Anschluss wird eine Analyse vorgestellt, die auf Basis dieser Detektorsimulation magnetische Monopole und simulierte Untergrundereignisse separiert. In diesem Fall sind Untergrundereignisse Standardmodell konforme Interaktionen von kosmischer Strahlung mit der Atmosphäre der Erde oder dem Eis der Antarktis. Die simulierten Untergrundereignisse werden mit einer Stichprobe der gemessenen Daten verglichen, um die Korrektheit der Untergrundsimulation zu zeigen.

Diese Stichprobe ist klein genug, um nur eine vernachlässigbar kleine Anzahl an magnetischen Monopolen zu beinhalten, wodurch eine Beeinflussung des Analysedesigns ausgeschlossen werden kann.

Mit einer Sequenz von univariaten Entscheidungen werden magnetische Monopole im nieder-relativistischem Bereich selektiert. Diese univariaten Entscheidungen basieren sowohl auf bereits in **IceCube** im Einsatz befindlichen und definierten Variablen, als auch auf speziell für diese Arbeit entworfenen. Im Anschluss wird ein auf maschinellem Lernen basierender Selektionsschritt verwendet, um die finalen magnetischen Monopol Kandidaten zu selektieren. Da in diesem Schritt die Untergrundsimulation bereits statistisch limitierend ist, wird ein Stichprobenwiederholungsverfahren eingesetzt, um Aussagen über die zu erwartenden Verteilungen von Variablen nach diesem finalen Selektionsschritt zu treffen.
2524.6 Tage von **IceCube** Daten, aufgezeichnet zwischen dem April 2011 und Juli 2018, werden untersucht. Basierend auf der Untergrundsimulation sind in diesem Zeitraum $0 \leq N_B \leq 10$ Untergrundereignisse mit einem Mittelwert $\overline{N}_B = 2$ zu erwarten. Da die Untergrundsimulation statistisch limitierend ist, wurde die Projektion der Anzahl zu erwartenden Untergrundereignisse ebenfalls mit der zuvor genannten Stichprobe der gemessenen Daten ausgeführt. Aus dieser werden $0 \leq N_B \leq 4$ Untergrundereignisse mit einem Mittelwert $\overline{N}_B = 0$ für den zuvor erwähnten Zeitraum prognostiziert.
Nach Anwendung der Analyse auf die Daten verblieben zwei Ereignisse. Dies liegt innerhalb des zu erwartenden Untergrunds. Somit konnten keine magnetischen Monopole nachgewiesen werden.
Die Ereignisse wurden visuell als auch mit statistischen Methoden untersucht. Sie scheinen untergrundartig zu sein. Als konservativer Ansatz wurden die beiden Ereignisse als magnetische Monopole angenommen und ein geschwindigkeitsabhängiger oberer Fluss von magnetischen Monopolen im nieder-relativistischen Bereich berechnet. Dieser unterbietet den bisherigen oberen Fluss um zwei Größenordnungen und limitiert den Fluss von magnetischen Monopolen auf bis zu unter $9.6 \times 10^{-19}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-2}$.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The existence of magnetic monopoles seems like one of the safest bets that one can make about physics not yet seen.

J. Polchinski [Pol04]

Magnetic monopoles (**MM**s) are hypothesized objects exhibiting the attributes of a magnetic source. First proposed in 1931 by P. A. M. Dirac, their existence would explain the quantitation of the electric charge and remedy the classical asymmetry between the electric and magnetic fields. For most of the last century, many theories have included them, prominently any Grand Unified Theory (**GUT**), unifying the strong interaction, weak interaction, and electromagnetic interaction into a single interaction. An overview of the field of **MM** search is outlined in Chapter 3 including different theories, kinematic attributes at Earth, interactions with matter, and current bound on the upper limit of **MM** flux by multiple experiments.

As yet, no experimental confirmation of the existence of **MM** has occurred. Bounds on the upper limit of **MM** flux at the location of the Earth have been established both by cosmology and experiments. Even assuming the weakest cosmological bound on the upper limit of the **MM** flux, already superseded by experiment by orders of magnitude, the *Parker bound* (see Section 3.9), the flux of **MM** at the Earth is only in the order of $4\,\mathbf{MM}\text{s}\,\text{m}^{-2}\,\text{kyr}^{-1}$. Thus, any search for **MM** requires detectors with huge effective **MM** detection areas to be able to detect **MM** in a reasonable amount of time.

The IceCube Neutrino Observatory (**IceCube**) features such detectors with its in-ice arrays instrumenting a km^3 of Antarctic ice with Digital Optical Modules (**DOM**s). In this work, ice always means water-based ice. Designed to detect direct Cherenkov light from secondaries of neutrino-ice interactions, its large effective area and instrumentation density allow for the search of **MM**. A description of **IceCube** is outlined in Chapter 2.

Depending on the velocity and theory of **MM**, different light production mechanisms exist which facilitate light production within **IceCube**'s detection volume. In this thesis, the velocity range of 0.1 c to 0.55 c, called the low relativistic regime, is of interest. In contrast to the mildly and relativistic regime above this velocity, no direct and barely any indirect Cherenkov light is produced. Instead, luminescence light becomes the dominant light production mechanism.
Luminescence light is isotropically, delayed emitted light due to excitation of the molecular or atomic structures in a medium by a passing particle. Both direct and indirect Cherenkov light and luminescence light are described in Chapter 5.
Searches in the mildly relativistic and relativistic regime have already been conducted at **IceCube** setting competitive limits on the flux of **MM**s. This work sets out to extend the **MM** sensitivity of **IceCube** down to the low relativistic regime. A blinded analysis, meaning it is only based on simulated responses of the detector, is conducted to detect **MM**s. Luminescence light as a new light production mechanism is implemented in the simulation framework of **IceCube**. Specialized signal simulation for **MM**s in the low relativistic regime is conducted and described in Chapter 6.
The analysis, described in Chapter 7, consists of multiple single variable based selection steps up to a machine learning (**ML**) based selection step for the final selection sample. Remedies to low background statistics are applied in the form of bootstrap aggregating (**bagging**).
The final event sample is discussed in detail in Chapter 8. A new bound on the upper limit of **MM** flux at the location of the Earth in the low relativistic regime is calculated, superseding previous limits by two orders of magnitude.

Chapter 2

IceCube Neutrino Observatory

Fortune would be in a hard mood indeed if it allowed such a combination of knowledge, experience, ability, and enthusiasm to achieve nothing.

R. F. Scott [Sco13]

The IceCube Neutrino Observatory (**IceCube**), described in references [Ach+06; Aar+17], is situated close to the geographic South Pole. It was primarily designed to detect statistically significant numbers of high energetic neutrinos. Due to the low cross-section of neutrino-matter interactions, large volumes of material are needed to be instrumented to achieve reasonably high neutrino count rates. Additionally, neutrino based events become rarer with increasing energy. On the other hand, higher instrumentation density enables detection of lower energetic events which produce less light.

As the cost of a detector increases with raised total instrumentation, some form of compromise between the density and volume has to be reached. **IceCube** accomplishes this by combining multiple detectors optimized for different energy ranges. The multi-detector architecture of **IceCube** is described in Section 2.1, including the utilized detection medium. Subsequently, the data acquisition chain to record data is outlined in Section 2.2.

2.1 Detector Architecture of IceCube

IceCube applies a multi-detector design utilizing natural transparent materials as a detection medium for neutrinos and other particles to balance light sensitivity and cost. Two of its detectors utilize the natural Antarctic ice: the **IceCube** in-ice array (**IC**) and the **IceCube** DeepCore array (**DC**). Both consist of multiple Digital Optical Modules (**DOM**s) embedded in the Antarctic ice on vertically drilled strings. In total, 5160 **DOM**s are installed. The **IC** features a higher volume but lower instrumentation density than the **DC**. A schematic of **IceCube** can be seen in Figure 2.1. While other detectors exist at **IceCube**, they are not utilized in this thesis and hence are omitted from description.

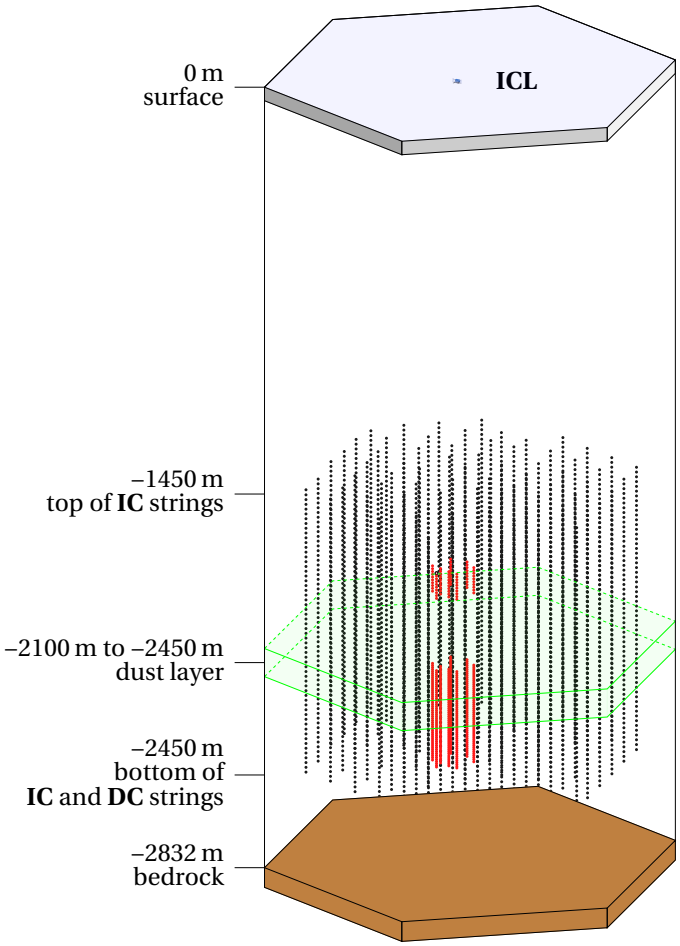


Figure 2.1: **Schematic of IceCube**
*A schematic of the positioning of all **DOM**s on their respecting strings in both the **IC** and the **DC** is shown. The strings are each connected to the **ICL** on top of the ice. **DC** exclusive **DOM**s are marked in red. The dust layer, a region in the ice with high light absorption due to impurities in the ice, is marked in green. The schematic is to scale besides the radii of the **DOM**s which have been upscaled by a factor of 50 for better visibility.*

All detectors are connected to the **ICL**. The **ICL** contains all the support infrastructure needed to operate **IceCube**, including all servers, data storage, power distribution infrastructure, and infrastructure for external satellite communication. Data is collected from all sub-detectors and processed centrally while preprocessing might be applied in each sub-detector.

Depending on different criteria, the data may be deleted or stored locally. If additional criteria are met, the data is also transmitted via satellite to the **IceCube** data center outside of Antarctica for immediate analysis. The locally stored data will be shipped on hard disks to the data center once per year for further analysis. While **IceCube** was designed to detect primarily neutrinos, the size of fiducial volume facilitates the search for presumably rarer, exotic events.

2.1.1 Antarctic Ice

Antarctica is mostly covered by ice. This ice is the accumulation of snow in Antarctica which was compressed down into firn and then ice under its own weight over time. The oldest ice at the bottom, directly on the bedrock, is about 165 kyr old as described in reference [PWC00]. Due to variations in snowfall rate and impurities in the snow, the optical properties of the ice can change depending on age and hence depth.

These optical properties are calibrated for by the **IceCube** collaboration (**ICC**), described in reference [Ron19], by using specialized impurity measurement devices lowered into the ice or by in-situ measurements utilizing light sources within the **DOM**s. The most prominent impurity relevant to this thesis is called the *dust layer*. At a depth between –2100 m to –2450 m, a region with high content of dust exists. This dust reduces the absorption length of light drastically and can be observed in many event renders of recorded events by **IceCube**.

2.1.2 Digital Optical Modules

The Digital Optical Module (**DOM**) is the common data collection module of the **IC** and **DC**. Each **DOM** consists of a glass pressure sphere in which a 25.4 cm-diameter photomultiplier tube (**PMT**) is placed at the bottom half, facing outward. As a support structure, a collar is molted onto the neck of the **PMT**. Around the collar of the **PMT**, multiple interconnected electronic boards are placed: the High Voltage Control Board, the **PMT** base board, the Flasher Board, and the Delay Board. A schematic of the inside of a **DOM** is depicted in Figure 2.2.

The High Voltage Control Board generates and monitors the high voltage required to operate the **PMT**. The high voltage is supplied to the **PMT** base board which distributes it to the individual cathodes of the **PMT**. Furthermore, it contains filters and adjusts the coupling of the output signal which is recorded by the Main Board.

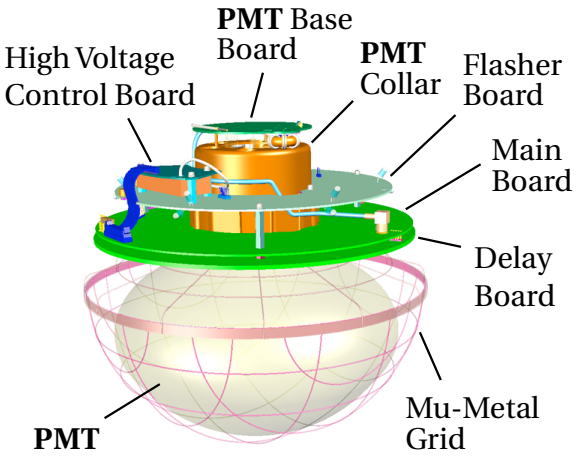


Figure 2.2: **Schematic of a DOM**
*A schematic of a **DOM** without the pressure sphere is depicted. The downward facing the **PMT** sits at the bottom. Magnetic shielding is implemented via a Mu-Metal Grid surrounding the **PMT** from the bottom to the neck, shown here in pink. Depicted in orange, the **PMT** collar is molted onto the neck. It serves as a support structure for all electronic boards. In light green, the Delay Board is illustrated while the **PMT** Base Board is painted in dark green. Graphic taken and modified from reference [Aar+17].*

The Main Board is the central management and data recording system of the **DOM**. All communication inside or outside the **DOM** is implemented in the Main Board. Once the output signal of the **PMT** base board reaches a configurable threshold, a trigger signal is raised. The nearest and second-nearest neighboring **DOM**s on the same string are notified of the trigger signal via specialized cabling. Locally to the **DOM**, the trigger signal causes the data acquisition to be started.

Data acquisition is performed by two complementary sampling systems: the Fast Analog to Digital Converter (**fADC**) and the Analog Transient Waveform Digitizer (**ATWD**). The **fADC** has a sampling speed of 40 Msps. Data is collected for 6.4 µs after the trigger signal was raised. The **ATWD** has a sampling speed of 300 Msps. Data is recorded from 75 ns before to 352 ns after the trigger signal was raised. Data before the trigger signal can be recorded as the signal is passed through the Delay Board. The Delay Board is a passive board which consists of an approximately 10 m of printed electrical circuit delaying the output signal before it arrives at the **ATWD**. After it has been delayed, it is passed through three parallel amplifiers with different gain factors. The resulting signals are recorded by three channels of the **ATWD** increasing the dynamic range of the amplitude of the signal. The combination of all recorded data is called a *hit*.

A hit which is coincident within ± 1 µs with an external trigger signal from a nearest or second-nearest neighboring **DOM**s is called a hard local coincidence (**HLC**) hit. Any other hit is called a soft local coincidence (**SLC**) hit. For both, the first three samples of the **fADC** are stored along with the time of the local trigger signal. In case of a **HLC** hit, the full waveforms are also stored albeit not every digitized channel of the **ATWD**. Only channels required to measure the dynamic range of the recorded signal are kept.

The last board is the Flasher Board. While the Main Board contains a low power light source for calibration purposes of the **DOM** local **PMT**, the Flasher Board contains a high intensity light source that can be used to calibrate other **DOM**s or perform in-situ measurements of the surrounding ice.

2.1.3 IceCube In-Ice Array

The **IceCube** in-ice array (**IC**) is an in-ice detector of **IceCube**. It consists of 78 vertical strings which are embedded in the ice of Antarctica in a hexagonal formation. The inter-string spacing is 125 m. On each string, there are 60 **DOM**s. Placement of the **DOM**s starts at a depth of –1450 m and ends at a depth of –2450 m with a 17 m vertical inter-**DOM** spacing. In total, this instruments a 1 km^3 of ice.

2.1.4 IceCube DeepCore Array

The **IceCube** DeepCore array (**DC**) is an in-ice detector of **IceCube**. It is situated within the detection volume of the **IC** creating a higher instrumented region within the ice. It consists of 7 strings shared with the **IC** and 8 extra strings dedicated to the **DC**. These dedicated strings feature a closer vertical **DOM** spacing of either 7 m or 10 m and lower the average horizontal inter-string spacing to 72 m with a minimum distance of 41 m and a maximal distance of 105 m. The **DOM**s on these strings also feature **PMT**s with a higher quantum efficiency. While this detector does not feature the detection volume of the **IC**, it is much more sensitive to dimmer events than the **IC**.

2.2 Data Acquisition of IceCube

IceCube is operated with a triggered data acquisition system which selects data acquisition events based on **HLC** hits. A data acquisition event is the combination of all **HLC** and **SLC** hits within a defined time frame. Once a data acquisition event has been detected, it is always recorded locally at the **ICL**, but might not be transmitted via satellite based on different filters that select specific event signatures. In the next sections, the selection criteria of the triggered data acquisition system of **IceCube**, relevant to this thesis, are depicted. The signal processing up to the **ICC** internally named *level 0* is outlined. In this analysis, only events which pass the MonopoleFilter_16 (**MPF16**) filter, described in Section 7.1.2.1, are considered and hence only selection criteria which are relevant to the aforementioned filter are examined.

2.2.1 Global Triggers

HLC hits happen within the **IC** and **DC** at a rate between 5 Hz to 25 Hz per **DOM**. Selection criteria are applied to select a time window to build a data acquisition event. These selection criteria are referred to as *global triggers*. Such data acquisition events are recorded with a yearly median rate of 2.7 kHz as reported in reference [Aar+17].
In this thesis, four global trigger criteria are used. Two are based on a simple multiplicity trigger (**SMT**) requirement. **SMT** means a number of **DOM**s have signaled a hit within a sliding time window of a configurable size independent of any other topological constrain but the array the hits occurred within. The other two triggers, the Volume and String triggers function the same as the **SMT** selection criteria but add topological constrains. The Volume trigger applies the **SMT** selection criteria only on a cylinder around each hit **DOM** with a radius of 175 m and a height of 75 m while the String trigger applies the **SMT** criteria to 7 consecutive **DOM**s on a string. The sliding time window sizes and number of required to be hit **DOM**s is listed in Table 2.1.

Name	Array	Hit **DOM**s	Window / µs	Topology
SMT-8	**IC**	8	5	
SMT-3	**DC**	3	2.5	
Volume	**IC**	4	1	cylinder, $r = 175$ m, $h = 75$ m
String	**IC**	5	1.5	7 consecutive **DOM**s on string

Table 2.1: **Trigger configuration settings**
*A summary of the trigger configuration for each trigger upstream of the **MPF16** filter is depicted. Individual values are taken from reference [KI14].*

2.2.2 Waveform Deconvolution

To approximate the physical time a photon has hit a **PMT** from the recorded waveforms of the associated **DOM**, the recorded waveform is deconvoluted. The characteristic response functions of the **fADC** and **ATWD** is taken into account as well as calibration data taken for the whole detector. The final result is the number of reconstructed, calibrated photons over time which caused the waveform in units of the most probable value of charge a single photon would cause (pe) as described in reference [Aar+14a]. An example of this deconvolution is depicted in Figure 2.3.

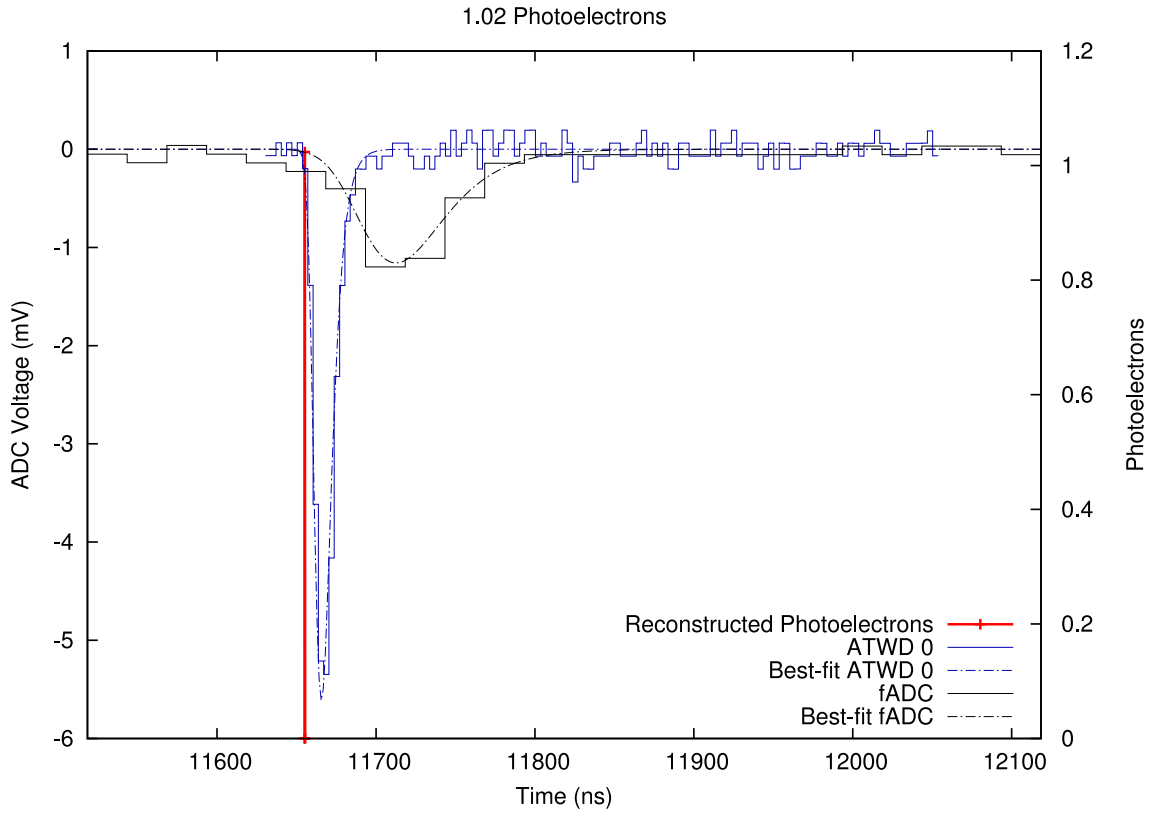


Figure 2.3: **Exemplary Waveform Deconvolution**
*A plot with the measured voltage of the **fADC** and **ATWD** (called ADC voltage in the graphic) and the reconstructed number of photons, both as a function of time is presented. Time is indicated by the shared horizontal axis, the recorded voltage of the signal, corresponding to the non-red colors in the plot, is indicated by the left hand vertical axis while the number of reconstructed photons in red is indicated by the right hand vertical axis. The same single photon waveform is recorded by the **fADC** and **ATWD** yet significant differences can be observed in the recorded response of the two digitizer systems. The **fADC**s response is much wider, lower in amplitude, and occurs later than the response of the **ATWD**. A fit is applied to both digitizer systems to derive a non-discrete description of the recorded signal. After unfolding, both systems reconstruct the time and amplitude of the initial arriving photon, here drawn in red. Graphic is taken from reference [Aar+14a].*

Chapter 3

Magnetic Monopoles

One would be surprised if nature had made no use of it.

P. A. M. Dirac [Dir31]

Magnetic monopoles (**MM**s) are sinks and sources of the magnetic field. Classically interpreted as hypothetical particles, they are carriers of elementary magnetic charge and no electric charge. Particles which carry both electric and magnetic charge are called dyons. Both are highly ionizing, stable, massive particles (**SMP**s).

In multiple theories, mathematical solutions are known that can be identified with **MM**s. While this makes the theories compatible with **MM**s, they are not required to exist. Due to the diverse nature of the theories, a wide phase space for the rest mass of **MM** exists ranging from 10^3 GeV as described in reference [Pin+09] up to or even surpassing the Planck mass m_P of 10^{19} GeV.

In this chapter, an overview of different theoretical models describing **MM**s, their expected kinematic attributes at Earth as well as exclusion limits of **MM** both by theoretical assumptions and reports by experiments is given. Equations are given in Gaussian units unless otherwise stated.

This chapter is limited to only touch upon theories related to this thesis. Velocities will be given as fractions of c (β) or, in the high relativistic regime, by the Lorentz factor $\gamma = \frac{1}{\sqrt{1-\beta^2}}$. Vectors are set in bold.

3.1 Maxwell's Equations

In 1861, J. C. Maxwell published his paper describing the interactions between the electric field, magnetic field, and electric charges in reference [Max61]. His equations feature an asymmetry between the electric and magnetic force: the magnetic field is source-free in contrast to the electric field. This is often referred to as *Gauß's Law of magnetism*. There is no fundamental physical reason why this should be the case. Assuming such sources, referred to as **MM**s, exist, Maxwells equations become symmetric for both the electric and magnetic field as seen in the following equations, given for vacuum:

$$\nabla \cdot \boldsymbol{E} = 4\pi\rho_e, \qquad \text{(Gauß's Law)}$$

$$\nabla \cdot \boldsymbol{B} = 4\pi\rho_m, \qquad \text{(Gauß's Law of magnetism)}$$

$$-\nabla \times \boldsymbol{E} = \frac{1}{\mathrm{c}}\frac{\partial \boldsymbol{B}}{\partial t} + \frac{4\pi}{\mathrm{c}}\boldsymbol{j}_m, \qquad \text{(Faraday's Law of induction)}$$

$$\nabla \times \boldsymbol{B} = \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} + \frac{4\pi}{\mathrm{c}}\boldsymbol{j}_e. \qquad \text{(Ampère's Law with Maxwell's extension)}$$

$\rho_{e,m}$ is the density of electric or magnetic charges, $\boldsymbol{j}_{e,m}$ the corresponding electric or magnetic current, $\boldsymbol{E}$ and $\boldsymbol{B}$ are the electric and magnetic flux densities. The terms marked in red are 0 in case of the non-existence of **MM**s. Their absence gives rise to the aforementioned asymmetry. Including these terms, a transformation of the form

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} = R \in \mathbf{SO}(\mathbf{2}), \tag{3.1}$$

$$\begin{pmatrix} \rho_e' \\ \rho_m' \end{pmatrix} = R\begin{pmatrix} \rho_e \\ \rho_m \end{pmatrix}, \qquad \begin{pmatrix} \boldsymbol{j}_e' \\ \boldsymbol{j}_m' \end{pmatrix} = R\begin{pmatrix} \boldsymbol{j}_e \\ \boldsymbol{j}_m \end{pmatrix}, \qquad \begin{pmatrix} E' \\ B' \end{pmatrix}, = R\begin{pmatrix} E \\ B \end{pmatrix}, \tag{3.2}$$

will leave the equations unchanged. This indicates the existence of an internal symmetry which can be described by the $\mathbf{SO}(\mathbf{2})$, the special, orthogonal group of dimension 2. Note that $\mathbf{SO}(\mathbf{2})$ is locally equivalent to $\mathbf{U}(\mathbf{1})$, the unitary group of dimension 1. Phenomena which are the result of these symmetries can manifest in different theories as long as the underlying symmetry is also present, a principal described in reference [Cur94].

3.2 Dirac Magnetic Monopoles

In 1931, P. A. M. Dirac investigated the problem of the quantization of electric charge. By studying nodal line endpoints and their phase changing effect on particle wave functions, he constructed a formation of magnetic field lines which are in parts all directional aligned and compressed to a single line, a Dirac string, described in reference [Dir31]. As the field lines are closed, they wrap at the ends of the Dirac string back and point into the direction of the other end of the Dirac string. The configuration of field lines and the resulting Dirac string are depicted in Figure 3.1.

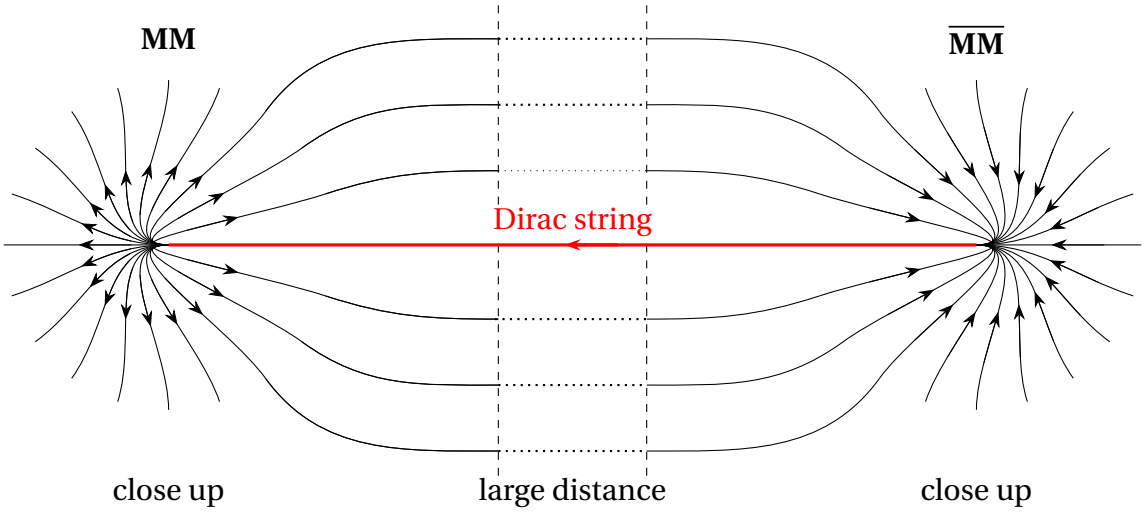


Figure 3.1: **Sketch of the magnetic field lines of an MM with Dirac string**
*A sketch of the magnetic field lines of a Dirac **MM**, shown on the left, and an Dirac anti **MM**, shown on the right, is depicted. In red, a Dirac string is shown connecting the two objects. In the center, a region bordered by two vertical dashed lines has been drawn where the spatial scaling has been changed indicating a far distance between the two objects. While both objects seem to carry a magnetic charge in the far field due to the magnetic field lines indicated by the black lines, a reverse magnetic flux is carried by the Dirac string connecting the two objects closing all of the magnetic field lines, thus, the magnetic field is still source-free.*

Without being able to detect the Dirac string, this formation matches the field line structure of a magnetic source. Such a Dirac string would cause a detectable phase change similar to the Ehrenberg–Siday–Aharonov–Bohm effect described in reference [AB59], depicted in Figure 3.2, unless

$$\mathrm{N}_{\mathrm{g_D}} = 2\frac{q_e q_m}{\hbar \mathrm{c}} \in \mathbf{Z}^+, \tag{3.3}$$

the Dirac quantization criteria described in reference [Dir31], is satisfied. In this equation, q_e and q_m are the electric and magnetic charge, $\mathbf{Z}^+$ is the set of positive integers, and $\hbar$ is the reduced Planck constant. As this is virtually independent of any assumptions apart from low level quantum mechanics, the existence of a single **MM** would require all electric charges to be quantized.

The minimal magnetic charge such a Dirac **MM** can have is called the Dirac charge g_D. By applying $\min \mathrm{N}_{\mathrm{g_D}} = \min\left(\mathbf{Z}^+\right) = 1$ to Equation 3.3, the Dirac charge can be calculated to be

$$g_D = \frac{\hbar \mathrm{c}}{2q_e} = \frac{q_e}{2\alpha} \approx 68.5 q_e \tag{3.4}$$

where α is the fine structure constant and q_e is the elementary electric charge, usually assumed to be the charge of an electron. Any **MM** has to have a multiple of the Dirac charge. As **MM**-matter interaction strength scales with the magnetic charge, the Dirac charge is used to calculate the minimal interaction strength between matter and **MM**s.

While Equation 3.4 is derived classically, the same can be done in quantum theory by examining the gauge condition of the scalar potential Φ and the vector potential $\boldsymbol{A}$. Maxwell's equations become

$$\boldsymbol{E} = -\frac{1}{\mathrm{c}}\frac{\partial \boldsymbol{A}}{\partial t} - \boldsymbol{\nabla}\Phi \tag{3.5}$$

$$\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}. \tag{3.6}$$

While $\boldsymbol{E}$ and $\boldsymbol{B}$ are physical, Φ and $\boldsymbol{A}$ can be gauge transformed utilizing an arbitrary scalar function Ξ like this:

$$\Phi' = \Phi - \frac{\partial \Xi}{\partial t} \qquad\qquad \boldsymbol{A}' = \boldsymbol{A} + \boldsymbol{\nabla}\Xi. \tag{3.7}$$

This is known as a $\mathbf{U}(\mathbf{1})$ symmetry.

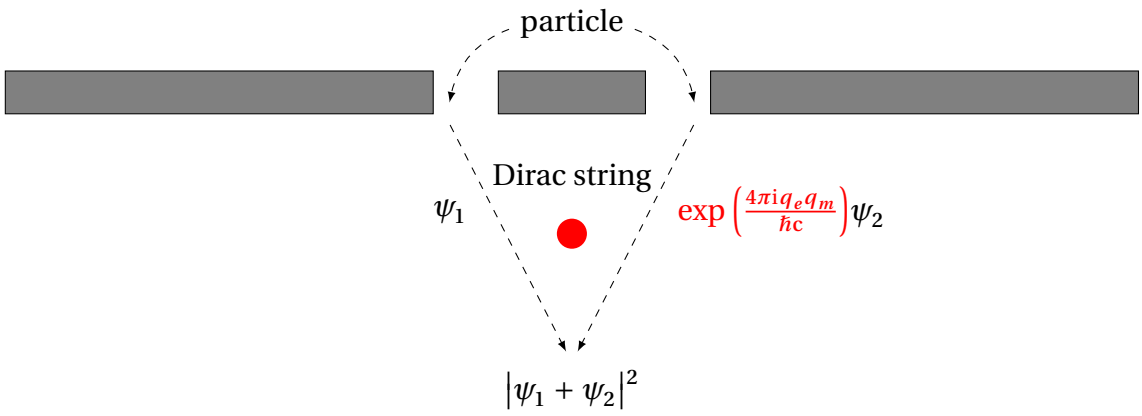


Figure 3.2: **Ehrenberg–Siday–Aharonov–Bohm setup**
A schematic of the Ehrenberg–Siday–Aharonov–Bohm effect measurement setup is displayed. A particle is sent through a double slit. $\psi_{1/2}$ denotes the wave function resulting from passing through slit 1 or 2. The superposition of the two wave functions $|\psi_1 + \psi_2|^2$ is measured. If a Dirac string is placed between the paths of the two wave functions, indicated by a red dot in the schematic, a phase shift of $\exp\left(\frac{4\pi i q_e q_m}{\hbar c}\right)$, also in red, is introduced.

Depending on the chosen gauge transformation, **MM**s are allowed, but poles are introduced into the $\boldsymbol{A}$-potential. These poles take the form of infinitely small lines and can be identified with the aforementioned Dirac string. Considering a double-slit experiment as depicted in Figure 3.2, the probability density to measure a particle behind the double-slit becomes either $|\psi_1 + \psi_2|^2$ without or $\left|\psi_1 + \exp\left(\frac{4\pi i q_e q_m}{\hbar c}\right)\psi_2\right|^2$ with a Dirac string being present. Consequently, as long as

$$\exp\left(\frac{4\pi \mathrm{i} q_e q_m}{\hbar \mathrm{c}}\right) = 1, \tag{3.8}$$

the Dirac string is undetectable. As the exp function is periodic with $2\pi\mathrm{i}$, this is true as long as

$$2\frac{q_e q_m}{\hbar \mathrm{c}} \in \mathbf{Z}^+. \tag{3.9}$$

which is the same as Equation 3.3.

3.3 't Hooft-Polyakov Magnetic Monopoles

The Higgs field is modeled as a vector field that at some point during the cooling process of the Universe started to uniformly have the same value and direction. Assuming that different domains of the Universe independently selected a random Higgs field direction which then propagated outwards, the borders of these domains will have the Higgs field point into different directions forcing a reorientation of the Higgs field in at least one of these domains. In 1974, both G. 't Hooft and A. M. Polyakov independently studied the effects such Higgs field boundaries would have. They showed that Higgs field configurations exist, like the Hedgehog configuration shown in Figure 3.3, that cannot reorient themselves into the vacuum state and hence are stable topological defect. As the Higgs field still must be smooth at all points, it has to vanish at the origin conserving the symmetry laws that existed during the Big Bang, consequently creating a non-zero density at the origin of the topological defect which imparts rest mass to the topological defect. Additionally, they derived the far-field electromagnetic field tensor resulting from this topology to be

$$F_{\mu\theta} = \frac{1}{q_e \, |\boldsymbol{r}|^3} \varepsilon_{\mu\theta a3} \boldsymbol{r}_a \tag{3.10}$$

and the resulting magnetic field becomes

$$\boldsymbol{B} = \frac{\boldsymbol{r}}{q_e \, |\boldsymbol{r}|^3}. \tag{3.11}$$

In this equation, ε is the Levi-Civita tensor and $\boldsymbol{r}$ the distance vector from the origin, the location of the topological defect, with indices 0 to 2 denoting the 3 individual components. Thus, a Hedgehog configuration exhibits a magnetic field just as a magnetic field source of twice the Dirac charge g_D as described in references [t H74; Pol74].
As the energy density solution is non-singular in the origin, a mass estimate can be derived and places the 't Hooft-Polyakov **MM** mass at roughly 100 GeV which is reachable by modern accelerators. No such **MM** was ever detected.

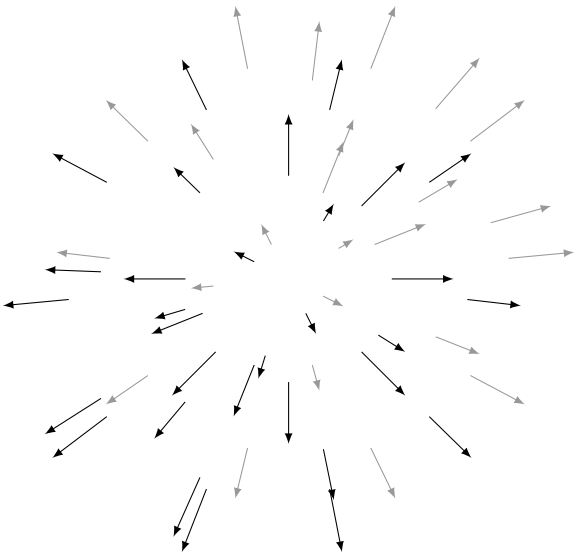


Figure 3.3: **Hedgehog Higgs field configuration**
*The configuration of the Higgs field surrounding a 't Hooft-Polyakov or Grand Unified Theory (**GUT**) **MM** is displayed. The **MM** is situated at the origin and the arrows indicate the direction of the Higgs field pointing away from the **MM**. The opaque arrows point into the page, all other arrows points out of the page.*

3.4 Magnetic Monopoles in Grand Unified Theories

In 1974, H. Georgi and S. L. Glashow published their Georgi–Glashow model of particle physics in reference [GG74]. It attempted to unify the electromagnetic, the weak, and the strong force into one Grand Unified Theory. This theory is based on an **SU**(**5**) symmetry that breaks down into a **SU**(**3**) × **SU**(**2**) × **U**(**1**) symmetry at low energies. It also predicts Higgs field configurations like the Hedgehog configuration of the 't Hooft-Polyakov **MM** but places the mass at 10^{16} GeV which is not probable by anthropogenic particle accelerators. It also predicts that the proton can decay with a live time of the order of 10^{31} yr which has already been ruled out by experiments placing lower bounds on the proton live time of 10^{34} yr as described in reference [Tan+20].
Nonetheless, it can be shown that the 't Hooft-Polyakov **MM** solutions are not an artifact of this specific theory but arise from the fundamental underlying symmetry. Any **GUT** which features an explicit **U**(**1**) subgroup allows for the existence of **MM**s as described in reference [Pre84].
In contrast to the 't Hooft-Polyakov **MM**, an **SU**(**5**) or higher dimension **GUT MM** would be able to catalyze nucleon decay. This is due to a stable high energetic region at the core of the **MM** where the **SU**(**5**) symmetry is unbroken just as it was shortly after the Big Bang as depicted in Figure 3.4.

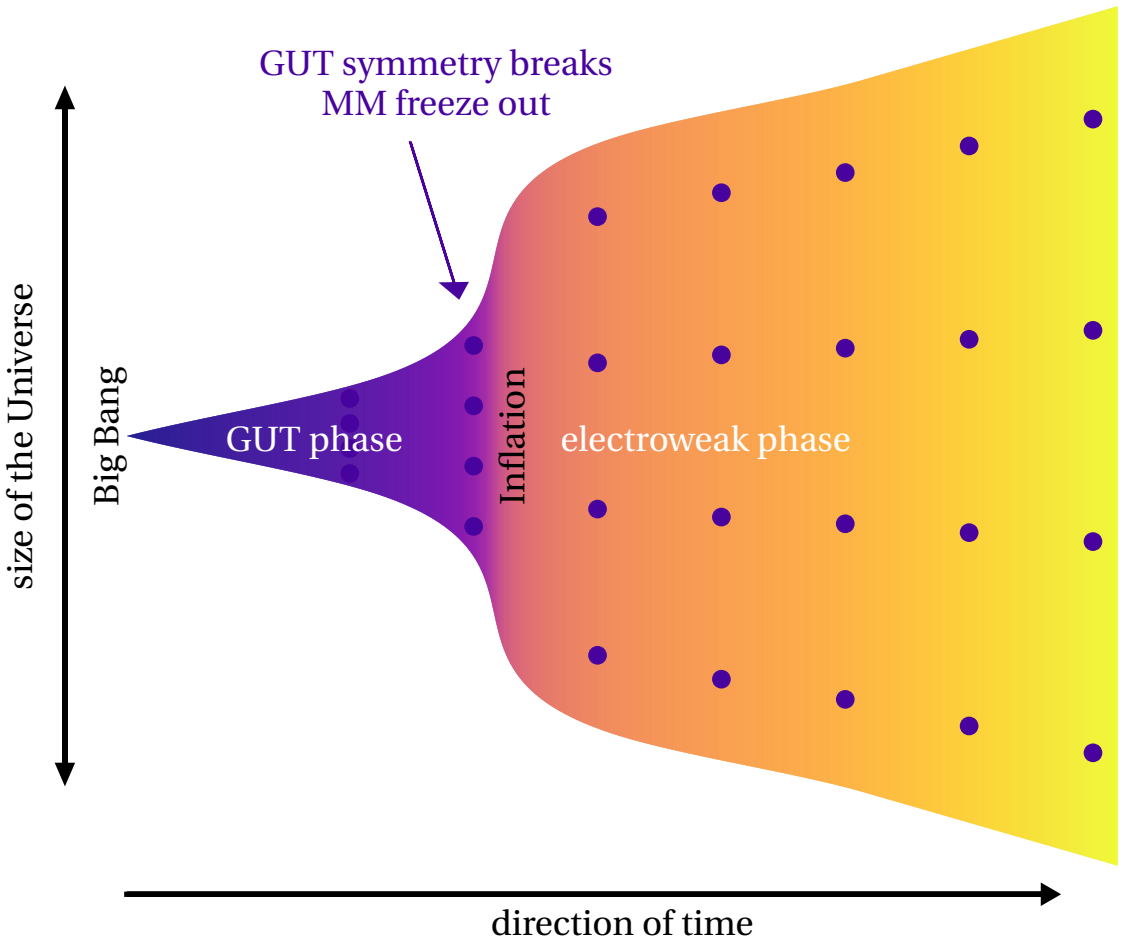


Figure 3.4: **Sketch of cosmological MM production**
*A sketch of the **MM** production during the cosmological evolution of the Universe is depicted using **SU**(**5**) **GUT** magnetic monopoles as an example for all **MM**s. Time evolution of the Universe is indicated from left to right, the size of the Universe is sketched vertically albeit arbitrary. At the beginning of the Universe, referred to as the Big Bang, it was both very small and energy dense. During the expansion of the Universe, the energy density falls, forcing the Universe into multiple phase transitions. The phase transition from the **GUT** phase, where the **SU**(**5**) is unbroken, to the electroweak phase with a **SU**(**3**) × **SU**(**2**) × **U**(**1**) symmetry is highlighted. In this case, the energy density is too low to further produce **MM** and the existing **MM**s freeze out limiting the total number of **MM**s in the Universe. Afterwards, the Universe underwent massive inflation, increasing its size by tens of orders of magnitude, dramatically reducing the density of **MM**s. Colors are used to emphasize the falling energy density in the Universe with increasing size from dark blue on the left to yellow on the right. The color of the **MM** corresponds to the energy density at the time of freeze out.*

3.5 Rest Mass

As previously stated, various theories predict varying different rest masses for **MM**s. In many symmetry group based theories, the mass of the predicted **MM** is tied to the energy Λ_{sym} at which the symmetry is broken spontaneously. The mass can then be estimated by

$$M_{MM}c^2 \gtrsim \frac{\Lambda_{sym}}{\alpha_{sym}} \tag{3.12}$$

with α_{sym} being the coupling constant in question. Thus, a rest mass of 10^2 GeV to 10^5 GeV is expected for electroweak **MM** like the 't Hooft-Polyakov **MM** and more than 10^{13} GeV is expected for **GUT** scale **MM**s like the **SU**(**5**) **GUT MM**s. There are also theories in which the symmetry is not broken directly but in multiple steps which allows for masses between 10^5 GeV to 10^{13} GeV, referred to as intermediate mass magnetic monopoles (**IMM**s). Much of this mass phase space cannot be probed by anthropogenic particle accelerators and consequently, searches can only be performed for relic **MM**s which were produced during the Big Bang as sketched in Figure 3.4.

3.6 Kinetic Energy

Relic **MM**s are assumed to have negligible kinetic energy at their creation as they would be created in thermal equilibrium with the surrounding Universe. During their propagation through the Universe, they are accelerated by magnetic and gravitational fields. As gravitational field acceleration only becomes dominant for **MM** with rest masses larger than 10^{17} GeV, they are ignored in this work. The total gained energy by magnetic fields can be calculated by

$$E_{\text{kin}} = q_B \cdot B \int_C \boldsymbol{\xi} \cdot d\boldsymbol{v} \tag{3.13}$$

where q_B is the magnetic charge of the **MM**, B is the magnetic field strength, $\boldsymbol{\xi}$ is the field's direction and coherence length, C is the propagation path, and $d\boldsymbol{v}$ is the normalized **MM**s velocity vector at any given point of C as described in reference [Wic+03]. Assuming the **MM** is always traveling parallel to $\boldsymbol{\xi}$, the equation reduces to

$$E_{\text{kin}} = q_B \cdot B \cdot |\boldsymbol{\xi}| \tag{3.14}$$

and allows for a classical estimate of the expected velocity after a single pass through a magnetic field by

$$E_{\text{kin}} = q_B \cdot B \cdot |\boldsymbol{\xi}| \Leftrightarrow \tag{3.15}$$

$$\frac{1}{2} M_{MM} \cdot v_{MM}^2 = q_B \cdot B \cdot |\boldsymbol{\xi}| \Leftrightarrow \tag{3.16}$$

$$v_{MM} = \sqrt{\frac{2 \cdot q_B \cdot B \cdot |\boldsymbol{\xi}|}{M_{MM}}} \,. \tag{3.17}$$

Estimated values for the coherent field length and field strength for different fields, published in reference [Wic+03], are given in Table 3.1.

	B/µG	$\lvert\boldsymbol{\xi}\rvert$ /Mpc	$E_{\text{kin}}/\text{N}_{\text{g}_\text{D}}$/GeV
normal galaxies	$3-10$	10^{-2}	$(0.3-1)\cdot 10^{12}$
starburst galaxies	$10-50$	10^{-3}	$(1.7-8)\cdot 10^{11}$
active galactic nucleus jets	≈ 100	$10^{-4}-10^{-2}$	$1.7\cdot\left(10^{11}-10^{13}\right)$
galaxy clusters	$2-30$	$10^{-4}-10^{-1}$	$3\cdot 10^{11}-5\cdot 10^{14}$
extragalactic sheets	$0.1-1.0$	$1-30$	$1.7\cdot 10^{13}-5\cdot 10^{14}$

Table 3.1: **Estimated magnetic fields in the Universe**
*A table with different magnetic fields which are expected to contribute to the acceleration of **MM**s is presented. E_{kin} is given for a single pass through the field. Adapted from reference [Wic+03].*

An **MM** can be accelerated by the same field multiple times if it is trapped due to electric fields or gravity (e. g., inside galaxies) or passed through multiple similar fields due to the size of the environment (e. g., extra galactic fields). For the extra galactic domain, this effect increases the maximal kinetic energy to $E_{\text{kin}_{\text{max}}} = 10^{15}$ GeV as described in reference [Wic+03].
For this work, a velocity profile for **IMM** as expected at a detector situated at Earth of

$$\beta = \begin{cases} \text{c} & M \lesssim 10^{11}\,\text{GeV} \\ \text{c}\left(\frac{10^{11}\,\text{GeV}}{M}\right)^{\frac{1}{2}} & M \gtrsim 10^{11}\,\text{GeV} \end{cases} \tag{3.18}$$

is assumed, taken from reference [Tan+18]. This corresponds to an acceleration scenario of an **MM** of Dirac charge being accelerated from rest by the electric fields in the Milky Way galaxy. In Figure 3.5 the resulting velocity profile in the transition zone for different rest masses from the high relativistic regime to the non-relativistic regime can be observed.

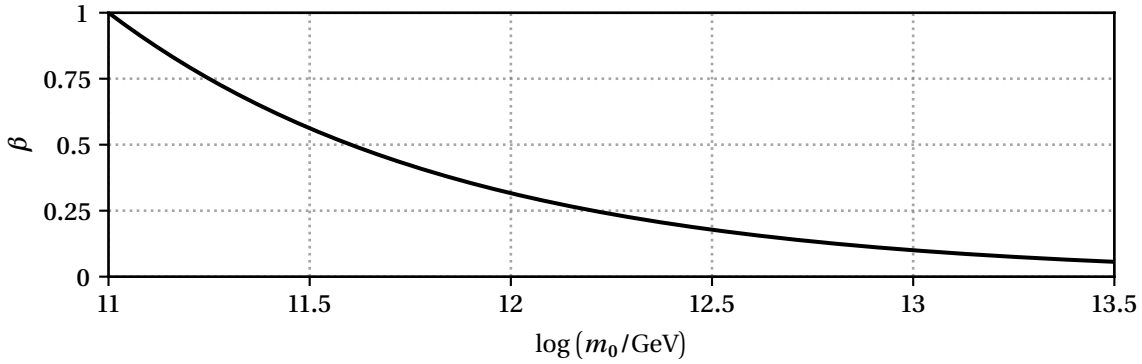


Figure 3.5: **Velocity of IMMs in dependence of rest mass**
*A plot with the expected β as a function of the rest mass of an **MM** in the intermediate mass range following Equation 3.18 is depicted. The horizontal axis is logarithmic while the vertical axis is linear. Below 10^{11} GeV, the expected velocity is close to c. The velocity drops to about 0.6 c at $10^{11.5}$ GeV, passes below 0.2 c at $10^{12.5}$ GeV, and converges towards 0 for the non-relativistic regime heavier than 10^{14} GeV.*

3.7 Stopping Power

MMs passing through matter lose parts of their kinetic energy to said matter. The energy loss per path length depends on the Dirac charge, passed material, and velocity. Different velocity ranges are dominated by varying energy losses. Starting at the high relativistic regime where $\gamma \cdot \beta \geq 1000$, the energy loss is dominated by stochastic loses. As this regime has been covered in multiple searches before and is of little interest to this work, the energy loss is excluded from description in this thesis. Between 0.05 c and $\gamma = 100$, **MM**s dominantly lose kinetic energy by excitation of electrons and the resulting ionization of the target material. The energy loss in this velocity range can be described by the Bethe-Bloch formula adjusted to magnetic charges as described in reference [Ahl78]

$$\begin{aligned}\frac{dE}{dx} = \frac{\pi \rho_e \hbar^2 \mathrm{N}_{\mathrm{g_D}}^2}{m_e} \Bigg[& \ln\left(\frac{2m_e \left(\mathrm{c}\beta\gamma\right)^2}{I_{\mathrm{target}}}\right) - 0.5 && (3.19)\\ & + \frac{k\left(\mathrm{N}_{\mathrm{g_D}}\right)}{2} && \text{(quantum electrodynamics correction)}\\ & - \frac{\delta\left(\log_{10}\left(\beta\gamma\right)\right)}{2} && \text{(density correction)}\\ & - B_m\left(\mathrm{N}_{\mathrm{g_D}}\right) && \text{(Bloch correction)}\\ \Bigg] & \end{aligned}$$

where $\rho_e = \rho_{\text{target}} \langle Z/A \rangle_{\text{target}}$ is the electron density of the target material, m_e is the mass of the electron, $\mathrm{N}_{\mathrm{g_D}}$ is the aforementioned number of Dirac charges the **MM**s possess, and I_{target} is the mean ionization energy of the target material.

$$B_m\left(\mathrm{N}_{\mathrm{g_D}}\right) = \begin{cases} 0.248, & \text{for } \mathrm{N}_{\mathrm{g_D}} = 1, \\ 0.672, & \text{for } \mathrm{N}_{\mathrm{g_D}} = 2, \\ 1.022, & \text{for } \mathrm{N}_{\mathrm{g_D}} = 3, \\ 1.685, & \text{for } \mathrm{N}_{\mathrm{g_D}} = 6, \\ 2.085, & \text{for } \mathrm{N}_{\mathrm{g_D}} = 9, \end{cases} \tag{3.20}$$

is the Bloch correction, described in references [Blo33; Der+98], and

$$k\left(\mathrm{N}_{\mathrm{g_D}}\right) = \begin{cases} 0.406, & \text{for } \mathrm{N}_{\mathrm{g_D}} = 1 \\ 0.346, & \text{for } \mathrm{N}_{\mathrm{g_D}} = 2 \\ 0.300, & \text{for } \mathrm{N}_{\mathrm{g_D}} \geq 3 \end{cases} \tag{3.21}$$

is the quantum electrodynamics correction, described in reference [Der+98]. $\delta\left(\log_{10}\left(\beta\gamma\right)\right)$ is the density correction due to the polarization of real media which can be modeled via the Sternheimer model described in reference [Ste81]

$$\delta\left(x\right) = \begin{cases} \delta_0 10^{(2(x-X_0))} & \text{for } x < X_0 \\ 2\ln\left(10\right)x - C + a(X_1 - X_0)^m & \text{for } X_0 <= x < X_1 \\ 2\ln\left(10\right)x - C & \text{for } X_1 <= x \end{cases} \tag{3.22}$$

where X_0, X_1, C, δ_0, and a are material specific constants derived from direct measurements of the materials. In this work, the values from reference [SBS84] are used.

Between 10^{-3} c to 10^{-2} c, the interaction between the **MM** and the material can be described by a free, possibly degenerate, electron gas. The energy loss can then be described by

$$\frac{dE}{dx} = \frac{\pi\rho_e \mathrm{N}_{\mathrm{g_D}}^2 \hbar c \beta}{2\left(3\pi^2\rho_e\right)^{\frac{1}{3}}}\left(\ln\left(2a_0\left(3\pi^2\rho_e\right)^{\frac{1}{3}}\right) - 0.5\right), \tag{3.23}$$

taken from reference [AK82], where a_0 is the Bohr radius and all other factors are the same as in Equation 3.19. For the velocity range between 10^{-2} c and 0.05 c, the energy loss was linearly interpolated for this work.

Below 10^{-3} c, **MM**s cannot excite atoms anymore but still lose energy via elastic collision. As **MM**s in this velocity range cannot be detected in this work, they are assumed to immediately stop within the target material without being detectable.

The resulting energy loss for different materials in the velocity ranges where energy loss has been modeled is depicted in Figure 3.6. Given an energy loss, the distance an **MM** can travel within a medium can be calculated by

$$R = \int_{E_{\text{min}}}^{E_0} \frac{1}{dE/dx} dE \tag{3.24}$$

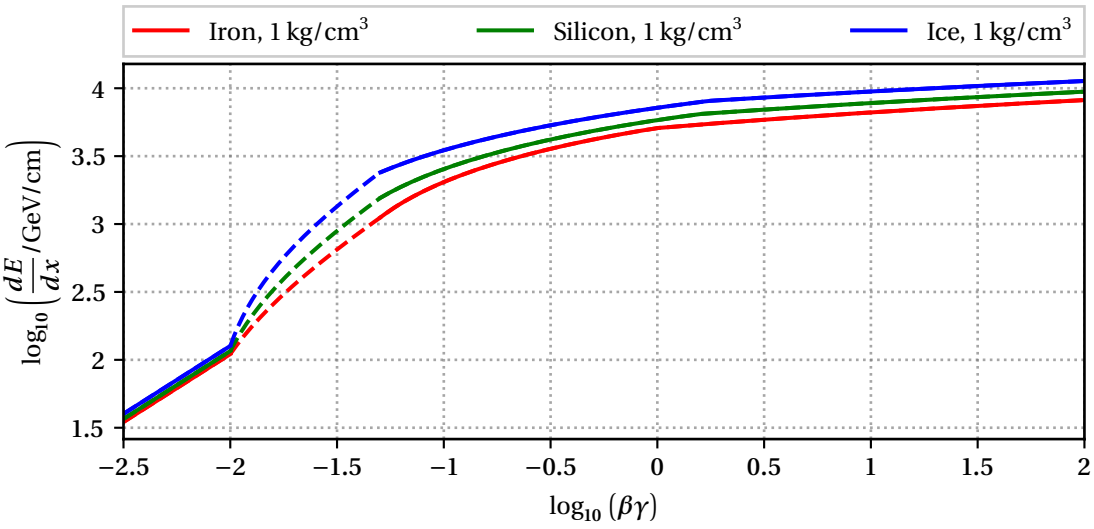


Figure 3.6: **Energy loss of an MM in different media**
*In the plot, the logarithmic energy loss of an **MM** of Dirac charge in GeV cm^{-1} as a dependence of the velocity expressed in logarithmic $\beta\cdot\gamma$ is depicted. As example materials, iron (red), silicon (green), and ice (blue) were chosen. The density was kept at 1 kg cm^{-3} for all three materials. The dashed zone marks the region of linear interpolation between Equations 3.19 and 3.23. The shape in the left and right speed regime is roughly the same for all materials. However, there is a differing offset between them which widens in the intermediate regime. The highest energy loss is seen in ice, second is silicon, and last is iron. This is mainly due to fixing the density for all materials. The energy loss linearly scales with the density.*

where E_{min} is the energy there the particle is assumed to be at rest. E_0 is the initial energy and $\frac{1}{dE/dx}$ is the inverse energy loss.

3.8 Earth Shielding

MMs traversing the Earth lose energy due to matter interactions. Consequently, their velocity decreases. Depending on the amount of Earth passed and the sensitivity to the velocity of a traversing **MM**, this may result in a measurable different velocity of an **MM** or even shield certain directions from incoming **MM**s.
In order to model this velocity difference for this work, a two material composition model for the Earth with only a radial dependency is assumed. In this model, the core of the Earth consists of iron and the outer crust is composed of silicon with a linear transition zone in between. The density profile of Earth is taken from reference [LM01]. The resulting material densities are depicted in Figure 3.7.

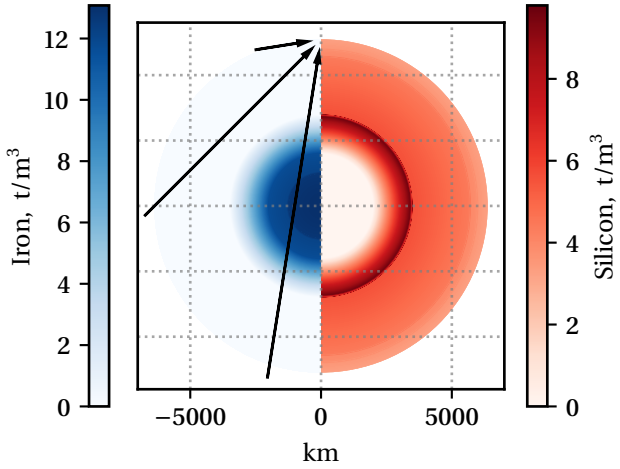


Figure 3.7: **Density and composition model of Earth**
*Model of Earth used for calculating the expected energy loss for **MM** at the in-ice arrays of **IceCube**. Left hand, blue shaded region is indicative of the density of iron, right hand, red shaded region is indicative of the density of silicon. The model starts out with a high density iron core and ends with a silicon outer shell. In between, the composition is linearly interpolated. The model is only dependent on the radius and assumes a perfectly spherical planet. The black arrows depict possible exemplary trajectories of **MM**s and point at the depth of the in-ice arrays of **IceCube**.*

The shielding depends on the mass in the path of the incident particle which varies as a function of the zenith angle in the aforementioned spherical Earth model. On the other side, there are free parameters from the model of the **MM**, namely the number of Dirac charges N_{g_D}, its rest mass m_0, and kinetic energy E_{kin}.
In Figure 3.8, the expected velocity profile at the in-ice arrays of the IceCube Neutrino Observatory (**IceCube**) assuming an initial velocity mass profile as described in Equation 3.18 corresponding to an **IMM** being accelerated by the Milky Way galaxy is depicted.

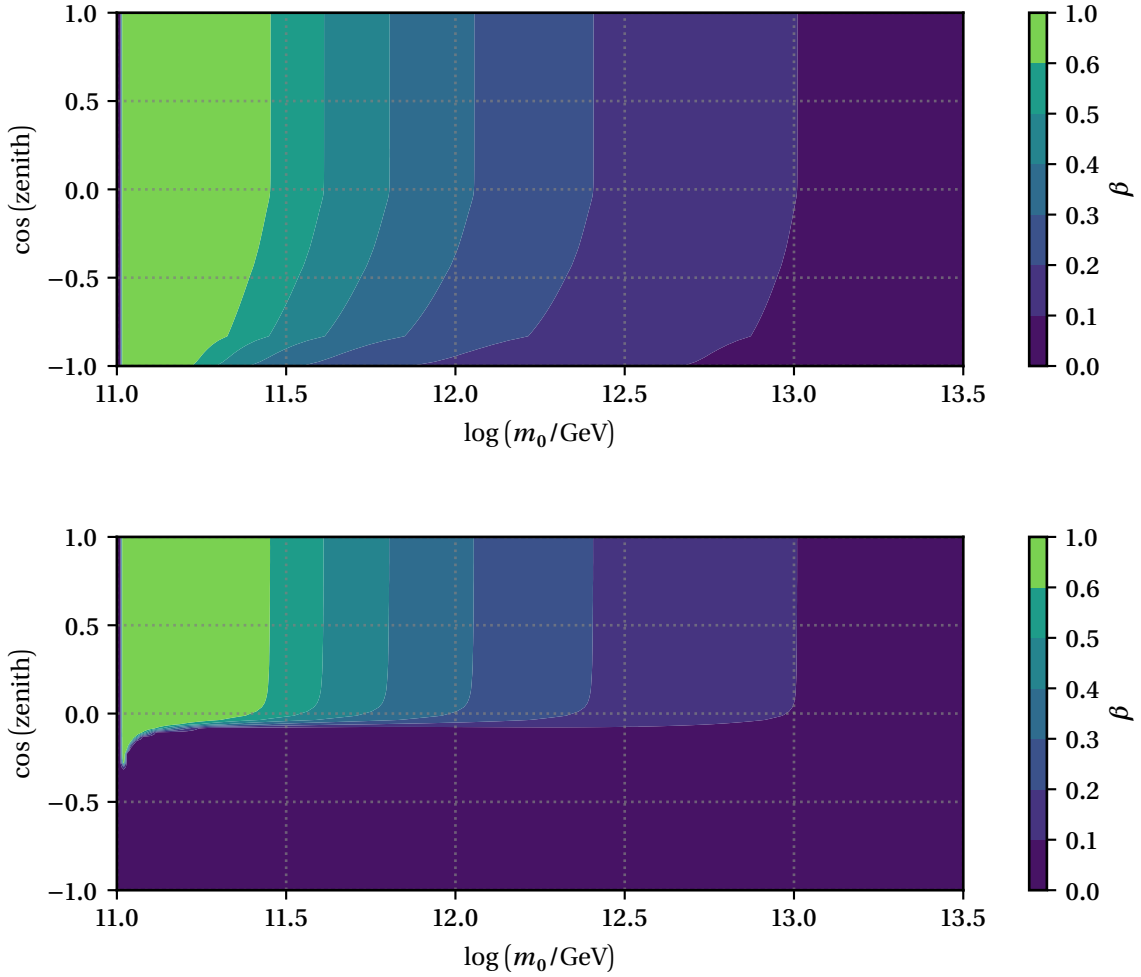


Figure 3.8: **Velocity of IMMs at IceCube for N_{g_D} = 1 and N_{g_D} = 9**
*In the two figures, the velocity of **IMM** at the in-ice arrays of **IceCube** as a function of their rest mass and zenith angle are on display. The first figure has been calculated assuming an **MM** with $N_{g_D} = 1$, the second figure with $N_{g_D} = 9$. Earth shielding depends on the magnetic charge, thus $N_{g_D} = 1$ represents a low bound for the shielding of Earth while $N_{g_D} = 9$ is presented to make the shielding effect of the Earth more prominent. In both figures, the horizontal axis is indicative of the rest mass of an incident **MM** in logarithmic presentation while the vertical axis is used to indicate different zenith angles in the coordinate system of **IceCube**, depicted in Figure 6.3. For each combination, an initial velocity as described in Figure 3.5 is assumed. For computing purposes,* c *is treated as* $\gamma = 1000$. *The **MM** is propagated through the Earth following the density and composition model described in Figure 3.7. Its arrival velocity at the in-ice arrays of **IceCube** is indicated by color, corresponding to a velocity as noted at the right hand color bar. Note that the color coding is not uniform but is deliberately fine grained in the sensitive region of this analysis (0.1 c to 0.6 c). For both N_{g_D}=1 and N_{g_D}=9, nearly no velocity change is observed for down going **MM**s. For up going **MM**s, total shielding is observed for N_{g_D}=9 while **MM**s with N_{g_D}=1 slow down slightly. Additional intermediate magnetic charges have been calculated and can be seen in Figures B.1 to B.7.*

3.9 Exclusion Limits for Magnetic Monopole Flux

Observations of the current state of the Universe limit the possible number of **MM**s in existence in the observable Universe. One such observation is the detection of a magnetic field inside the Milky Way galaxy. **MM**s are accelerated by magnetic fields as described in Section 3.6 and thus drain energy from magnetic fields. If the flux of **MM**s is too high, no magnetic field inside the Milky Way could be observed as they would have been completely drained. This is called the *Parker bound*, described in reference [Par70]. Thus, the exclusion limit of **MM** flux at the Earth is

$$F < \begin{cases} 10^{-15}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-2}, & M_{\mathrm{MM}} \lesssim 10^{17}\,\mathrm{GeV}, \\ 10^{-15}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-2}\,\frac{M_{\mathrm{MM}}}{10^{17}\,\mathrm{GeV}}, & M_{\mathrm{MM}} \gtrsim 10^{17}\,\mathrm{GeV}. \end{cases} \tag{3.25}$$

The same reasoning can be applied to all current or past magnetic fields observed in the Universe, but dependencies on the specific model of **MM**, especially the assumed rest mass of the **MM**, become more prominent making the bounds less universal. In the low relativistic to highly relativistic regime, the Parker bound is already superseded by current exclusion limits for **MM**s flux as can be seen in Figure 3.9.

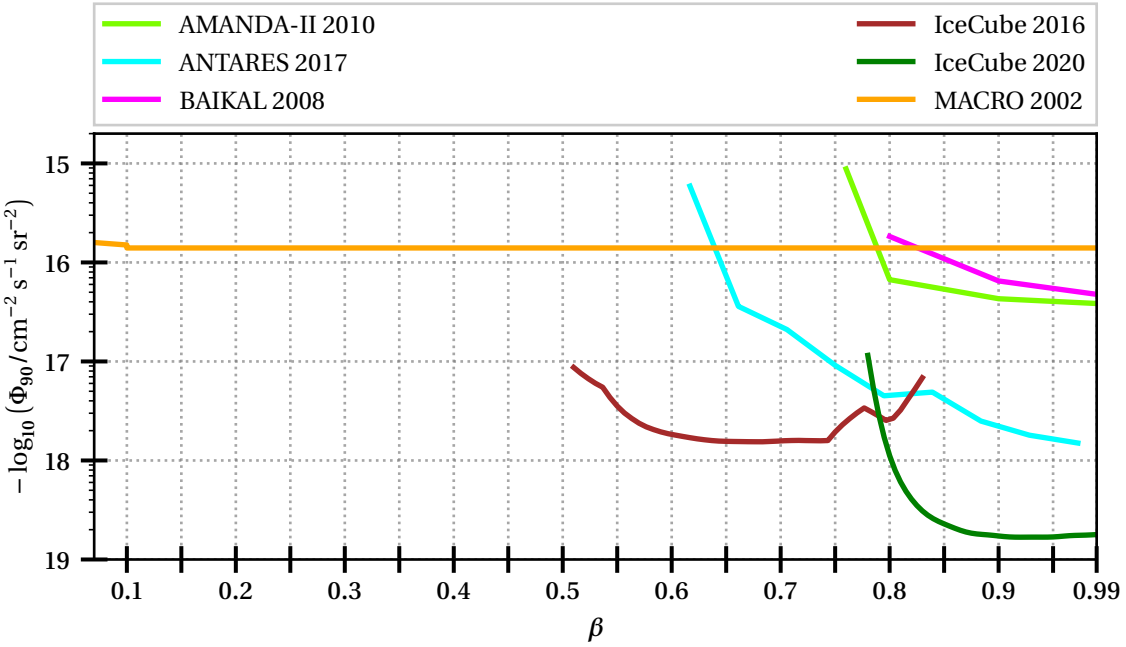


Figure 3.9: **Previous flux limits for MMs**
*A plot with the result of previous searches for **MM**s is depicted. The horizontal, linear axis is the velocity in units of c from 0.1 to 0.99. The vertical, logarithmic axis is the excluded level of flux for **MM** as reported by different experiments and collaborations in references [Abb+10b; Alb+17; Ayn+08; Aar+16; Bur20; Amb+02], sorted in the legend by alphabetical order.*

Chapter 4

Machine Learning

Machine learning, in the simplest terms, is the analysis of statistics to help computers make decisions based on repeatable characteristics found in the data.

V. K. Agrawal [Agr20]

Machine learning (**ML**) is a subfield of artificial intelligence. This subfield is limited to deriving or *training* estimators, here called *predictors* from a set of data points, called the *training set*, with the goal of generalized performance towards disjunct data points. All points are part of the same feature space, the space spanned by the product of all used features. In this thesis, **ML** is used in the final analysis step to select candidate events. This is called binary regression or *binary classification*. The predictors are trained based on data points with known classes. This subclass of **ML** is called supervised learning. In this chapter, the required nomenclature and methods used for that step is explained.

4.1 Feature Space

A *feature* of an event is any information derived reproducible from said event. In this thesis, every feature is a scalar value. Multiple features can be derived from a single event.

Feature space is the abstract space where all possible points a predictor can be applied to reside in. It is the product of the individual features which span the space.

For any **ML** based algorithm, choosing proper features is crucial. Adding more features increases the size of the feature space. As processing time is limited, usually only a finite number of data points within the feature space can be utilized. This can result in the density of sampled data points in sub-samples to vary wildly and in turn influence the performance of a predictor towards different sub-samples of the feature space which is not desirable. Furthermore, features influence the maximal performance achievable by a predictor.

4.2 Sampled Supervised Binary Classification

Sampling is the process by which discrete data points are derived from an underlying probability density function. In this thesis, only the case of finite numbers of sampled points is discussed.

Binary classification is the process by which a sampled data point is assigned one of two possible labels or *classes*. In this thesis, these classes are *background-like* or *signal-like*.

Once a data point is classified, it can be removed if it is background-like or selected as a candidate event for further use. The class of a data point may be known or unknown depending on the origin the data point is derived from, e. g., simulated events have known classes while measured events have unknown classes.

In supervised learning, predictors are trained using data points with known or assumed class. After the training has concluded, the predictors are applied to data points with unknown class assigning each data point a probability or *score value* between 0 and 1 where 0 stands for background-like while 1 stands for signal-like.

A predictor approximates the mapping $\mathfrak{M}_{\mathrm{TS}} : \boldsymbol{x}_i \rightarrow c_i$ between the data points $\boldsymbol{x}_i$ and their associated class $c_i \in \{0, 1\}$ within a training set $\boldsymbol{x}_i \in h_{\mathrm{TS}}$. It is crucial that this mapping $\mathfrak{M}_{\mathrm{TS}}$ is representative for all possible data points. This can be achieved by randomly selecting the data points from all possible points. If only subspaces are utilized, special precautions are required to ensure that the subspaces have the same underlying mapping.

4.2.1 Bias–Variance Trade-Off

The bias of a predictor towards a given set of data points is the average offset between the predicted class of all data points within the set by the predictor and the true class of these data points. In the ideal case, the bias towards all possible data points is minimal, i. e., the predictor has the best mean performance for all possible data points. Thus, the variance of this bias in different sub-samples of all possible data points becomes minimal as well. The absolute value of this bias is influenced by the structure and thus complexity of a given predictor.

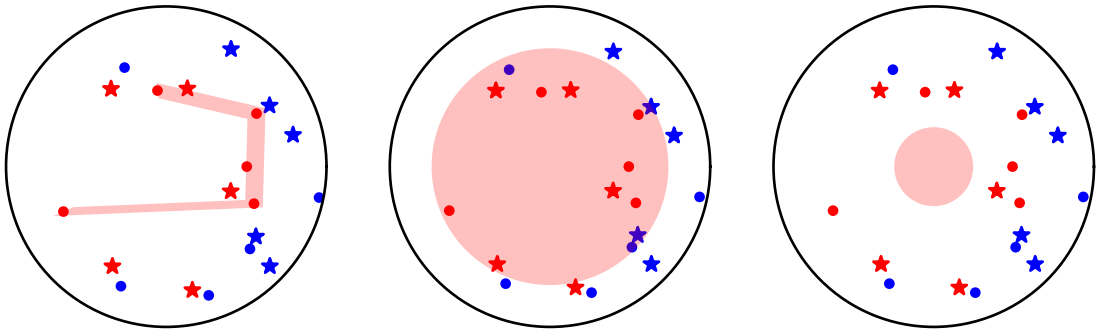
During the training of a predictor, only a finite set of data points can be regarded. Consequently, the aforementioned average bias and variance for all possible data points cannot be determined. Instead, only the bias towards a specific subset, e. g., the training set, and the variance between different sets, e. g., the training set and a disjunct comparison set, can be derived.

The bias for a specific set of data points can be driven to zero with high enough complexity of a predictor and consequently can be lower than the theoretical minimal bias towards all possible data points. Such a predictor is overspecialized towards this specific set of data points, also referred to as *over-trained*. This over-training is accompanied by a high variance of the predictor.

On the other hand, a trivial, low complexity predictor, e. g., a predictor with a constant predicted value, usually has a high bias but a low variance. Such a predictor is called *under-trained*.

The training of a predictor needs to be balanced between the under-trained and over-trained regime. The first will generate a predictor with a sub-optimal performance, also called a *weak* predictor. The latter will result in a predictor with apparent great performance, even better than the unknown theoretical optimum, but which does not generalize outside the training set. Examples of predictors with different biases and variances can be seen in Figure 4.1.

In this thesis, this balance is accomplished by adding regularization systems during the training phase. Regularization systems limit the degree of complexity a predictor can achieve by penalizing high complexity over low complexity.



(a) *over-trained* (b) *ideal* (c) *under-trained*

Figure 4.1: **Predictors in different bias and variance regimes**
Three polar plots are presented, containing the same data points. Two classes of data points are used, represented by the colors red and blue. For each class, two sets of data points are sampled, one set used to train the predictor, the other set to compare the performance of the predictor outside the training set. These sets are illustrated by either a circle, indicating the training set, or a star, indicating the comparison set. Each data point has two features: the radius and the angle. The radius is sampled from a uniform probability distribution ranging from 0.5 to 1 for the red class while the blue class is uniform between 0.8 and 1.3. The angle is randomly distributed and hence does not allow for any separation power between the two classes. For each plot, a binary class predictor has been trained on the same data points. The training result is indicated by a red shaded region to indicate areas in which data points are assigned to the red class. The left plot is an example of a highly over-trained predictor. All training data points are correctly identified and thus the bias is zero, but testing the predictor on the comparison set yields a high bias result. Thus, the variance of the predictor is high. The predictor on the right is under-trained. No data point in the training set is correctly identified, a performance matched by the comparison with the test set. While the bias of the predictor is high for both sets, the variance is low. In the middle plot, an ideal predictor is presented as a comparison to the previous two predictors. In this case, both the bias and the variance are as small as theoretical possible given the aforementioned distribution of the radius.

4.2.2 Training Goal: Minimize the Objective Function

The *objective function* is the mathematical description of the objective to be achieved. During training, the goal is to converge the predictor to a global minimum of the objective function. Depending on the objective function, different minima with different performance characteristics for the predictor may exist.

In this thesis, the objective is to correctly predict the class of all data points while only a subset of data points is available. The objective function is the sum of two subfunctions, the *loss function* and *regularization function*.

The *loss* is the qualification of how well the predictions of a predictor match the classes of data point. The *loss function* is the algorithm to derive this loss. In this thesis, an estimator of the cross entropy, H_P, is used as the loss function.

The cross entropy H, first mentioned in reference [Goo56], is an extension of the Kullback–Leibler divergence described in reference [KL51]. It is a metric for the distance between two probability measures O and P with the same topological support Q. It is defined as

$$H(O,P,Q) = -\sum_{q\in Q} o(q)\cdot \log p(q) \tag{4.1}$$

where o and p are the respective probability density functions of O and P. Assume O is an unknown probability measure and a finite number N_q of randomly sampled points Q_O following o are available. Then, an estimator of the cross entropy $H_P(Q_O)$ can be defined as

$$H_P(Q_O) = -\frac{\sum_{q\in Q_O} o(q)\cdot \log p(q) + (1-o(q))\cdot \log(1-p(q))}{N_q}. \tag{4.2}$$

By identifying

P as the probability measure of the predictor and thus
$p(q) \in [0,1]$ as the probability density of the predicted binary class,
$o(q)$ as the true class label, and
Q_O as the training set,

it can be observed that minimizing $H_P(Q_O)$ lowers the differences between the predictions of the predictor and the class labels as presented by the sampled data points used for training. This approximates the underlying, unknown class distribution $\mathfrak{M}_{\mathrm{TS}}$ that the data points in the training set were sampled from.

The *regularization function* is a penalty applied to a given predictor based on various factors. These penalties are applied in order to limit the complexity of the predictor or to facilitate convergence of the predictor. In this chapter, the regularization function itself is split into different functions and techniques described in Section 4.4.

4.3 Machine Learning Based Estimators

Different algorithms exist in **ML** to build predictors. In this section, the algorithms relevant for this thesis are summarized. Only binary classification is regarded.

4.3.1 Singular Variate Decision

The singular variate decision (**SVD**) separates a set of data points h into two subsets $h_{l/r}$ based on the value of a single feature f and a defined value c via

$$h_l = \{x \in h | f(x) \leq c\} \tag{4.3}$$

and

$$h_r = \{x \in h | f(x) > c\}. \tag{4.4}$$

The class represented by h_l and h_r can vary. In this thesis, c is called a *cut* value. Multiple techniques exist to determine c. Two different techniques are used in this work, described in Section 4.5 and Section 7.3.6.

4.3.2 Decision Tree

A set of predictors can be combined into a single predictor with a higher performance than the sum of the individual predictors as described in reference [Sch90]. A decision tree (**DT**) is an ensemble of multiple, weak **SVD**s arranged in a binary tree. It can be applied to multifeature data points albeit only a single feature is utilized in an **SVD**. Different strategies exist to select a feature for each **SVD**. The algorithm used in this thesis is described in Section 4.5.
Each **SVD** forms a *branch* of the tree, splitting the training set into subsets which in turn the next **SVD**s are based upon. A subset which is not split by an **SVD** is called a *leaf*.
An example of a binary tree is depicted in Figure 4.2. The maximal number of consecutive **SVD**s to reach all leafs within a **DT** is called the depth of the **DT**.

4.3.3 Boosted Decision Trees

Similar to a **DT** being an ensemble of **SVD**s, a boosted decission trees (**BDT**) is an ensemble of consecutively added **DT**s forming a single predictor.
The **DT**s are built in succession, adjusting the objective to be optimized by each individual learner by altering or *boosting* the training data based on the predictions of the previously built ensemble of **DT**s. Various strategies for boosting exist, the technique utilized in this thesis is described in Section 4.5. A diagram of the process is presented in Figure 4.3.

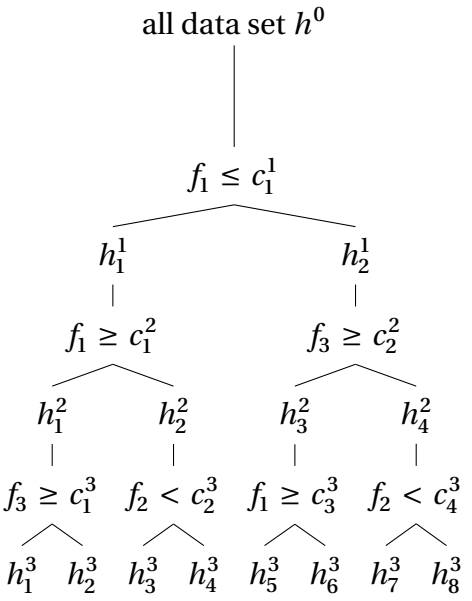


Figure 4.2: **Example of a DT**
*An example of a **DT** with a depth of three is depicted. Data points with three features, f_1, f_2, and f_3 are assumed. At the top, all data points used for training are within one set h^0. A feature f_1 is used to split h^0 into two subsets, h_1^1 and h_2^1 depending if the data points have a f_1 value below of c_1^1. This is repeated twice, resulting in a total of eight subsets.*

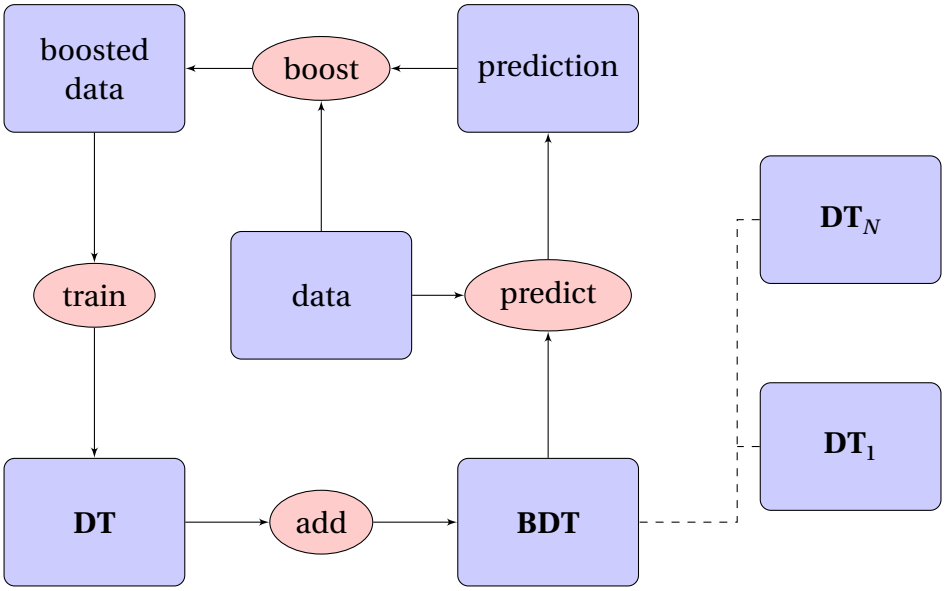


Figure 4.3: **Flow diagram of the principal of BDT training**
*A flow diagram of the iterative training process of a **BDT** is presented. As the first step, a **DT** is trained based on the training data. The resulting **DT** is added to the **BDT**. The predictions of the **BDT** for the training set is calculated. The predictions and training data are compared and the training data is boosted based on the comparison. A new, boosted training data set is generated. Another **DT** is trained on the boosted data and added to the **BDT**. The cycle repeats until a stop condition is met.*

4.4 Regularization Systems and Techniques

Different techniques and systems exist for regularization of the training depending on the task at hand. These can vary from penalizing high complexity in a predictor or ensemble of predictors, defining a stopping condition to training, or augmenting the available training data. This results in a minimal reachable bias of the predictor towards the training set instead of allowing the bias to reach zero, and thus, as it has been discussed in Section 4.2.1, keeps the predictor from becoming over-trained.

4.4.1 Number of Predictors

One of the basic regularization for ensembles of predictors is the number of predictors within the ensemble. For **DT**s, this is called the maximal depth as explained in Section 4.3.2. For **BDT**s, it is called the maximal number of **DT**s. Note that this describes only the maximal number of possible predictors, other techniques may limit the number of added predictors based on different factors. For example, early stopping can be utilized while building a **BDT** to stop adding **DT**s once no further improvement of the performance of the **BDT** has been observed for a given number of added **DT**s. In this case, the **BDT** is assumed to have already reached the theoretical best performance given the regularization function.

4.4.2 Sub-sampling the Training Data

Sub-sampling is a technique in which only a fraction $S \in (0, 1]$ of the training set is utilized to build the individual predictors of an ensemble of predictors. Every time, sub-sampling is applied, each data point has a probability of S to be selected. As not every data point is included in the training of the predictor, potential over-trained regions of feature space in each predictor average out while regions common in all predictors get amplified.

4.4.3 Dropout Regularization

Dropout regularization, originally proposed as a regularization for Neural Network based **ML** applications in reference [Hin+12], is a technique in which a random set of previously trained predictors in an ensemble is disregarded for a given prediction.

In this thesis, dropout regularization is applied during the training of **BDT**s, described in Section 4.3.3. An example flowchart of dropout during the training of a **BDT** is depicted in Figure 4.4. It is described by two parameters: the probability to apply the regularization $D \in [0, 1]$ and the probability for each **DT** to be ignored in the prediction $R \in (0, 1]$. The sampling of **DT**s to be ignored is performed in a weighted fashion, resulting in higher performance predictors being dropped more often and therefore equalizing the overall performance over all **DT**s. As a performance metric, the overall reduction of the objective function per **DT** is utilized.

Without this form of regularization, early predictors in the training achieve high performance while subsequent added predictors contribute less. By only selecting a subset of predictors, overall beneficial additions to the predictors are included in multiple predictors, spreading the performance over all predictors while suppressing overtraining additions in early predictors. Additionally, settling on a local minimum of the objective function over the global minimum is suppressed.

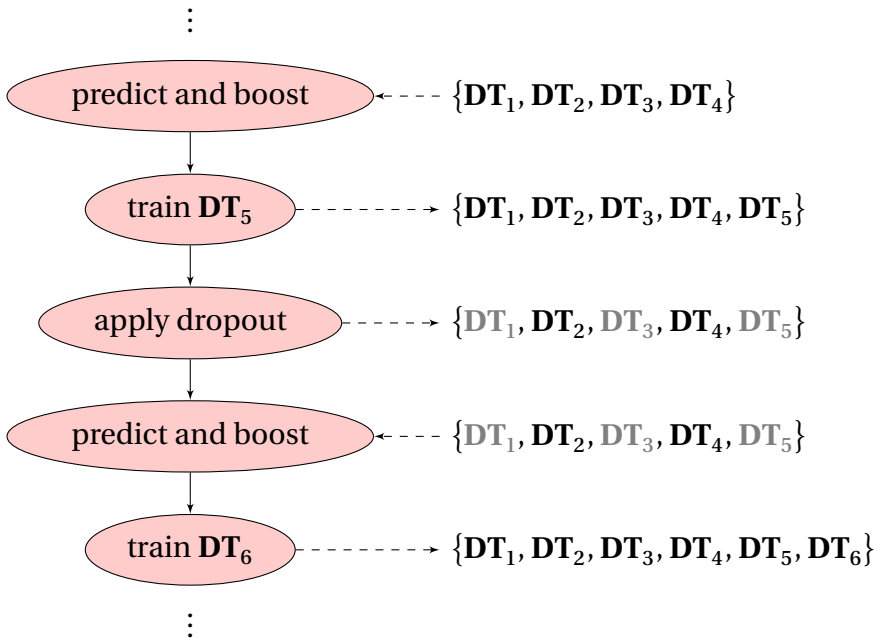


Figure 4.4: **Example flowchart of dropout regularization**
*An exemplary flowchart of dropout regularization is depicted. After four training cycles of a **BDT**, presented in Figure 4.3, four **DT**s have already been trained. $\boldsymbol{DT}_5$ is trained regularly. By chance, the next cycle has dropout regularization applied. $\boldsymbol{DT}_1$, $\boldsymbol{DT}_3$, and $\boldsymbol{DT}_5$ are disabled. Only $\boldsymbol{DT}_2$ and $\boldsymbol{DT}_4$ are used for the next prediction. $\boldsymbol{DT}_6$ is trained based on the boosted data based on this prediction. For the next cycle, all **DT**s are enabled unless dropout is applied once again.*

4.4.4 Tikhonov Regularization

Tikhonov regularization is a set of techniques to smooth linear ill-posed problems and make them potentially numerically solvable. A regularization parameter $\boldsymbol{w}$ is added to the solution and subsequently minimized. In the case of gradient based **BDT**s, described in Section 4.5, $\boldsymbol{w}$ is the bias of the **BDT** towards all data points in the data set. An example of gradient based optimization is presented in Figure 4.5.
For this work, a special kind of Tikhonov regularization is utilized by adding the *L2 regularization term on weights* $\|\boldsymbol{w}\|_2^2$, described in reference [Ng04], to the regularization function. Thus, any newly added **DT** which introduces large deviations from the previous predictions is suppressed in favor of a **DT** with small changes to the previous predictions, a prerequisite for convergence towards a minimum.

4.5 EXtreme Gradient Boosting

In this thesis, the **ML** is implemented in the eXtreme Gradient Boosting (**XGBoost**) framework, introduced in reference [CG16]. In this framework, boosting is implemented via *gradient boosting* with multiple different supported boosting algorithms. The *Dropouts meet Multiple Additive Regression Trees* (DART) booster, presented in reference [VG15], has been chosen as it allows for dropout regularization during boosting as discussed in Section 4.4.3.
In gradient boosting, the individual predictors are not directly trained using the class of a given data point but by *pseudo-residuals*. Let $\{p_1, \dots, p_{n-1}\}$ be the set of already trained **DT**s, $\hat{y}_i$ the class of the data point $\boldsymbol{x}_i$, and L a chosen loss function. Then, the pseudo-residuals are

$$R_{i,n} = -\frac{\partial L\left(\hat{y}_i, p_{n-1}\left(\boldsymbol{x}_i\right)\right)}{\partial p_{n-1}\left(\boldsymbol{x}_i\right)}. \tag{4.5}$$

The next **DT** p_n is trained to predict these pseudo-residuals, i. e., in the ideal case $p_n\left(\boldsymbol{x}_i\right) = R_{i,n}$. Before the new **DT** is added to the set of all already trained **DT**s, the predictions are scaled with an arbitrarily chosen scaling factor, the *learning rate* η. With η the rate of descent towards a minimum can be controlled. The higher η, the faster and more prone to non-convergence the ensemble of predictors becomes, the lower η, the more predictors are needed to descend towards a minimum.
During the training of the **DT**, sub-sampling to the training data can be applied at each **SVD**. Each **SVD** is calculated following the *Exact Greedy Algorithm* described in reference [CG16].

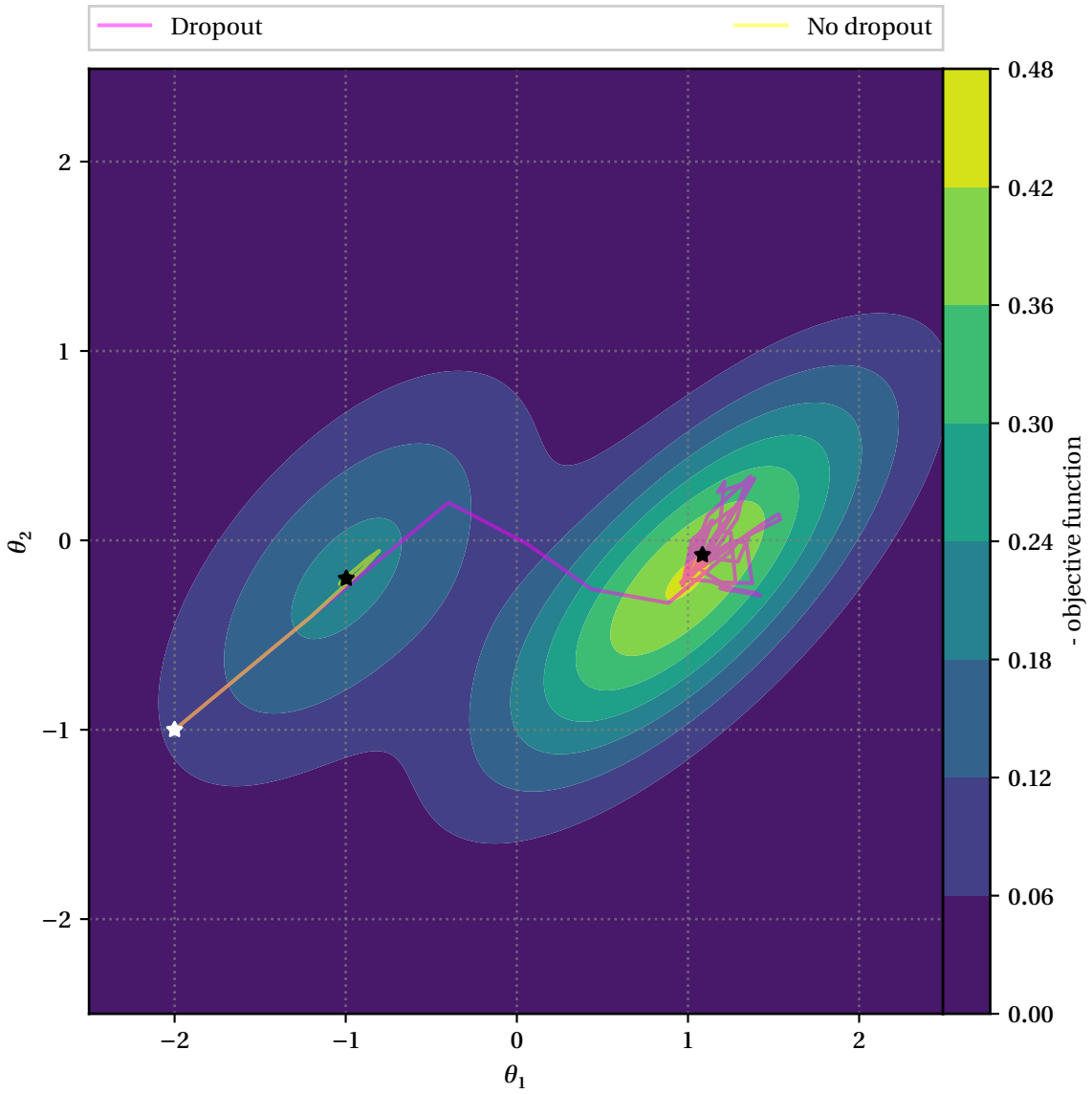


Figure 4.5: **Gradient descent example with and without dropout**
The two dimensional surface of an arbitrary chosen objective function as a function of two varying parameters $\theta_{1/2}$ is presented. The objective function is emulated by the sum of two Gaussian distributions, one forming a local minimum, the other the global minimum. An arbitrary starting point, indicated by a white star, is chosen. Following the gradient of fastest descend, 100 steps are performed with a decreasing step width for numeric stability. The endpoint is marked by a black star. In one case, only the current position and hence all previous steps are regarded, i. e., the no dropout scenario. It can be observed that the found minimum is the local minimum. With dropout, the path also directs into the local minimum at first but as previous steps are dropped out, an escape is made out of the local minimum to the global minimum. Note that this is an arbitrary constructed example, emulating the setup of gradient descent, to highlight the principal.

Chapter 5

Light Emission of Magnetic Monopoles in Ice

Wherever there is light, one can photograph.

A. Stieglitz [Ada80]

Magnetic monopoles (**MM**s) passing through ice cause light emissions. These emissions can be grouped into two subcategories. The first are classical light emission processes caused by the charge of the incident particle. The second are hypothesized light emissions which arise from the structure of the **MM** and might differ between various theories. The latter is mostly of interest in the non-relativistic regime where **MM**s are theorized to catalyze nucleon decay and hence is out of the targeted velocity range of this analysis.

Classical light emissions are modeled by the measured light emissions of already detected particles. The light yield is adjusted to the difference in emission from one electric charge to one magnetic charge. For this thesis, three classical light emissions are of interest: direct Cherenkov radiation, indirect Cherenkov radiation, and luminescence light emissions, described in the following sections.

5.1 Direct Cherenkov Radiation

Coherent light produced by the polarization of a dielectric medium by a propagating electromagnetic field with a velocity higher than the local phase speed in the medium is called Cherenkov radiation. This effect was first described in references [Čer34; Čer37]. Each volume element of the dielectric medium becomes polarized locally due to the electromagnetic field. Once the electromagnetic field does not force the material into a polarized state anymore, the material returns to a non-polarized state and in turn emits a spherical light wave. Due to the polarization being induced faster than the local phase speed, these light waves can positively interfere creating a plane wave front which can be detected as macroscopic light.
A common case of direct Cherenkov radiation occurring are electrically charged particles traversing a dielectric medium close to the speed of light. In this case, the positive interference occurs under a specific angle from the particle trajectory as seen in Figure 5.1.

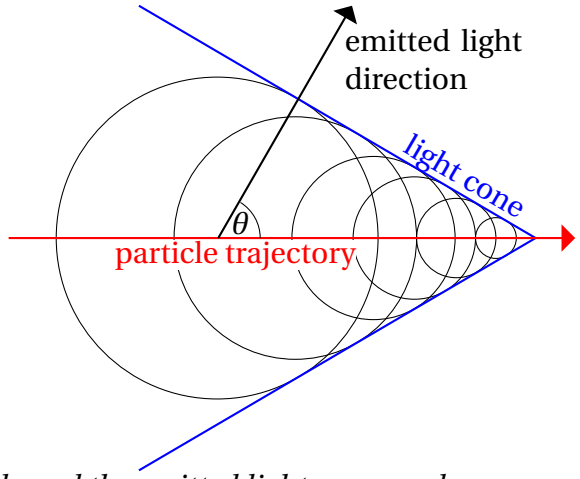


Figure 5.1: **Direct Cherenkov light**
A sketch of the induction of Cherenkov light is shown. A charged particle, drawn in red, traverses a medium. The medium is polarized locally. Once the material loses its polarization, a wave of light is emitted indicated by black circles. The waves of light from different locations on the trajectory positively interfere and form a light cone indicated by two blue lines. The angle θ between the trajectory of the particle and the emitted light cone can be observed.

The angle θ is defined by

$$\cos\theta = \frac{1}{n(\lambda)\,\beta} + \frac{p_{\text{pho}}}{2p_{\text{par}}}\left(1 - \frac{1}{n^2(\lambda)}\right) \tag{5.1}$$

with

n as the refractive index of the medium for a specific wavelength λ,

β as the ratio of the speed of the particle to the absolute speed of light,

p_{pho} as the momentum of the photon, and

p_{par} as the momentum of the particle, taken from reference [Gin40].

The term $\frac{p_{\text{pho}}}{2p_{\text{par}}}\left(1-\frac{1}{n^2(\lambda)}\right)$ can be ignored for most media as the momentum of the photon is negligible. The angle θ does not depend on the rest mass of the particle but only on its velocity.

The number of emitted photons in a defined wavelength range per distance traversed by the particle is

$$\frac{dN_{\gamma_{dCh}}}{dx} = \frac{q_e^2}{c^2}\int_{\lambda_1}^{\lambda_2}\left(1-\frac{1}{n^2(\lambda)\cdot\beta^2}\right)\frac{d\lambda}{\lambda^2} \tag{5.2}$$

where λ_1 to λ_2 is the wavelength range. It is also independent of the rest mass of the particle. It only depends on the velocity of the particle as described in reference [TF37].

To adapt the aforementioned formulas derived for electrically charged particles to magnetically charged particles, a transformation as described in Equation 3.2 can be applied. Choosing $R = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, i. e., switching the electric and magnetic components, and applying it to Equation 5.2 the number of Cherenkov photons per path length can be determined to be

$$\frac{dN_{\gamma_{\text{dCh}}}}{dx} = \frac{q_m^2}{c^2}\int_{\lambda_1}^{\lambda_2}\left(n^2(\lambda)-\frac{1}{\mu\cdot\beta^2}\right)\frac{d\lambda}{\lambda^2} \tag{5.3}$$

where μ is the magnetic permeability.

For ice, assuming $n(\lambda) = n \approx 1.3$ and $\mu \approx 0$, the ratio of Cherenkov photons emitted by an **MM** with one Dirac charge to a muon is $\approx \frac{q_m^2\cdot n^2}{q_e^2} = \frac{\left(\frac{q_e\cdot n}{2\alpha}\right)^2}{q_e^2} \approx 8400$. Another difference is a change in the polarization between the two. This is of no consequence for this thesis.

5.2 Indirect Cherenkov Radiation

Depending on the charge and velocity of an incident particle, the particle can ionize the target medium inducing free electron emission around its path. If enough energy is transferred to the electrons, they themselves become relativistic and induce Cherenkov light on their own. A sketch of the setup is depicted in Figure 5.2.

These relativistic electrons are referred to as δ-electrons. The Cherenkov radiation emitted by δ-electrons is called indirect Cherenkov light. In contrast to direct Cherenkov light, there is no sharp velocity at which this effect takes place. Instead, the light yield decreases with reducing velocity until no δ-electrons are created.

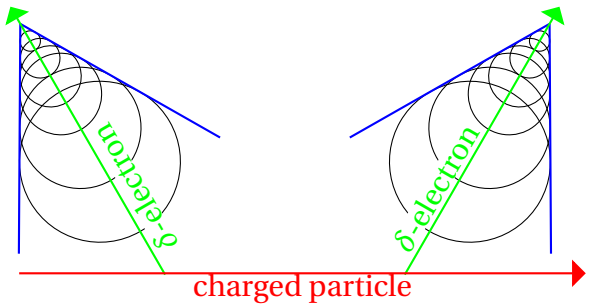


Figure 5.2: **Indirect Cherenkov light** *A sketch of the induction of indirect Cherenkov light is shown. A charged particle, here drawn in red traverses a medium. δ-electrons, drawn in green, are knocked out of the medium by the traversing particle. These in turn produce direct Cherenkov light as shown in Figure 5.1.*

The structure and yield of this light emission channel is defined by the emission characteristic of δ-electrons. The emission of δ-electrons caused by an **MM** can be modeled following the Kazama Yang Goldhaber (**KYG**) cross-section, described in reference [KYG77]. In this thesis, this is implemented the same way as in a previous search for **MM**s described in reference [Pol18]. A fitted form factor, depicted in Figure 5.3, is applied to the Rutherford differential cross-section described in reference [Rut11]

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{Rutherford}} = \left(\frac{q_{\text{target material}}\, q_m}{4E}\right)^2 \frac{1}{\sin^4 \frac{\vartheta}{2}} \tag{5.4}$$

where E is the initial energy of the incident particle and ϑ is the reflection angle.
By integrating over the number of photons produced per δ-electron and the δ-electron production rate, the total number of photons per unit length can be calculated via

$$\frac{dN_{\gamma_{\text{iCh}}}}{dx} = 2\pi \int_{E_0}^{E_m(\beta,M)} \left(\frac{d\sigma\left(E,\vartheta\left(E,\beta\right)\right)}{d\Omega}\right)_{\text{KYG}} \left(\int_{E_0}^{E} \frac{dN_{\gamma_{\text{dCh}}}\left(E'\right)}{dE'} \cdot dE'\right) \cdot dE. \tag{5.5}$$

E_0 is the minimal kinetic energy needed for an electron to produce direct Cherenkov light, $E_m\left(\beta, M\right)$ is the maximal energy transferred to a recoiled electron from an **MM** with mass M, and $\vartheta\left(E, \beta\right)$ is the scattering angle in dependence of the incidents particles velocity β and the transferred energy to the electron E.

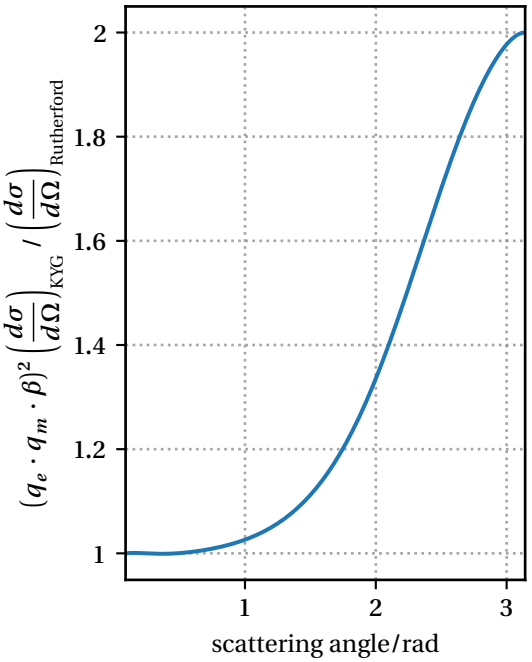


Figure 5.3: **KYG form factor**
*On the vertical axis, the ratio between the Rutherford cross-section and the **KYG** cross-section is shown as a function of the scattering angle on the horizontal axis. A scaling factor based on the magnetic and electric charge is included. The graph is interpolated from the tabulated values published in reference [KYG77], applying a six dimensional polynomial fit. For low scattering angles, the **KYG** and Rutherford cross-section are roughly the same, but the **KYG** cross-section increases by a factor of two approaching an scattering angle of π rad.*

5.3 Luminescence Light Emissions

Energy is conserved and thus, any energy deposited into a medium by an incident particle is either stored as potential energy or emitted in some form from the medium. There are two energy emission methods which directly emit light: black body radiation due to the heating of the material and luminescence light due to molecular or atomic excitation. For this work, black body radiation is neglected as no detectable shift in temperature of the medium can be expected. Luminescence light is the sum of the light emitted during individual molecular or atomic processes, each associated with an individual emission spectrum of light and decay time. An example of the resulting photon emission is illustrated in Figure 5.4. Consequently, it highly depends on the composition of the target material as described in reference [Bel53].

Ice emits luminescence light when illuminated with charged particles. Depending on the preparation process of the water and ice, different photo spectra can be observed as described in references [QTS82; Fre+84; Lee+09].

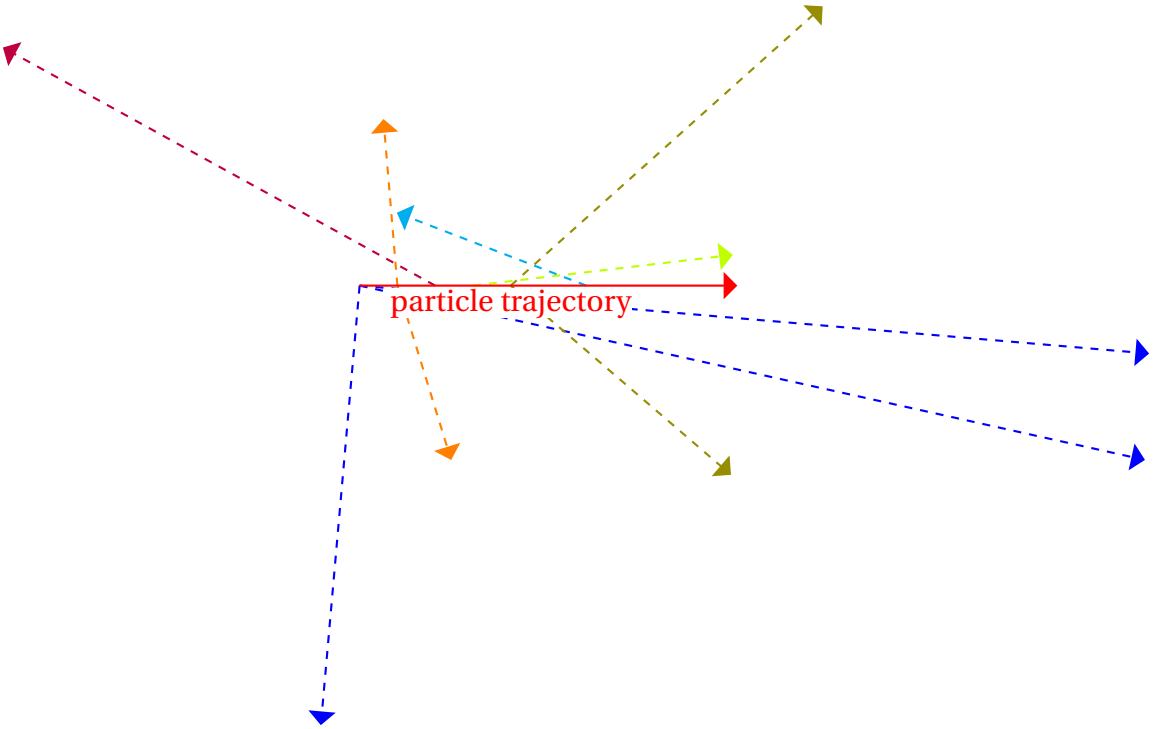


Figure 5.4: **Emission of luminescence**
A two dimensional sketch of the emission of luminescence light around a particle trajectory at a fixed point in time is shown. A particle traverses a medium at 0.5 c. Excitation of the medium is simulated in fixed steps. Individual photons are emitted indicated by arrows from the excitation points. The photons are emitted isotropically. An exponential time delay from the time of excitation to the emission of the photon is simulated. The location of the arrow indicates the covered distance. The color of the photons encodes the time and consequently the location of the excitation of the medium.

Temperature influences the spectra as described in reference [Qui+91]as does the fragmentation of the ice as described in reference [Bak+88]. Measured decay times range from 15 ns in reference [Fre+84] to 250 ns in reference [Bak+88]. As the IceCube Neutrino Observatory (**IceCube**) cannot distinguish different wavelengths of light, these photo spectra can be integrated into an average effective luminescence light yield $\frac{dN_\gamma}{dE}$ within the wavelength acceptance of **IceCube**. Hence, the luminescence light yield becomes

$$\frac{dN_{\gamma_{\text{lum}}}}{dx} = \frac{dN_\gamma}{dE} \cdot \frac{dE}{dx} \tag{5.6}$$

where $\frac{dE}{dx}$ is the energy loss per unit length of the incident magnetic monopole as described in Section 3.7. Notably, the average effective luminescence light yield varies depending on the preparation of the sample and excitation mode by roughly an order of magnitude from 0.2 γ/MeV in reference [Tro92] to 2.4 γ/MeV in reference [SSH72]. Precise knowledge of the particular ice is thus required.

Measurements of the effective light yield at **IceCube** have been performed and are described in reference [Pol19]. Based on this reference, an effective luminescence light yield for magnetic monopoles of $\frac{dN_\gamma}{dE} = 1\,\gamma/\text{MeV}$ is assumed. While four decay times have been preliminarily identified in reference [Pol19] $\left(2.4(2)\,\text{ns}, 190(30)\,\text{ns}, 5.03(6)\,\mu\text{s}, 56(6)\,\mu\text{s}\right)$, the relative ratio of the decay modes to the absolute amount of luminescence is not available. Assuming the first two decay modes dominate the luminescence production, an average decay time of 500 ns is assumed during simulation in this thesis to be conservative towards the first two decay modes.

5.4 Expected Light Yields

Depending on the assumed velocity, the light yield of **MM**s is dominated by different processes, described in Section 5.1, Section 5.2, and Section 5.3. A comparison of the expected light yields in ice is illustrated in Figure 5.5. While luminescence light has the lowest expected light yield, it is mostly constant over the whole velocity range while direct Cherenkov light has the highest expected light yield, but drops off sharply once the incident particles speed falls short of the local speed of light. Indirect Cherenkov light also shows a change with velocity, but does not have a sharp drop off like direct Cherenkov light. Instead, it decreases until it becomes undetectable.

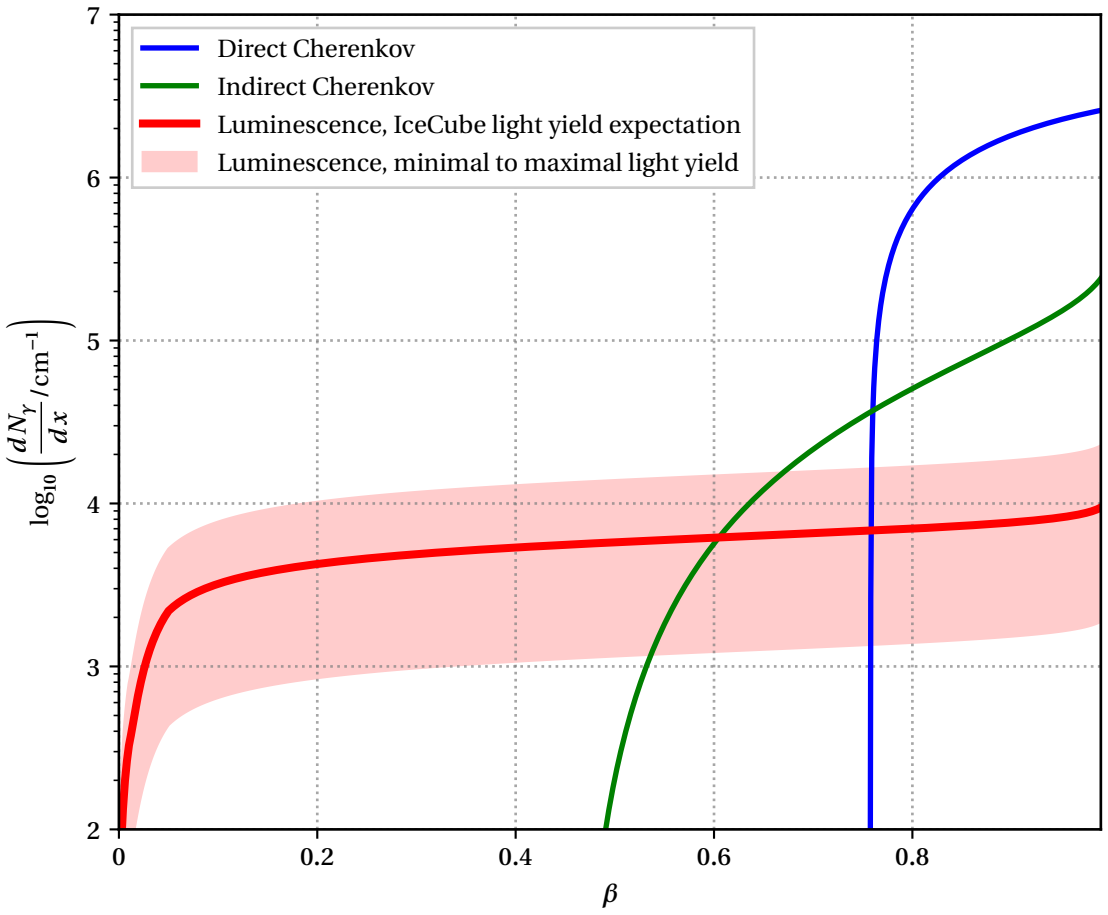


Figure 5.5: **Expected light yields for MMs in IceCube**
*The theoretically predicted light yield for **MM**s as a function of velocity within the ice of **IceCube** is shown. On the horizontal, linear axis, the velocity in relation to the speed of light between 0 c and 0.99 c is shown. On the vertical, logarithmic axis, the expected number of emitted photons per unit length is depicted. Three light production mechanisms are regarded, described in Section 5.1, Section 5.2, and Section 5.3. For luminescence, both the expected value for the in-ice arrays of **IceCube** and the minimal and maximal reported light yields from reference [Tro92] and reference [SSH72] is drawn.*

Chapter 6

Simulation

Physicists like to think that all you have to do is say, these are the conditions, now what happens next?

R. P. Feynman [Fey70]

Modern experiments have complicated response functions to physical events which are challenging to derive analytically. Instead, Monte Carlo Simulations (**MCS**s) are utilized to derive the response function. This is especially true for theoretical physical signatures like magnetic monopoles (**MM**s).

MCS is based on the law of large numbers. Many possible physical events are simulated and the expected detector response is calculated. Attributes of the physical event which are of no interest or are not observable to the experiment, i. e., random fluctuations in one of the experiments components, are averaged out. Assuming a large enough number of simulated events, the sum of the detector responses converges against the true detector response.

In this chapter, the **MCS** used within the IceCube Neutrino Observatory (**IceCube**), and the specific simulation conducted for the analysis of this thesis is summarized. The creation of the simulated datasets, the sets containing the detector response for specific sampled points, utilized in this thesis is chosen as the transition point to Chapter 7.

To reduce the complexity of the **MCS** for a given experiment, the detector response can be decomposed in discrete sub-responses. For the **MCS** computed for **IceCube**, this decomposition has three sub-responses: the light produced inside the detection medium by a physical signature, the propagation of aforementioned light to the Digital Optical Modules (**DOM**s) of **IceCube**, and the response of the **DOM** and data acquisition (**DAQ**) of **IceCube** to the light. Thereafter, simulated and physically measured events are treated the same.

The following subsections will highlight each step, starting at the **DAQ**, described in Section 2.2.1, applied to both simulated and physically measured events tracking back to the initial location of a photon induced by an incident particle. The photon production originating from physical events is described in Section 6.3.

6.1 Digital Optical Module Response

The **DOM** response translates the first electron emitted from the photocathode, also called the photoelectron, of the **DOM** to the digitized signal including the creation of soft local coincidences (**SLC**s) and hard local coincidences (**HLC**s) as described in Section 2.2. The **DOM** response is composed of two sub-responses: the **PMT** response and the Main Board response.

6.1.1 Photomultiplier Tube Response

The **PMT** type used within each **DOM** has been characterized in reference [Abb+10a]. Multiple physical effects which influence the response function have been determined by the **IceCube** collaboration (**ICC**) and are included in the **PMT** response simulation. A *pulse* consists of a measured amplitude and time at the end of the **PMT** response simulation. A sketch of a **PMT** and the simulated pulses is illustrated in Figure 6.1. *Normal pulse* refers to a pulse that is only subject to the first two simulated effects.

The first effect is a random time delay between the photoelectron and the cascade of electrons arriving at the anode of the **PMT** referred to as *time jitter*. A Gumbel distribution with a mean of ≈1.26 ns and a variance of ≈6.05 ns is utilized to simulate a random time offset for each photoelectron.

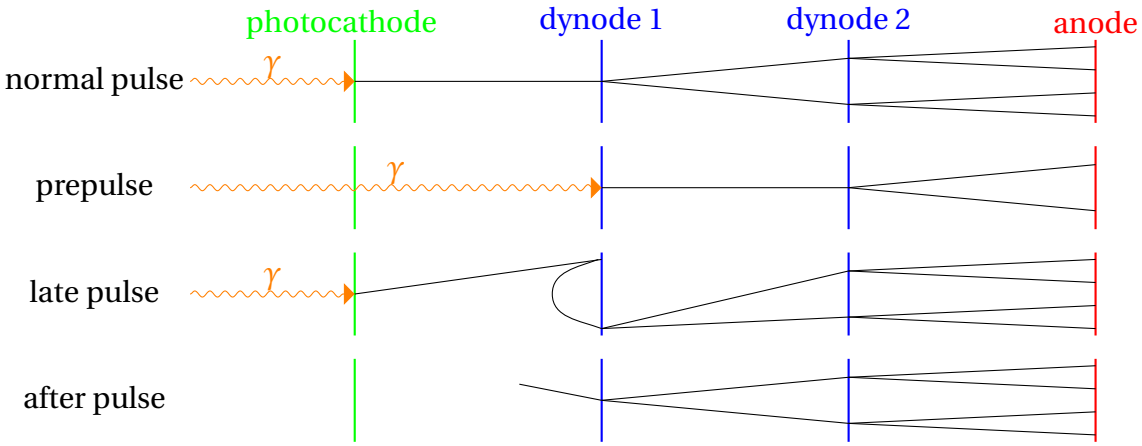


Figure 6.1: **Sketch of PMT with different pulses**
*A sketch of a **PMT** with two dynodes is illustrated. Four situations are shown each corresponding to one simulated situation described in Section 6.1.1. Electrons are drawn as black lines.*

The second effect is a non-constant amplification of a photoelectron. The *TA0003* model, published in reference [Aar+20], is used to map each photoelectron to a measured quantity in pe. The definition of pe can be found in Section 2.2.2. In this model, a probability density function is used to sample the derived charge for each photoelectron. The probability density function consists of a Gaussian distribution around 1 pe with an additional exponential distribution which dominates in the 0 pe to 0.5 pe range.

The third effect arises from photoelectrons originating from the first dynode of the **PMT** instead of the photocathode. These are called *prepulses*. The prepulses arrives earlier at the anode than normal pulses. The amplitude is also lower. This is simulated by applying a chance of 0.3 % to each photoelectron to be a prepulse. If it is a prepulse, a further time shift of –31.8 ns is applied and the signal amplitude is reduced by a factor of 12 in contrast to a normal pulse. Values can be found in reference [Aar+20].

The fourth effect results from a photoelectron back-scattering from the first dynode before initiating the amplification process. These pulses are referred to as *late pulses*. Each photoelectron is assigned a 3.5 % probability to be a late pulse, as described in reference [Aar+20]. These pulses have the same amplification as normal pulses but arrive later due to the extra distance traveled. The extra time delay for each late pulse is sampled from the combination of six Fisher-Tippett distributions derived in reference [Aar+20].

The fifth effect does not originate from a photoelectron but from ionized gas molecules inside the **PMT** which may be ionized by an amplification cascade caused by a photoelectron as described in reference [Ma+11]. These are called *after pulses*. About 5.9 % of all pulses are accompanied by an after pulse as mentioned in reference [Aar+20]. This is modeled by a probability density function derived from measured data. The probability density function consists of eleven Gaussian peaks combined with center time delays ranging from 500 ns to 7833.4 ns with varying widths as described in reference [Aar+20].

The last effect is caused by the electron current inside the **PMT** being limited by the **PMT** Base Board, described in Section 2.1.2. Once the current inside the **PMT** is greater than the supplied current, the amplification is reduced. This is called *saturation*. Saturation is taken into account by folding the simulated current with an exponential decay template with a decay time of 2.2 ns. The amplification is scaled accordingly. As this analysis is conducted in the low brightness regime, no saturation effects are expected.

6.1.2 Main Board Response

The simulation of the Main Board, described in Section 2.1.2, response consists of three sub-responses. The discriminator, the detection of **SLC**s and **HLC**s, and the digitization of the simulated **PMT** anode pulse are simulated.

Starting with the discriminator, the simulated **PMT** pulses, described in Section 6.1.1, are folded with calibration templates to take into account the electronic components between the **PMT** and the actual discriminating electronic element of the discriminator. Afterwards, every amplitude crossing over a preset discrimination value is collected and used in the next simulation step without any further processing. They are examined to detect coincident crossings thus making them **HLC**s, otherwise they are marked as **SLC**s. The same logic as described in Section 2.1.2 is used to build all hits.

For the simulation of the digitization, the waveform created for the discriminators is sampled in regular intervals. The sampling rate and dead times are taken into account. Electronic noise is added to each sample. The noise is modeled by a Gaussian distribution with a mean value of 0 and different widths for the different digitizers inside the **DOM**. The phase of sampling is randomized and a random phase noise is added to digitizers situated on different **DOM**s.

6.2 Photon Propagation Inside the Ice

IceCube utilizes the natural ice of Antarctica as a detection medium for particles. While "Antarctic ice is the cleanest, clearest natural solid on Earth" as mentioned in reference [Bra13], it is not completely isotropic, transparent, or scatter free. While the bulk of the original ice has these properties, parts of the ice had to be melted and refrozen during the instrumentation phase of **IceCube**. As the ice refroze from the outside to the inside of the boreholes, impurities were pushed towards the center altering the ice properties measurably. Due to this, this *hole ice* has different attributes than the original *bulk ice*.

To model the light transmission from an incident particle passing through the in-ice arrays of **IceCube**, specialized studies have been performed. Flashes of light were emitted in a controlled manner from **DOM**s which in turn were recorded by other **DOM**s. With this flasher data, South Pole ICE (**SPICE**) models were built starting with a Mie scattering based model called ***SPICE** Mie*. Over the past years, updates and adjustments to this model have been made. In simulation specifically conducted for this thesis, the **SPICE** 3.2.1 model is used described in reference [Ron19] which is the latest iteration of **SPICE** models. **SPICE** 3.2.1 is a complex, multi-parameter model describing the propagation and absorption of light inside the detection medium of **IceCube** depending on the initial direction and location of the light inside the in-ice arrays. This includes modeling impurities in the ice, most noticeably the dust layer, and anisotropies, due to the movement of the glacier.

While **SPICE** 3.2.1 is the current best model to describe the ice used by **IceCube**, it is not perfect. Sensitivity to systematic deviations of the model from reality are explored by systematic variations of the model parameters of the ice. The systematic variations related to the ice can be found in Section 7.3.9.

Within the simulation, photons originating from an incident particle are tracked through the detector, scattering as described by the aforementioned **SPICE** model, until they either are absorbed by the ice or a **DOM** where they induce a photoelectron.

6.3 Physical Signature to In-Ice Photons

Before any photon can be propagated through the ice, the time, direction, and initial location of the photon is required. While the individual photons do not interact with each other, the relative timing of photons and multi-photon detection at a given **DOM** need to be considered. This increases the size of the possible phase space immensely.

Within the active volume of **IceCube**, charged leptons are the only expected light inducing particles within the Standard Model of particle physics. While limiting the phase space to all possible signatures inducible by any number of charged leptons passing through the active detection volume of **IceCube** lowers the photon phase space tremendously, it is still virtually impossible to calculate the response function for the whole photon phase space.

Thus, the sampling needs to be further reduced down to physical signatures which can reasonably be expected to be measured. While there are no direct background events expected in this analysis as all other particles have a velocity which is at least twice as high as the signal particles, indirect background events from mis-reconstructed charged leptons are expected. In particular, coincident muons are of interest as these can be reconstructed as a single event passing through the detector at sub-luminal velocities.

For the in-ice arrays of **IceCube** in this thesis, the dominant injector of charged leptons are cosmic ray induced air showers described in Section 6.3.1. The simulation of light production by an **MM** is outlined in Section 6.3.3.

6.3.1 Cosmic Ray Induced Air Showers

The atmosphere of Earth is constantly subjected to a flux of high energetic cosmic particles historically called *cosmic rays*. Due to inelastic scattering between these primary cosmic particles with the nuclei of the atmosphere, secondary particles can be created. These in turn scatter elastically with the atmosphere losing energy until they are at rest or inelastically to produce further secondary particles. On top of this, they can decay into particles with lower rest mass.

The secondary particles can be grouped, based on their attributes, into the muonic component, the hadronic component, and the electromagnetic component. The electromagnetic and hadronic components are absorbed by either the atmosphere or the ice above the in-ice arrays of **IceCube** and are hence not detectable by the **IceCube** in-ice array (**IC**) or the **IceCube** DeepCore array (**DC**). The muonic component can reach the in-ice arrays. A single air shower can produce many muons which are causally linked inside the in-ice arrays of **IceCube**. A schematic view of an air shower with its components is depicted in Figure 6.2.

To simulate the event signature of cosmic ray induced air showers at the ground level, a software package originating from the KArlsruhe Shower Core and Array DEtector (**KASCADE**) experiment called COsmic Ray SImulations for **KASCADE** (**CORSIKA**), described in reference [Hec+98], is used. The primary and the produced secondaries are propagated until they interact with the atmosphere. The probabilities for all further secondaries are calculated and one specific interaction is selected. The secondaries produced in the selected interaction are then added to the particles being propagated through the atmosphere.

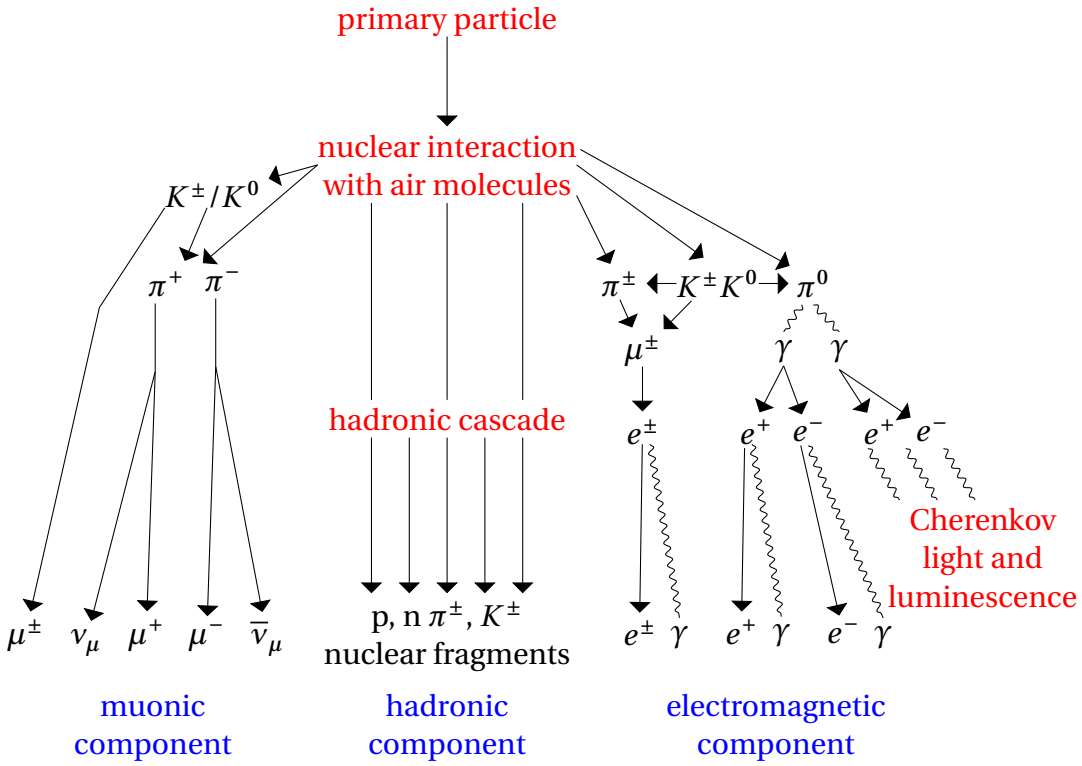


Figure 6.2: **Schematic air shower**
*A schematic cosmic ray induced air shower is shown with the different components being spatially separated. Only the muons and neutrinos are able to reach the in-ice arrays of **IceCube**. The separation of components is only used for illustrative purposes and does not happen in a real extended air shower. Graphic is inspired by reference [RS12].*

Particles at ground level are stored. If stopped inside the atmosphere due to lack of kinetic energy, they are removed until no particle is left to be tracked. Once no particle remain to be tracked, the simulated air shower can be used as an input to the **IceCube** simulation.

As simulating an air shower is computational expensive, individual showers can be used multiple times to generate a signature inside the in-ice arrays of **IceCube**. This process is called *over sampling*. The location of the primary and the relative rotation of the primary to the in-ice arrays of **IceCube** is randomized resulting in different signal signature for the same air shower. For each over sampled air shower, only the muons reaching the ground are propagated through the Antarctic ice to the in-ice arrays of **IceCube** as only these particles have a reasonable chance to be detected by the in-ice arrays. The propagation through the ice is done utilizing PROpagator with optimal Precision and Optimized Speed for All Leptons (**PROPOSAL**), described in reference [Koe+13]. Inside the in-ice arrays, the muons produce light via direct Cherenkov light, as described in Section 5.1.

While neutrinos are also able to reach the in-ice arrays, they do not directly produce light inside the ice. Instead, they have to interact with the ice to produce charged leptons which in turn produce light. As the probability of this process is rather low in contrast to the light production of the air shower muons, neutrinos are neglected in the **CORSIKA** simulation. Instead, specialized neutrino simulation described in Section 6.3.2 is conducted.

The flux of primary cosmic particles F in dependence of kinetic energy E can be described by a power law with a power law factor γ being a function of the energy and the type of primary

$$\frac{dF}{dE} \propto E^{-\gamma}. \tag{6.1}$$

In simulation, this is approximated by simulating five different kinds of primary particles: hydrogen nuclei, helium nuclei, nitrogen nuclei, aluminum nuclei, and iron nuclei. The ratio between the nuclei can be adjusted and a different γ factor for each type can be applied. The combination of used primary nuclei, energy dependent γ factor, and total rate of each primary nuclei is called the flux profile. After simulation, the events can be reweighted to follow any flux profile instead of the original profile they have been simulated with.

Three **CORSIKA** based datasets are used in this thesis: the cosmic ray induced air showers (2012) set, the cosmic ray induced air showers (2016) set, and the cosmic ray induced air showers (2016)c set. Each dataset consists of multiple subdatasets covering different energy ranges described in Tables 6.1 to 6.3. The subdatasets have been centrally simulated by the **ICC** and are designed to be used by multiple analyses. Thus, the parameters for each subdataset are not specifically chosen for this analysis. Each subdataset has an **ICC** internal, unique identification number or *ID* which can be used to identify the subdataset by the **ICC**. While of no use to the public, the IDs are given here to uniquely identify the sets used in this thesis.

As a brief summary, the cosmic ray induced air showers (2012) set consists of **CORSIKA** based cosmic air showers conducted at **IceCube** in 2012 but has been updated to use the 2016 filter settings of **IceCube**. Due to this, the dataset is well understood and used by many other analyses but has limited statistics. The cosmic ray induced air showers (2016) set and the cosmic ray induced air showers (2016)c set consist of new simulations conducted by the **ICC**. Both utilize software, filters, and settings present at **IceCube** in 2016. The simulation process was started in the mid of 2019 and therefore are not studied by many analyses yet. As they deploy a new technique, over sampling of simulated cosmic ray induced air showers, they have a much higher statistic.

While both cosmic ray induced air showers (2016) and cosmic ray induced air showers (2016)c are created with the 2016 filter settings of **IceCube**, they were created using two different versions of the SIBYLL event generator, described in reference [Fle+94], which determines the interactions within the cosmic air showers. The cosmic ray induced air showers (2016) set uses version 3.2 described in reference [Rie+16], the cosmic ray induced air showers (2016)c set uses version 3.2 c described in reference [Fed+19]. Latter is an updated version of the former. As this analysis should not be sensitive to the difference of these two interaction models, both datasets are used to enhance the statistics.

All **CORSIKA** based datasets are weighted to simulate the same primary particle flux and composition. For this work, the *GaisserH3a* flux model, described in reference [Gai12], has been chosen to be comparable to other **IceCube** based analyses.

ICC ID	N_{Shower}	$N_{\text{Over sampling}}$	$E_{\text{min}}^{\text{primary}}$ / GeV	$E_{\text{max}}^{\text{primary}}$ / GeV
11499	$5 \cdot 10^{11}$	/	600	10^5
11937	$2.4 \cdot 10^{10}$	/	10^5	10^{11}

Table 6.1: **Subdatasets of the cosmic ray induced air showers (2012) set**
*In the table, the subdatasets constituting the cosmic ray induced air showers (2012) set with the number of simulated cosmic air showers are shown. No over sampling was used, hence one shower is equivalent to one potential event inside the in-ice array of **IceCube** and the times each shower was used to create a potential event for the in-ice arrays of **IceCube** are shown. All subdatasets have been created with **CORSIKA**, version 74001pg, SIBYLL 2.1, and the **SPICE** model* lea *described in reference [WI14].*

ICC ID	N_{Shower}	$N_{\text{Over sampling}}$	$E_{\text{min}}^{\text{primary}}$ / GeV	$E_{\text{max}}^{\text{primary}}$ / GeV
20777	10^{11}	20	600	$3 \cdot 10^4$
20779	10^{11}	50	600	$3 \cdot 10^4$
20782	$2 \cdot 10^{11}$	50	600	$3 \cdot 10^4$
20787	$3 \cdot 10^8$	2	10^6	10^{10}
20783	$3 \cdot 10^8$	2	10^6	10^{10}
20780	$3 \cdot 10^{10}$	10	600	10^{11}
20781	$3 \cdot 10^{10}$	2	600	10^{11}

Table 6.2: **Subdatasets of the cosmic ray induced air showers (2016) set**
*In the table, the subdatasets constituting the cosmic ray induced air showers (2016) set with the number of simulated cosmic air showers and the times each shower was used to create a potential event for the in-ice arrays of **IceCube** are shown. All subdatasets have been created with **CORSIKA**, version 75600, SIBYLL 3.2, and the **SPICE** model 3.2.1.*

ICC ID	N_{Shower}	$N_{\text{Over sampling}}$	$E_{\text{min}}^{\text{primary}}$ / GeV	$E_{\text{max}}^{\text{primary}}$ / GeV
20788	10^{11}	50	600	$3 \cdot 10^4$
20891	$5 \cdot 10^{11}$	50	600	$3 \cdot 10^4$
20789	$3 \cdot 10^9$	50	$3 \cdot 10^4$	10^6
20848	$3 \cdot 10^8$	2	10^6	10^{10}

Table 6.3: **Subdatasets of the cosmic ray induced air showers (2016)c set**
*In the table, the subdatasets constituting the cosmic ray induced air showers (2016)c set with the number of simulated cosmic air showers and the times each shower was used to create a potential event for the in-ice arrays of **IceCube** are shown. All subdatasets have been created with **CORSIKA**, version 75600g, SIBYLL 3.2 c, and the **SPICE** model 3.2.1.*

6.3.2 Neutrinos

Neutrinos are electrically neutral particles. There are three flavors of neutrinos named after their charged lepton partner, the electron neutrino ν_e, the muon neutrino ν_μ, and tauon neutrino ν_τ. The sum of their rest masses can be constrained to be below 0.2 eV as reported in references [SSM06; CS06]. As they are neutrally charged, they do not induce Cherenkov light. Instead, only light produced by secondaries from a neutrino-ice reaction can be detected. The cross-section of these processes, while energy dependent, is $\mathcal{O}\left(10^{-7}\text{ b to }10^{-34}\text{ b}\right)$ as described in reference [FZ12].

If a charged lepton is created via a neutrino-ice interaction and has enough kinetic energy, Cherenkov radiation can be induced as described in Section 5.1. Alternatively, the lepton can induce a cascade of secondary particles inside the ice which in turn induce light. Former is the dominant process for muons, latter the dominant process for electrons and tauons.

To simulate these interactions, a neutrino interaction is assumed and the resulting particles are propagated through the ice. Afterwards, the event is assigned a weight based on an assumed flux of neutrinos with the specific kinetic attributes.

Different sources for neutrinos exist exhibiting different flux characteristics. For **IceCube**, two fluxes are of interest: atmospheric neutrinos originating from cosmic ray induced air showers described in Section 6.3.1 and the extra terrestrial flux caused by cosmic neutrino generators like the Sun or Supernovae. The same simulated neutrino event can be used for different neutrino sources by applying different weighting based on the assumed flux attributes.

In this thesis, only atmospheric electron and muon neutrinos are regarded as background sources. The number of events originating from atmospheric electron and muon neutrinos passing the very first selection step is orders of magnitude lower than the number of background events originating from cosmic air shower muons. This also does not change over the course of the analysis. As an extraterrestrial neutrino flux is simulated by reweighting the same events to a different, orders of magnitude lower flux, their contribution can be neglected.

Two separate datasets have been used in this thesis to simulate the flux of atmospheric electron and muon neutrinos: the atmospheric ν_μ set and the atmospheric ν_e set. In Table 6.4 and Table 6.5, the subdatasets constituting the datasets are listed including energy ranges and statistics at generation level.

The current **SPICE** model at the time of simulation production has been used, namely **SPICE** 3.2. Afterwards, the datasets have been weighted to an atmospheric flux by assuming a *HKKM2006* flux as described in reference [Hon+07], again to be consistent with previous **IceCube** analyses.

ICC ID	N_{ν_e}	E_{min} / GeV	E_{max} / GeV
20321	$1.2 \cdot 10^9$	0.1	5
20364	$0.6 \cdot 10^9$	5	10^4
20407	$0.1 \cdot 10^9$	10^4	10^5

Table 6.4: **Subdatasets of the atmospheric ν_e set**
*In the table, the subdatasets constituting the atmospheric ν_e set with the number of simulated neutrinos to create a potential event for the in-ice arrays of **IceCube** are shown. All subdatasets utilize the **SPICE** model 3.2.*

ICC ID	N_{ν_μ}	E_{min} / GeV	E_{max} / GeV
20450	$1.2 \cdot 10^9$	0.1	5
20493	$0.6 \cdot 10^9$	5	10^4
20536	$0.1 \cdot 10^9$	10^4	10^5

Table 6.5: **Subdatasets of the atmospheric ν_μ set**
*In the table, the subdatasets constituting the atmospheric ν_e set with the number of simulated neutrinos to create a potential event for the in-ice arrays of **IceCube** are shown. All subdatasets utilize the **SPICE** model 3.2.*

6.3.3 Magnetic Monopoles

An **MM** signal was simulated by isotropically illuminating the in-ice arrays of **IceCube** with **MM**s with different kinematic and physical configurations. A disc with a radius of 850 m, its central axis pointing to the center of the **ICC** internal coordinate system, roughly in the middle of the **IC**, was placed 1000 m from the aforementioned center. While distance and size of the disc can be varied, the optimal values for an isotropic illumination described in reference [Chr11] have been used.
For each **MM**, the placement of the creation disk, expressed as two angles, the zenith θ and azimuth φ, is randomized. Afterwards, the **MM** is emitted parallel to the central axis from a random position on the disc. The initial velocity is sampled from a configurable distribution. The resulting **MM** flux is isotropic around the in-ice arrays of **IceCube**. A sketch of the setup is shown in Figure 6.3.

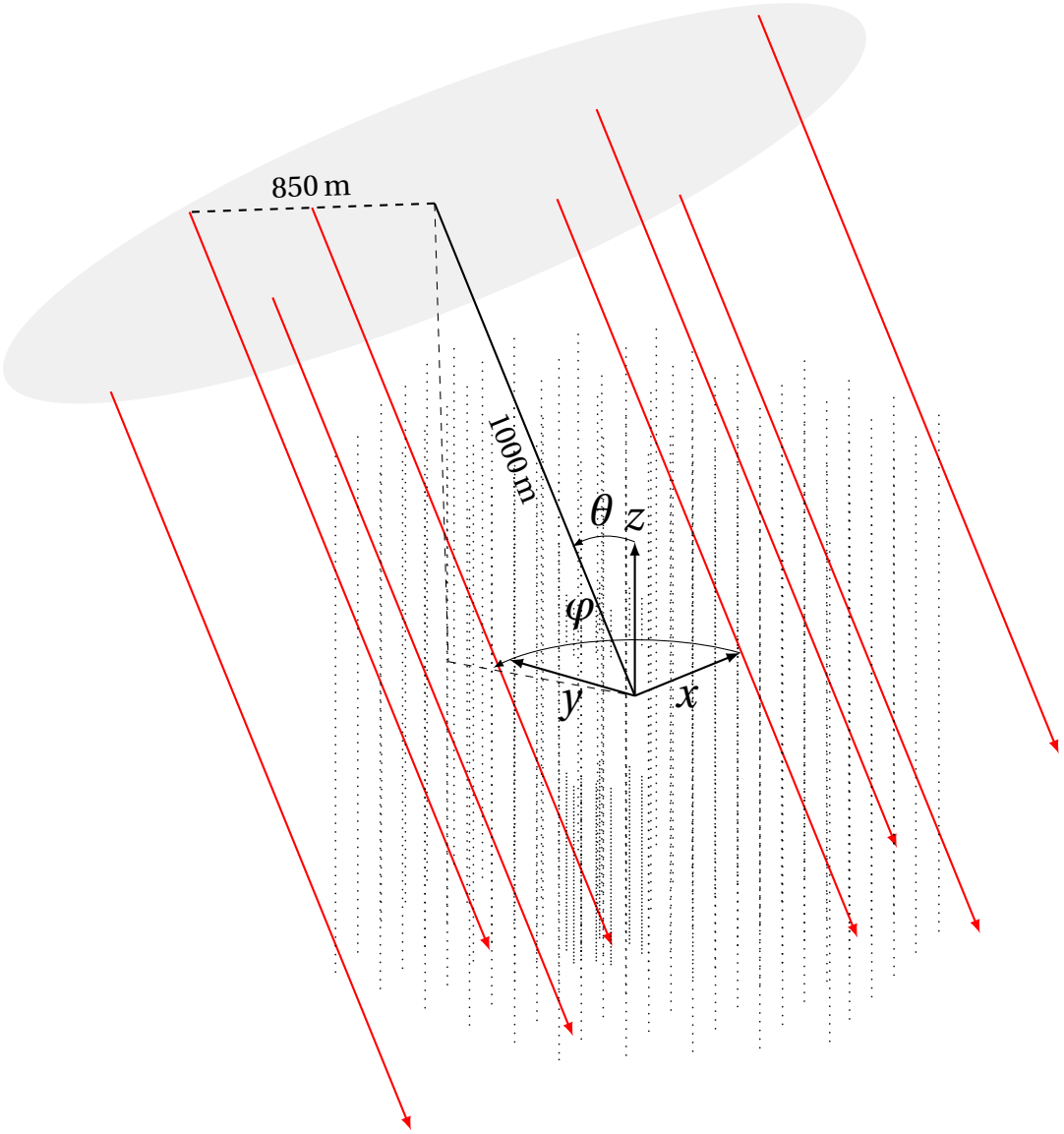


Figure 6.3: **Sketch of MM generation disc**
*A to scale sketch of the generation disc in gray relative to the in-ice arrays of **IceCube** is depicted. In red, potential **MM** propagation paths are drawn. The generation disc is placed at an example zenith θ and azimuth φ which are randomized for each generated **MM**. In the center, at the origin of the **IceCube** coordination system, the three directions of the coordinate system, the zenith, and azimuth of the generation disc are displayed.*

A mass of 10^{11} GeV was selected as the rest mass of all simulated **MM**s as this is the lower mass limit for intermediate mass magnetic monopole (**IMM**). Thus, any effect on the velocity of an **IMM** would be maximal. No measurable change in velocity within the volume of the in-ice arrays of **IceCube** is observed during the simulation of **MM** inside the volume of the in-ice arrays of **IceCube**. Hence, an **MM** can be assumed to remain at its initial velocity within the detection volume of **IceCube**.

While the **MM** is propagated through the ice, the energy loss due to interactions with the ice are calculated and its total energy is adjusted. Light production due to Cherenkov light is calculated as described in Section 5.1 and Section 5.2 and was already implemented by previous analyses.
The production of luminescence light was implemented and validated in the context of this thesis. The number of emitted photons per distance is calculated as described in Section 5.3. Each photon is emitted from the path of the **MM** in a random direction due to the isotropic nature of luminescence light. The initial position of the photons on the path is also randomly chosen. To simulate the delayed light emission, the emission time of each photon is shifted by a random time delay sampled from an exponential distribution.
Support for multiple pairs of luminescence light yield and decay time has been implemented albeit only simulation with one pair has been used in this thesis. The implementation was validated by comparing the expected light from simulation with the expected light from theory. Specialized simulations with fixed velocities were created. The combined light arriving at all **DOM**s of the in-ice arrays of **IceCube** and the light recorded by **IceCube** was compared to the theoretical light production per distance. The results is depicted in Figure 6.4. Events with a non-physically high light yield and long decay time were simulated and inspected by eye to observe the delayed light emission.

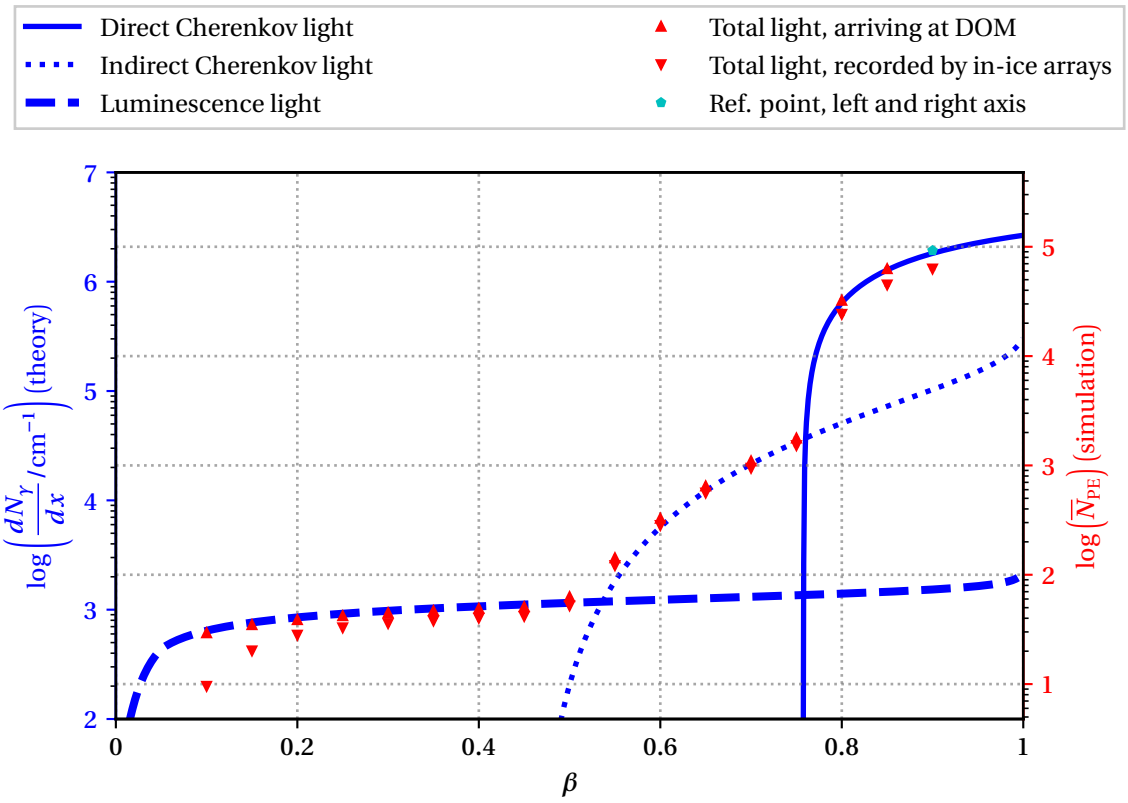


Figure 6.4: **Simulated and theoretical light production of MMs**
*An overlay of the expected arriving light integrated over all **DOM**s of the in-ice arrays of **IceCube**, the registered light by **IceCube** after digitization and triggering, and the expected number of photons per distance from theory as a function of initial velocity of an **MM** is depicted. The left, blue, vertical, logarithmic axis is the reference axis for all lines representing different light production mechanism: direct Cherenkov light (solid, Section 5.1), indirect Cherenkov light (dotted, Section 5.2), and luminescence light (dashed, Section 5.3). The right, red, vertical, logarithmic axis is the reference axis for all triangles indicating the mean of the total integrated light per simulated event. Two different metrics are used to determine the light per event. The upwards pointing triangles are based on light which reached a **DOM**. The downwards pointing triangles are based on light recorded by **IceCube** after the trigger process described in Section 2.2. Former is an indicator of the light induced by an event, latter is an indicator of the recorded light available for further analysis. An artificial scaling point was chosen indicated by a cyan colored point to align the two scales. It can be observed that the upwards facing, red triangles closely follow the expected light production from theory indicating a correct light yield simulation. While the downward facing triangles match mostly to the upwards facing ones, they divert in the high velocity, high light region at about 0.9 c due to saturation effects in the **DOM**s as described in Section 6.1.1 and the low velocity, low light region below 0.25 c due to a reduced chance to trigger on the simulated event. A worst case luminescence light yield was chosen (0.2 γ/MeV) for this test making the difference between the upward and downward facing triangles also the worst case scenario for this analysis.*

Multiple datasets to simulate **MM**s have been created, each covering a certain velocity range. A flat velocity spectrum was always chosen as there is no information to indicate a strong dependency of the expected **MM** flux to the expected velocity around the in-ice arrays of **IceCube** nor is there a strong dependency of computational cost per event to the aforementioned velocity.
Datasets have been created with two different luminescence light yields: $0.2\,\gamma$/MeV and $1\,\gamma$/MeV. The former is the lowest measured luminescence light yield reported by lab measurements of water-based ice and can be treated as the worst case scenario for light production. Latter is the expected luminescence light yield from measurements conducted in-situ at **IceCube** described in reference [Pol19]. For both, a velocity range starting at $0.1\,c$ to $0.6\,c$ has been chosen, split into multiple section covering $0.05\,c$ each. $2 \cdot 10^6$ **MM** events per $0.05\,c$ have been simulated.
While weighting for **MM**s can be chosen arbitrarily, each dataset is weighted to an integrated flux at generation time of $1.68 \times 10^{-16}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-2}$, the best flux limit in this velocity range before this analysis described in reference [Amb+02]. With the chosen weighting, each dataset can be interpreted as an **MM** flux within the given velocity range consistent with the previous best flux limit. Thus, in any graph comparing natural rates of simulated background events and simulated signal events in this work, the minimal required further separation to detect **MM**s can be observed. This makes the difference in natural rate between simulated background and signal in a given distribution the minimal required further separation needed before the signal can become dominant over background.

Chapter 7

Event Selection and Reconstructions

He who never made a mistake, never made a discovery.

S. Smiles [Smi66]

In this chapter, the methods and event selection steps required to process the lowest available data stream of the IceCube Neutrino Observatory (**IceCube**) to the final event selection of this analysis are described. In Section 7.1, all processing and signal selection steps which are implemented invariant from this specific analysis at **IceCube** are summarized, in Section 7.2, all selection criteria up to the last selection step of this analysis are presented, and in Section 7.3 the final selection step of the analysis, including the boosted decission trees (**BDT**) training, variable selection, model rejection factor (**MRF**) calculations, and optimization, is discussed.

The selection criteria are designed based on simulation of the detector response described in Chapter 6. The sum of the simulated cosmic ray induced air showers (2012) set, the atmospheric ν_μ set, and the atmospheric ν_e set is the simulated background set. To validate that the simulated background and a specific variable is representative of the measured data, the simulation is compared to a subset of the measured data at each selection step. This subset is referred to as the *blinded* data set due to the drastically reduced statistics. It consists of every 10^{th} data run recorded by **IceCube** and thus is approximately 10 % of the total data.

7.1 IceCube Event Selection

In this section, the event selection starting at the combined data streams of all Digital Optical Modules (**DOM**s) in the **IceCube** laboratory (**ICL**) to the **IceCube** internally named *level 0*, *level 1*, and *level 2* event selections, is outlined. This is intended for readers not familiar with the data processing used at **IceCube** and can be skipped otherwise as it does not contain any information specific to this analysis. As a takeover point, the online Processing and Filtering (**PnF**) system of **IceCube** is chosen, all processing before that is described in Section 2.2.

7.1.1 Level 0: Events at Data Acquisition

At the beginning of the **PnF**, events are the collection of all deconvoluted, calibrated hard local coincidence (**HLC**) and soft local coincidence (**SLC**) hits within a certain time window. This is called level 0. The size and position of the time window is defined by the global triggers as described in Section 2.2.1. Each hit consists of the information which **DOM** registered the hit, the calibrated amplitude in units of pe, the time of arrival, and if a local coincidence with other hits was detected. A pe is the most probable number of photons which caused the amplitude measured by a given **DOM**. As calibration is applied individually to all **DOM**s, the resulting calibrated amplitudes are comparable for all **DOM**s.

7.1.2 Level 1: Filtering at the South Pole and Reconstruction of Events at the IceCube Data Center

After the data has been processed up to level 0, specialized, **IceCube** wide implemented selection steps called *filters* are applied to select events of immediate interest to the **IceCube** collaboration (**ICC**). If any filter selects an event, it is transmitted via satellite to the **ICC** data center outside Antarctica. Before being transmitted, the event is compressed with a possible lossy algorithm and any redundant information is removed to reduce the required satellite bandwidth. This is called level 1.

If an event is not immediately sent, it will still be transmitted to the data center once a year via airplane. This allows to apply changes retroactively to the level 1 processing.

For this analysis, only one filter is utilized selecting relevant events, the MonopoleFilter_16 (**MPF16**) filter. In the following sections, the selection criteria of this filter are described.

7.1.2.1 The MonopoleFilter_16 filter

The MonopoleFilter_16 (**MPF16**) filter is a specialized filter implemented at **IceCube** in 2016. It is designed to detect particles with a velocity between the low relativistic up to the mildly relativistic regime (0.1 c to 0.75 c) and a wide range of possible light production yields matching the signature of an intermediate mass magnetic monopole (**IMM**). It is applied to events selected by the **SMT**-8, the **SMT**-3, Volume, and String triggers, described in Section 2.2.1.
Events are selected from both the **IceCube** in-ice array (**IC**) and the **IceCube** DeepCore array (**DC**) with different requirements due to the different geometries of the in-ice arrays. As a convention, variables based only on hits detected by one array will be denoted with the name of the array in question as $_{\mathbf{IC/DC}}$.
As a first step, each event is cleaned from probable noise hits. Starting by selecting all **HLC** hits, each **SLC** hit is compared against the already selected list of hits. If a hit is within a sphere of radius 150 m and a temporal distance of 1000 ns, it is selected as well. This process is repeated until no further hits are selected.
Afterwards, the selected hits are split into **IC** and **DC** selections depending on which in-ice array they are detected by. For the **DC** hit selection, the total light seen by each **DOM** is calculated and the dimmer half of hit **DOM**s is removed from the selection. As a convention, $h_{\mathbf{IC}}$ is the set of all hits of the **IC** selection, $h_{\mathbf{DC}}$ is the set of all hits of the **DC** selection.
For both selections, different criteria are applied, described in the subsequent sections. If one of these two criteria is met, the event qualifies for selection based on the **MPF16** filter.

7.1.2.2 In-Ice Array Selection

All variables are calculated on the aforementioned **IC** hit selection. A subscript to the variables using the **IC** is omitted in this section. To pass the **IC** selection of the **MPF16** filter, events need to have at least 6 **DOM**s hit and a maximal temporal distance between first hits (**TBH**) at any **DOM**s of at least 4000 ns. Assuming t_i is the time of the first registered hit at the i^{th} **DOM**, then **TBH** is defined as

$$\mathbf{TBH_{IC/DC}} = \max_{i,j \in \mathbf{DOMs_{IC/DC}}} \left(t_i - t_j \right). \tag{7.1}$$

Eventually, a global **MPF16** track hypotheses (**GMF16TH**) is calculated by applying a improved LineFit (**iLF**) to the event as described in Appendix A to all hits within the **IC**. The **GMF16TH** needs to converge and to have a reconstructed velocity $\beta_{\mathbf{GMF16TH}}$ of below 0.8 c.

The following last two requirements are calculated relative to the **GMF16TH**. Only hits that are within a 100 m radius of the **GMF16TH** are taken into account. These hits are projected onto the **GMF16TH** where l_i denotes the position along the track hypothesis, w_i is the number of reconstructed photons for the specific hit in pe. As a convention, the hits are sorted so that $l_{i+1} \geq l_i$.
The next to last requirement is based on the first to last quartile **COG** distance (**FLQCD**) defined as the distance of the center of gravity (**COG**) of the time sorted first quartile $Q1$ and last quartile $Q4$ of hits projected onto the **GMF16TH** or

$$\mathbf{FLQCD_{IC/DC}} = \frac{\sum\limits_{i \in Q4} l_i \cdot w_i}{\sum\limits_{i \in Q4} w_i} - \frac{\sum\limits_{i \in Q1} l_i \cdot w_i}{\sum\limits_{i \in Q1} w_i}. \tag{7.2}$$

A sketch of the calculation is depicted in Figure 7.1. To be selected, the absolute value of **FLQCD** needs to be at least 250 m. **FLQCD** is an indicator of the length of the **GMF16TH** and indirectly signals the amount of active media being passed by an incident magnetic monopole (**MM**).
The last requirement is a maximal separation length of hits (**MSL**) of at most 200 m. This describes the maximal distance between two track-wise neighboring hits or

$$\mathbf{MSL_{IC/DC}} = \max_{i \in (0,\ldots,N-1)} \left(l_{i+1} - l_i \right). \tag{7.3}$$

Thus, **MSL** is an indicator of the biggest gap in the light production along the **GMF16TH**. An **MM** should have a small gap as the light production is smooth while background events, especially cascade-like events will show a greater separation.

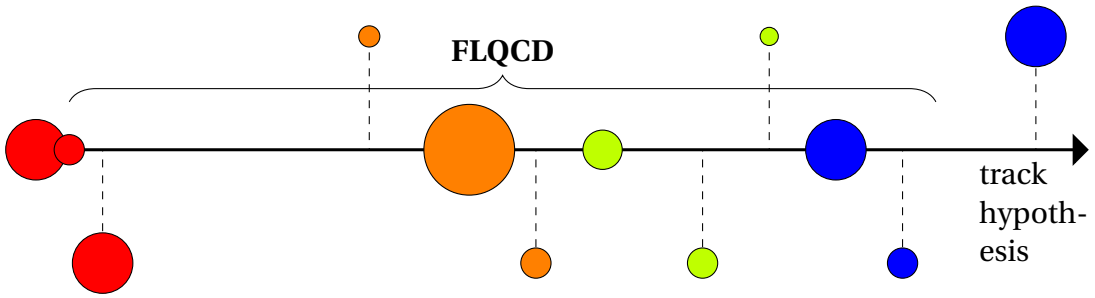


Figure 7.1: **Sketch of the first to last quartile COG distance**
*A sketch of the calculation of the first to last quartile **COG** distance (**FLQCD**) is drawn. An event consisting of twelve hits, indicated by spheres is shown. The radius of the sphere indicates the reconstructed amount of light. Based on these hits, a track hypothesis is drawn as a black arrow. Each quartile of hits is indicated by a different color: red for the first quartile, orange for the second quartile, lime for the third quartile, and blue for the fourth quartile. A bracket is utilized to indicate the distance between the **COG** of the first and last quartile of hits: the **FLQCD**.*

7.1.2.3 DeepCore Array Selection

All variables are calculated on the aforementioned **DC** hit selection. A subscript to this array is omitted in this section. Similar to the description in Section 7.1.2.2, the **DC** selection of the **MPF16** filter requires that at least 6 **DOM**s are hit and **TBH** needs to be bigger than 2750 ns. **TBH** has been introduced in Section 7.1.2.2.
The full width half maximum (**FWHM**) of the timing of the hits needs to be larger than 2500 ns. This is calculated assuming a Gaussian distribution of the hits via

$$\mathbf{FWHM_{IC/DC}} = 2\sqrt{2\ln 2}\sqrt{\sum_{i\in h_{\text{IC/DC}}}\left(t_i - \left(\frac{\sum_{j\in h_{\text{IC/DC}}} t_j}{\sum_{k\in h_{\text{IC/DC}}} 1}\right)\right)^2} \tag{7.4}$$

where t_i denotes the time at which a specific hit was detected. Note that no Bessel correction is included. As a last step, an **iLF** is applied to all hits within the **DC** as described in Appendix A. The **iLF** needs to converge and to have a reconstructed velocity below 0.7 c.

7.1.3 Level 2: MonopoleFilter_16 Filter Selection

At level 2, the main goal is to decompress the transmitted events from level 1 and reconstruct the deleted redundant information. For the **MPF16** filter, this is achieved by applying the same calculations as described in Section 7.1.2. Any variable utilized within the **MPF16** filter is accessible afterwards albeit with potential slightly different values as at level 1 due to the possible lossy compression of the event. In the context of this thesis, only the reconstructed values of level 2 are utilized.

7.2 Analysis Specific Quality Selection Steps

In this section, the analysis specific event selection steps are discussed up to the last step. These are implemented to increase the match between the simulated and measured data and to reduce the event rate. Latter allows more sophisticated, processing time intensive event selection algorithms to be used in the final selection step. While these selection criteria are commutative, they are presented in the order they have been developed in. Following nomenclature used in the **ICC**, each dataset after a given selection criteria is referred to as a *level*. The presented selection steps reduce the amount of background events by roughly a factor of 50 while keeping between 40 % to 80 %, depending on the specific velocity, of the signal selected. The exact numbers are presented in Table 7.1.

	Level 3 events		Level 8 events	
	Rate / Hz	$\mathcal{O}$ (#)	Rate / Hz	$\mathcal{O}$ (#)
Blinded data, 2011-2017	30.465(1)	$8 \cdot 10^8$	0.6021(2)	$2 \cdot 10^7$
Induced air showers (2012)	33.99(2)	$4 \cdot 10^6$	0.506(2)	$6 \cdot 10^4$
Atmospheric ν_e	4.44(1) µ	$1 \cdot 10^6$	0.207(2) µ	$3 \cdot 10^5$
Atmospheric ν_μ	10.83(1) m	$3 \cdot 10^6$	3.06(2) µ	$2 \cdot 10^5$
Monopoles (0.10 − 0.15 c)	7.26(1) µ	$3 \cdot 10^5$	2.662(8) µ	$1 \cdot 10^5$
Monopoles (0.35 − 0.40 c)	11.18(2) µ	$5 \cdot 10^5$	8.80(2) µ	$4 \cdot 10^5$
Monopoles (0.55 − 0.60 c)	12.88(2) µ	$5 \cdot 10^5$	5.46(1) µ	$2 \cdot 10^5$

Table 7.1: **Remaining events at different selection levels**
A table with the event rate and absolute number of events at level 3, the first analysis specific selection step described in Section 7.2.1, and level 8, the next to last selection step of this analysis described in Section 7.2.4, is presented.

7.2.1 Level 3: MonopoleFilter_16 filter, In-Ice Array Selection

Any event fulfilling the **IC** criteria of the **MPF16** filter is selected. This removes many of the complexities that arise from dealing with virtually two distinct detectors with different signal signatures. For this analysis, the focus is placed on the **IC** as it has a higher effective volume and consequently a higher flux of potential signal events.

7.2.2 Level 4-6: Coincident Particle Rejection

As most of the background events originate from coincident particles produced in cosmic ray induced air showers, multiple selection steps are created to reject these events. Each event is split time-wise in the middle. Two **iLF** are performed, one on each half. In case of an **MM** passing through the detector, these two **iLF**s should reconstruct similar kinematic attributes while multiple, coincident particles should be reconstructed with varying kinematic attributes. A sketch has been drawn in Figure 7.2 to illustrate the setup.

7.2.2.1 Level 4: Minimal Number of Digital Optical Modules hit

While the **MPF16** filter already has some requirements on the number of **DOM**s to be hit, as a prerequisite for the next level, a stronger requirement is added. As the event will be split, only half the number of hit **DOM**s, $N_{\text{hit DOMs}}$, is available for each **iLF**. To compensate for this, only events are considered with $N_{\text{hit DOMs}} \geq 25$. The distribution of the variable and the selected region is illustrated in Figure 7.3.

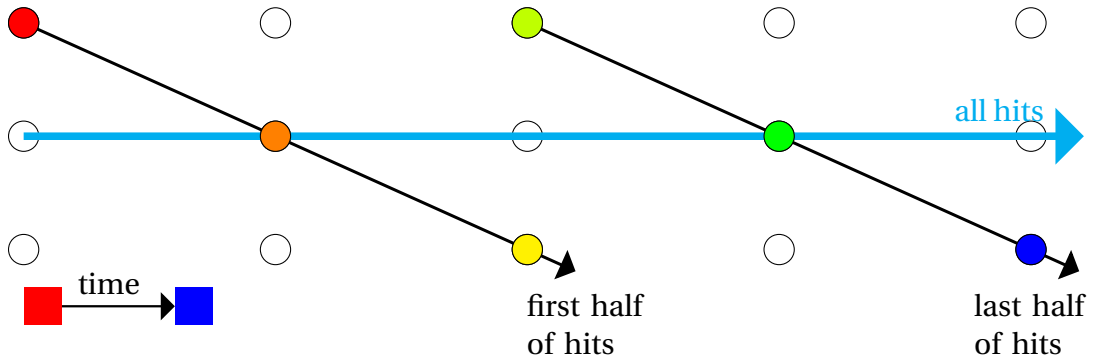


Figure 7.2: **Rationale for double splitting**
*A schematic of the logic behind event time-wise double splitting is presented. A two dimensional grid of circles, representing individual **DOM**s is drawn. The circles are colored to indicate the time when they are hit with white representing no hit at all. As an exemplary scenario, two particles traversing the grid in parallel but time-wise separate is chosen. Three hypothetical **iLF**s are drawn, the first is based on all hits with the resulting reconstruction indicated by a cyan arrow, the two other are computed on the first half and last half of hits individually with the resulting reconstructions indicated by black arrows. The time-split reconstruction fits the initial chosen scenario much better than the global **iLF**.*

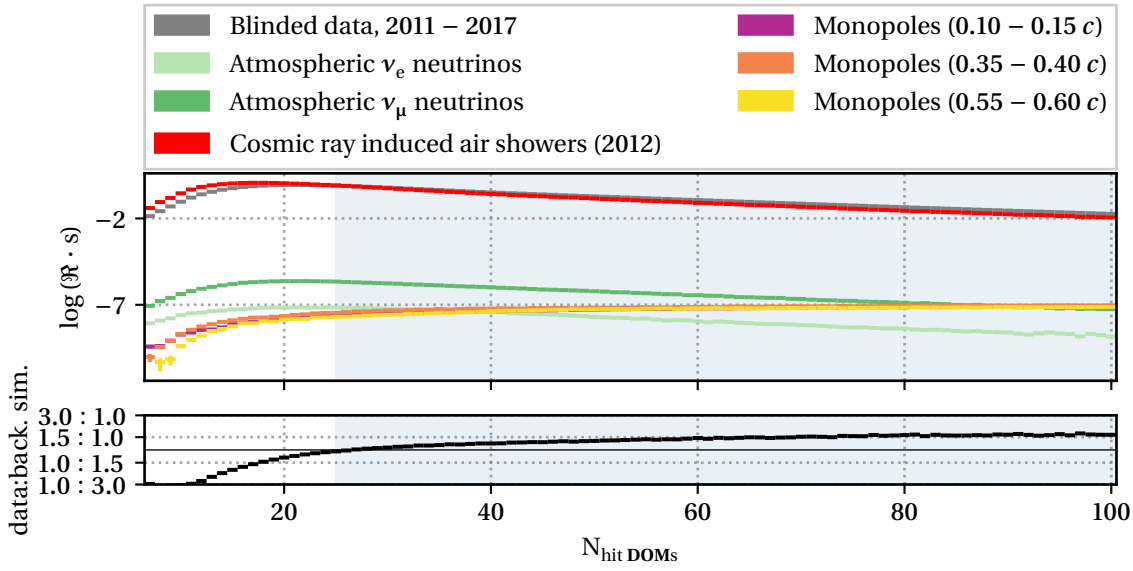


Figure 7.3: **Minimal number of DOMs hit selection**
Two vertically stacked plots are depicted. On the shared horizontal axis, $N_{hit\ \mathbf{DOMs}}$ is shown on a linear scale. $N_{hit\ \mathbf{DOMs}}$ is utilized in Section 7.2.2.1 as a selection variable. The top vertical axis is used to show the natural rate $\mathcal{R}$ on a logarithmic scale while the bottom vertical axis is used as an indicator of the ratio between the simulated background and the measured blinded data. The shaded region indicates the selection area as described in Section 7.2.2.1.

7.2.2.2 Level 5-6: Time-Wise Event Splitting

The event is split time-wise. This means that all hits within an event are separated into two groups of hit selection, h_{first} and h_{last}. Let t_{i_j} denote the time of the i^{th} hit at the j^{th} **DOM**, then the two hit sets are $h_{\text{first}} = \left\{t_{i_j} < \tilde{t}_{0_j}\right\}$ and $h_{\text{last}} = \left\{t_{i_j} > \tilde{t}_{0_j}\right\}$. $\tilde{t}_{0_j}$ is the median time of the first time any **DOM** is hit. Next, the time-wise double split **iLF**s (**TWDSiLF**) are computed on these hit selections resulting in two **iLF**s.

The level 5 event selection requires that the reconstructed velocity $\beta_{\text{first}}$is within 0.15 c to 0.65 c while the level 6 event selection requires that $\beta_{\text{last}}$is within 0.07 c to 0.8 c. Both distributions are depicted in Figure 7.4 and Figure 7.5. The selection criteria based on $\beta_{\text{first}}$is imposed first leading to a softer distribution of β_{last}additionally to a worse reconstruction of β_{last}in the first place. Thus, softer selection criteria where chosen for β_{last}.

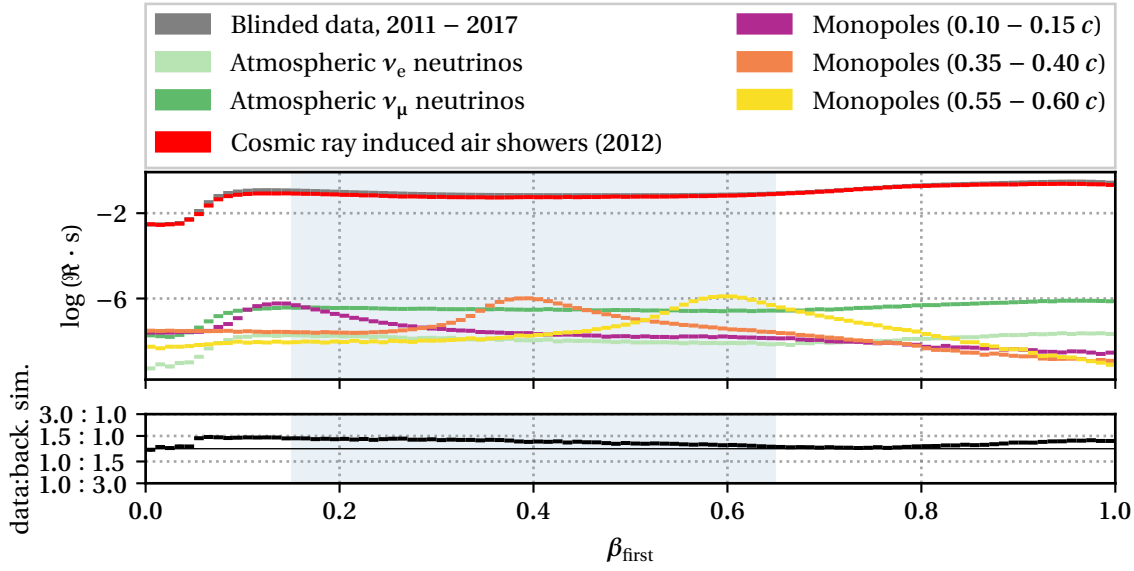


Figure 7.4: **Reconstructed velocity based on first half of event**
Two vertically stacked plots are depicted. On the shared horizontal axis, $\beta_{first}$is shown on a linear scale. $\beta_{first}$is utilized in Section 7.2.2.2 as a selection variable. The top vertical axis is used to show the natural rate $\mathfrak{R}$ on a logarithmic scale while the bottom vertical axis is used as an indicator of the ratio between the simulated background and the measured blinded data. The shaded region indicates the selection area as described in Section 7.2.2.2.

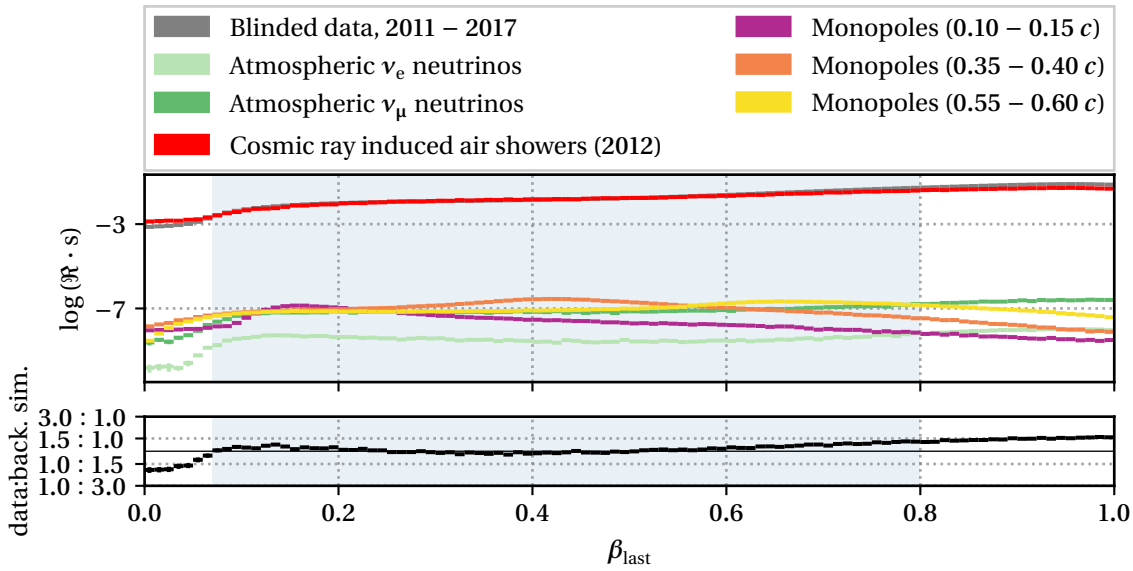


Figure 7.5: **Reconstructed velocity based on second half of event**
Two vertically stacked plots are depicted. On the shared horizontal axis, β_{last} is shown on a linear scale. β_{last} is utilized in Section 7.2.2.2 as a selection variable. The top vertical axis is used to show the natural rate $\mathscr{R}$ on a logarithmic scale while the bottom vertical axis is used as an indicator of the ratio between the simulated background and the measured blinded data. The shaded region indicates the selection area as described in Section 7.2.2.2.

7.2.3 Level 7: Global Linefit Cut

One requirement to pass the **MPF16** filter is a reconstructed velocity of the **GMF16TH** below 0.8 c. In this thesis, the low relativistic velocity regime is of interest. Consequently, this selection criteria is hardened. Only events with a reconstructed velocity of the **GMF16TH** within the aforementioned regime are selected, 0.1 c to 0.6 c to be precise.

While the low relativistic regime ends at 0.55 c, some overlap was chosen with a previous analysis searching for the same signal signature in the mildly relativistic regime, described in reference [Pol18]. The distribution of the reconstructed velocity is depicted in Figure 7.6.

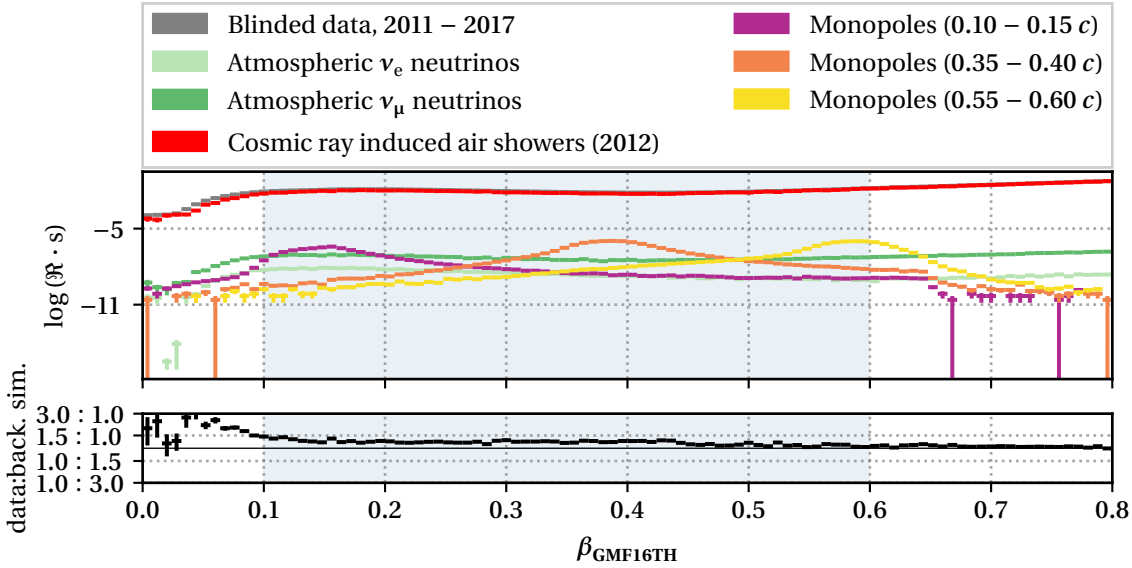


Figure 7.6: **Reconstructed velocity based on MPF16 filter**
Two vertically stacked plots are depicted. On the shared horizontal axis, $\beta_{GMF16TH}$ is shown on a linear scale. $\beta_{GMF16TH}$ is utilized in Section 7.2.2.2 as a selection variable. The top vertical axis is used to show the natural rate $\mathscr{R}$ on a logarithmic scale while the bottom vertical axis is used as an indicator of the ratio between the simulated background and the measured blinded data. The shaded region indicates the selection area as described in Section 7.2.3.

7.2.4 Level 8: Corner Clipper Removal

As a last quality selection step, corner clipping events are removed. These events happen in the corners of the detector and do not pass through much active detection medium making the differentiation of a track-like event from a cascade-like event challenging while offering virtually no benefit to signal selection. A cylinder is placed at the center of the active detection medium with a height and radius of 750 m. Afterwards, $l^{\text{r=750 m}}_{\text{h=750 m}}$, the length of the **GMF16TH** inside the cylinder with the given radius and height is calculated. The distribution of the lengths can be seen in Figure 7.7. A sketch of the cylinder is depicted in Figure 7.8. The track length within the cylinder needs to be at least 250 m for the event to be selected.

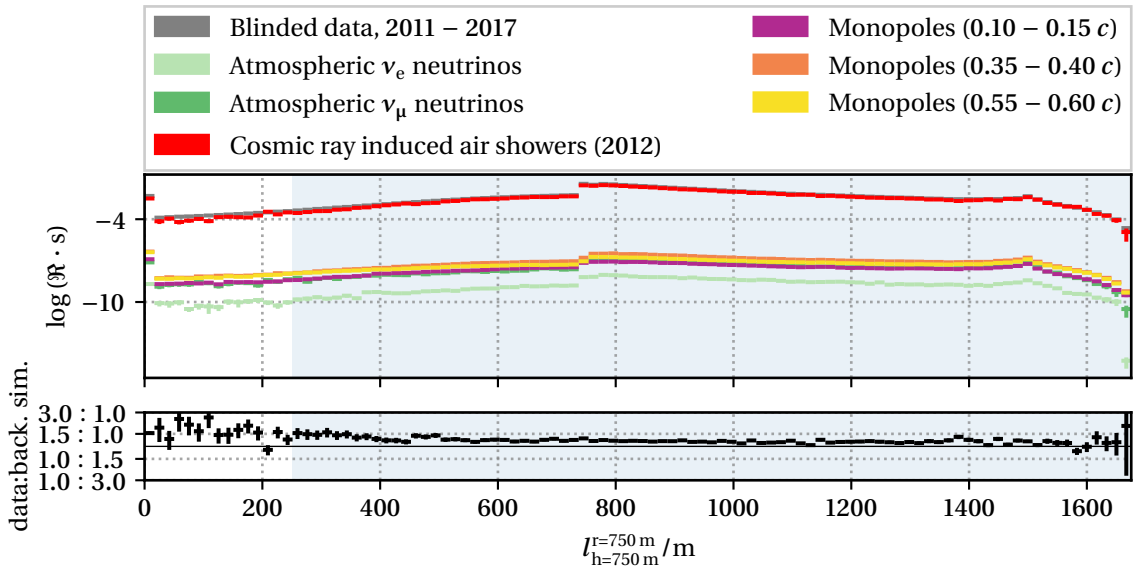


Figure 7.7: **Passed active volume**
Two vertically stacked plots are depicted. On the shared horizontal axis, $l^{r=750\,m}_{h=750\,m}$ is shown on a linear scale. $l^{r=750\,m}_{h=750\,m}$ is utilized in Section 7.2.4 as a selection variable. The top vertical axis is used to show the natural rate $\mathscr{R}$ on a logarithmic scale while the bottom vertical axis is used as an indicator of the ratio between the simulated background and the measured blinded data. The shaded region indicates the selection area as described in Section 7.2.4.

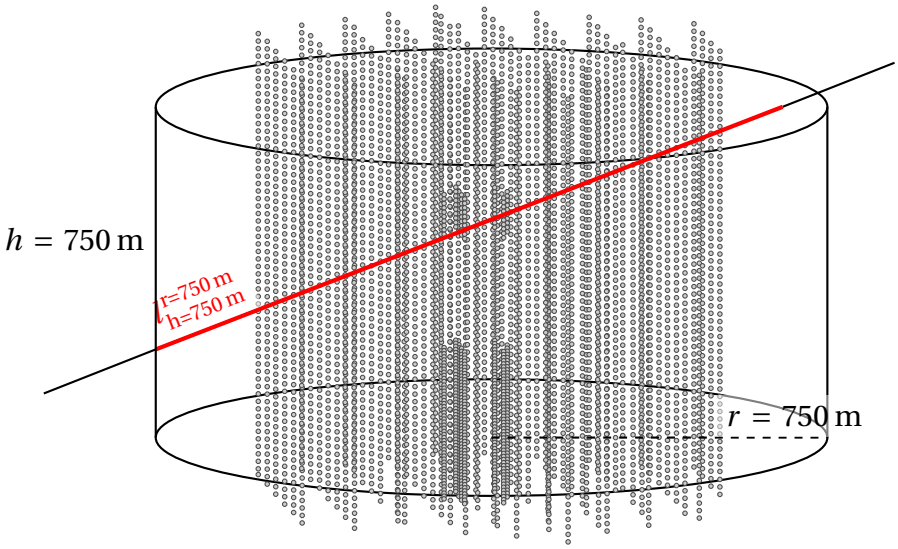


Figure 7.8: **Sketch of track length calculation inside cylinder**
*A sketch of the positions of the **DOM**s of **IceCube** is shown superimposed with a cylinder of a height and radius of 750 m. Position and relation of all elements are to scale besides an increased radius of the **DOM**s by a factor 50. A demo track is shown in black with the length inside the cylinder highlighted in red.*

7.3 Level 9: Machine Learning Based Selection

As a last event selection, a machine learning (**ML**) based, binary classification algorithm is utilized to select candidate events. In Chapter 4, the required nomenclature for this section is introduced.

A bootstrap aggregating (**bagging**) based approach, described in Section 7.3.4 is chosen. A set of **BDT**s is trained on subsets of the available training set. The resulting set of predictions for each event are either reduced to a single prediction, the mean of the set of predictions, or interpreted as the probability density of the event to pass the selection step.

The selection of events utilized for the training of the set of **BDT**s is described in Section 7.3.1, followed by the selection and description of all input features in Section 7.3.2 and training process in Section 7.3.3.

After the **BDT**s are trained, the last selection step is based on the predictions of the set of **BDT**s. The optimization of the final selection step is described in Sections 7.3.5 to 7.3.7. The reconstruction of the most probable true kinematic attributes of the remaining events is outlined in Section 7.3.8. The influence of potential systematic shifts from the assumed models of the detection medium, **MM**-ice interactions, and detector response is discussed in Section 7.3.9.

The expected result after the *unblinding* process, i. e., application of the analysis to all available events is described in Section 7.3.10.

7.3.1 Events Utilized for Training

To train the **BDT**s, a set of events with known class are required. All simulated events, both for cosmic ray induced air showers and atmospheric neutrinos are utilized as examples of background-like events. As the cosmic ray induced air showers (2012) set has very limited statistics, two extra datasets with simulated cosmic ray induced air shower events are included: the cosmic ray induced air showers (2016) set and the cosmic ray induced air showers (2016)c set with an additional $\mathcal{O}\left(4 \cdot 10^6\right)$ events each to compliment the $\mathcal{O}\left(6 \cdot 10^4\right)$ events remaining at this selection stage in the cosmic ray induced air showers (2012) set. These sets are created similarly to the cosmic ray induced air showers (2012) set, but have been simulated with updated settings and two different air shower interaction models.

Furthermore, 10 % of the events in the blinded data, 2011-2017 set are also included as background-like events, enhancing the statistics of the training set by $\approx 1.5 \cdot 10^6$ events. While the measured data may include **MM**s, the number of **MM**s consistent with previous flux limits at this selection stage in the 10 % of the blinded data, 2011-2017 set is < 20 **MM**s and can be neglected during the training stage.

The signal-like events utilized for training consist of the signal simulation with a flat velocity spectrum before any selection step. One half of the signal simulation is calculated for a luminescence efficiency of 1 γ/MeV, the other with 0.2 γ/MeV. A mixture of simulation with two light yields is used to harden the **BDT**s against potential systematic shifts in the luminescence light yield.

While the inter-set weighting is kept within both classes, the integrated weight of both the signal-like and the background-like class is re-normalized to 1 to compensate for the different statistics and debias the predictors from the total integrated natural rate of events of each class.

7.3.2 Input Feature Selection

For any **ML** based algorithm, choosing the correct input variables or *features* is crucial. Adding more features increases the amount of information space that can be sampled and consequently allows for stronger predictions to be made as described in Section 4.1.

33 features are suspected to be usable as separators between signal and background by visual inspection. These features are chosen from standardized reconstructions existing within the **ICC**, features which are reconstructed as by-products of the reconstruction steps in the **MPF16** filter, and previous selection steps. Features can directly separate, e. g., reconstructed velocity or indirectly, e. g., providing contextual information required to interpret other features.

For a feature to be usable for this analysis, a close match between simulated background and measured data at early selection stages is required. To ensure only suitable features are utilized, an iterative process is used.

A specialized **BDT** is trained to separate the background simulation set, the combination of the cosmic ray induced air showers (2012) set, cosmic ray induced air showers (2016) set, cosmic ray induced air showers (2016)c set, atmospheric ν_μ set, and atmospheric ν_e set from the blinded data, 2011-2017 set. The same training configuration as for the final **BDT**s is used.

The available training data is split into two similarly sized sets. One set is used for training, the other to evaluate the performance. The separation powers of the input features are calculated and the highest separating feature is removed.
The process is repeated until separation power is equalized between all features. Two metrics are utilized to define the separation power of a feature: the number of singular variate decisions (**SVD**s) operating on the feature, also called the *weight* of the feature, and the increase in accuracy of the model due to the inclusion of the feature, also called the *gain*. Note that the input features are correlated and thus might share importance based on the underlying correlation.
In total, 24 features remain. While this may seem like a rather high number of parameters, not every variable will be reconstructed for each event. 16 variables are defined by data taken within the **IC**, the other 8 are only reconstructed if enough hits occurred in the **DC**. All chosen variables are described below and are depicted in Figures D.1 to D.24.
In Figure 7.9, the maximal separation of simulated background and measured data, i. e., the score distribution of the final, feature rejection **BDT**, is illustrated. The two feature separation power metrics for this **BDT** are depicted in Figures E.1 and E.3. These can be used as a reference to compare the feature importance of the final, signal and background separating **BDT**s to highlight any potential influence of mismatches between simulated background events and measured data.

7.3.2.1 Global MonopoleFilter_16 Track Hypotheses Based Features

Two features are taken from the global **MPF16** track hypotheses (**GMF16TH**) as input features for the **BDT**s. The first feature is the reconstructed velocity $\beta_{\textbf{GMF16TH}}$ as described in Section 7.2.3. The second feature is the number of hits $N_{\text{hit GMF16TH}}$ the last linear regression of the **GMF16TH** is performed on. This number differs from the number of hits in the event due to internal event cleaning inside the **iLF** as described in Appendix A.

7.3.2.2 Double Splitting Based Features

Four features based on the time-wise double split **iLF**s (**TWDSiLF**) are utilized as **BDT** input features. A description of the **TWDSiLF** can be found in a Section 7.2.2. With the two **iLF**s expressed by

$$\boldsymbol{r}\,(t)_{\text{first/last}} = \boldsymbol{r}_{\text{first/last}} + t \cdot \boldsymbol{v}_{\text{first/last}}, \tag{7.5}$$

the input features are

$$\measuredangle\left(\boldsymbol{v}_{\text{first}}, \boldsymbol{v}_{\text{last}}\right) = \boldsymbol{v}_{\text{first}} \cdot \boldsymbol{v}_{\text{last}}, \tag{7.6}$$

the angle between the two **iLF**s,

$$\beta_{\text{first/last}} = \left|\boldsymbol{v}_{\text{first/last}}\right|, \tag{7.7}$$

the reconstructed velocity of the two **iLF**s, and the distance of closest approach (**DOCA**) between the two **iLF**s defined by

$$\textbf{DOCA} = \left|\boldsymbol{r}_{\Delta} - \boldsymbol{v}_{\Delta} \cdot \frac{\boldsymbol{r}_{\Delta} \cdot \boldsymbol{v}_{\Delta}}{(\boldsymbol{v}_{\Delta})^2}\right| \tag{7.8}$$

where

$$\boldsymbol{r}_{\Delta} = \boldsymbol{r}_{\text{first}} - \boldsymbol{r}_{\text{last}} \tag{7.9}$$

is the difference between the two support vectors and

$$\boldsymbol{v}_{\Delta} = \boldsymbol{v}_{\text{first}} - \boldsymbol{v}_{\text{last}} \tag{7.10}$$

is the difference of the two velocity vectors. $\beta_{\text{first/last}}$ is described in Section 7.2.2.2. $\measuredangle\left(\boldsymbol{v}_{\text{first}}, \boldsymbol{v}_{\text{last}}\right)$ is a scalar value which indicates how well the two **iLF**s directionally align while **DOCA** indicates any spatial or temporal offset between the two **iLF**s. Combining these features encodes the requirement that both **iLF**s describe a single particle traveling through the detector on a linear trajectory.

7.3.2.3 Hit Digital Optical Modules Projected on Global MonopoleFilter_16 Track Hypotheses

Seven features are based on the distribution of hits projected on the **GMF16TH**. Three distribution attributes are calculated for both the hits happening inside the **IC** and the **DC**, one is only calculated for the hits registered by the **DC**. Let $\{(t_0, l_0, d_0, w_0), \ldots (t_N, l_N, d_N, w_N)\}_{\textbf{IC/DC}}$ be the sets of all hits registered by either the **IC** or **DC** where t_i is the time the hit is detected, l_i is the projected position on the **GMF16TH**, d_i is the orthogonal distance of the hit from the **GMF16TH**, and w_i is the number of reconstructed photons which caused the hit. An illustration of these features is depicted in Figure 7.10. As a convention, the set is sorted in time so $t_i \leq t_{i+1}$ and the lengths are shifted so $\min l_i = 0$.
The first distribution feature is the smoothness of projected hits (**SPH**) defined by

$$\textbf{SPH}_{\textbf{IC/DC}} = \frac{j}{N} - \frac{l_j}{l_N} \tag{7.11}$$

where j is the first element to fulfill

$$\left|\frac{j}{N} - \frac{l_j}{l_N}\right| = \max_{i \in (1,\ldots,N)} \left|\frac{i}{N} - \frac{l_i}{l_N}\right|. \tag{7.12}$$

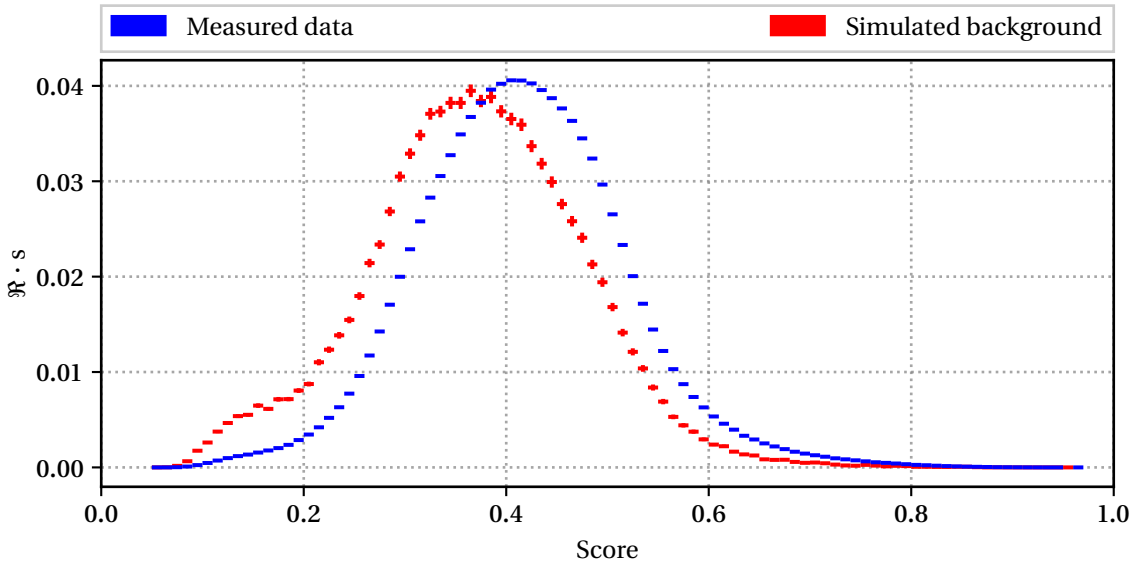


Figure 7.9: **Score distribution of final BDT of input feature selection**
*The distribution of events as predicted by the final **BDT** of the feature selection process as described in Section 7.3.2 is depicted. The vertical, linear axis indicates the rate of events in natural rate, the horizontal, linear scale indicates the predicted score for the events. In red, the distribution for simulated background events is shown while blue is utilized to indicate the measured data distribution. In an ideal case, a perfect match is supposed to be observed. Instead, two nearly Gaussian distributions with some offset can be seen. While this indicates some separation between the two datasets is still possible, no major separation capabilities for the given set of features can be expected.*

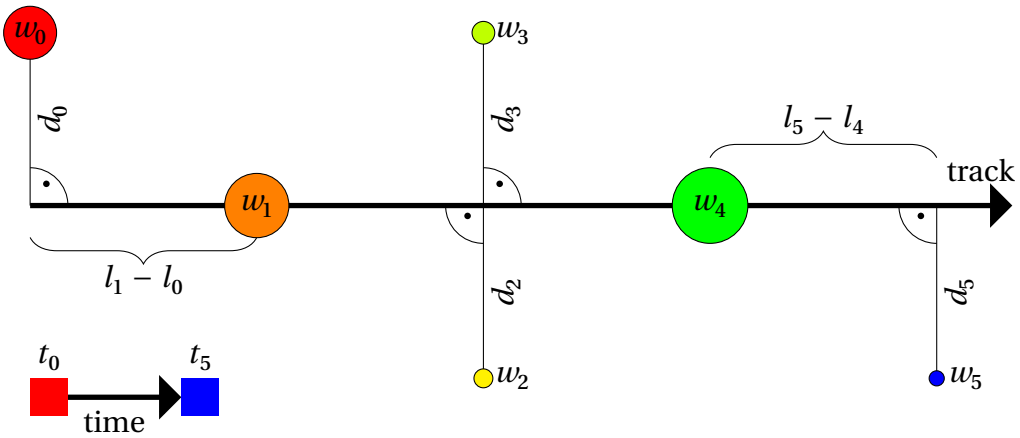


Figure 7.10: **Illustration of projected hits on track**
*All variables based on the relationship of hit **DOM**s and reconstructed track hypothesis are sketched out. Each sphere indicates a hit **DOM** with the radius indicating the recorded charge w_i and the color corresponding to the time t_i of the hit . l_i indicates the location of the hit projected onto the track while d_i is the orthogonal distance from the track hypothesis.*

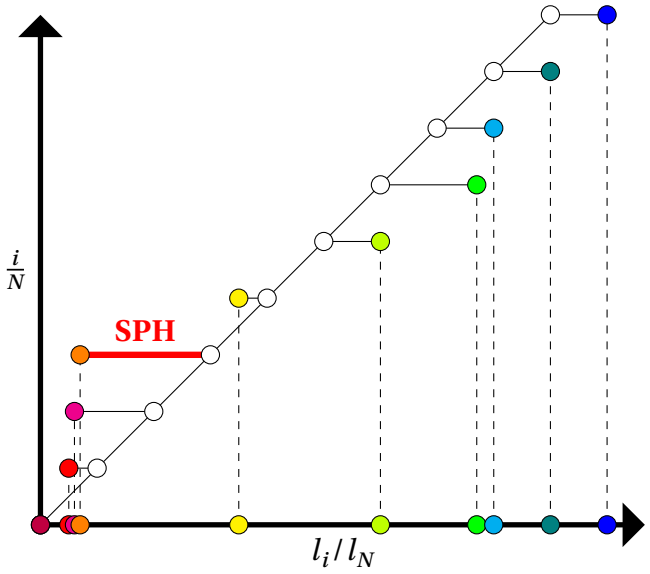


Figure 7.11: **Illustration of the smoothness of projected hits**
*A plot illustrating smoothness of projected hits (**SPH**) is shown. An event consisting of 10 hits, indicated by colored spheres, is utilized as an example. A horizontal arrow is drawn to indicate a track hypothesis. The hits are projected onto aforementioned track. The vertical direction is used to indicate the order in which the hits occurred. In the case of a perfectly homogeneously illuminated track, the hits would be situated on a 45 ° line between the two arrows. Their ideal position is indicated by non-filled spheres. The maximal distance between spheres and their ideal position is the **SPH**, here highlighted in red.*

This feature describes how even the hits are distributed on the track. A sketch of the setup is drawn in Figure 7.11. As an **MM** would constantly produce light along the track, the track of a signal event should be smooth and **SPH** is close to 0.

The other two dual array distribution attributes are the $\mathbf{MSL_{IC/DC}}$ and $\mathbf{FLQCD_{IC/DC}}$, described in Section 7.1.2.2. As the previous attribute is not independent of the total track length, this information is needed to utilize the attributes in a meaningful way. E. g., a **MSL** of 50 m might be acceptable for a long **GMF16TH** of 1 km but not for a short one of 100 m.

The last attribute is calculated only for the hits registered in the **DC**. The weighted orthogonal distance from the track hypothesis (**WODTH**) is also calculated based on the **GMF16TH** by

$$\mathbf{WODTH_{DC}} = \frac{\sum_{i=i}^{N} d_i \cdot w_i}{\sum_{i=1}^{N} w_i}. \tag{7.13}$$

This is only useful for the fraction of events that pass through the **DC**. It adds separation power as a signal-like event should form a cylinder around the track due to the homogeneous light emission while cascade events or mis-reconstructed coincident events vary depending on the topology.

7.3.2.4 Timing Based Features

Four timing based features are selected. One feature is calculated for both the **IC** and **DC**, the other two are only calculated for the **DC**. All features are derived based on the time of the first hit of any **DOM**. The first hit at a **DOM** has the most reliable information about the distance of a light inducing particle to the **DOM**. The hit inducing photon is most likely to have traversed directly between the particle and the **DOM**. Later hits may have been delayed due to scattering inside the detection medium. Let $\{t_1, \dots t_N\}_{\mathbf{IC/DC}}$ be the time sorted set of times of the first hit at a **DOM** in a given array. The first feature, the temporal maximal gap (**TMG**), is both calculated for the **IC** and the **DC** and is defined by

$$\mathbf{TMG}_{\mathbf{IC/DC}} = \max_{i \in (1,\dots,N-1)} t_{i+1} - t_i. \tag{7.14}$$

For a signal-like event, this feature should have small values as the light production is homogeneous while cascade-like events should have large values as in between of cascades, no light is being produced. The other two features are the $\mathbf{TBH}_{\mathbf{DC}}$ and $\mathbf{FWHM}_{\mathbf{DC}}$, both described in Section 7.1.2. Both are used in the **MPF16** filter, **DC** selection and are indicators of the temporal length of the event. Signal-like events should be temporally longer than background events as the incident particle has a lower velocity, requiring more time to pass through the detector.

7.3.2.5 Topology Based Features

Seven topology based features are selected. Three are based on the track length within different areas of the detection medium of the **IC**. Another three describe the height of the event within the **IC** while the last is the number of strings with any hits in the **DC**. The latter is referred to as $N\mathrm{strings}_{\mathbf{DC}}$.

The next three are $l^{\mathrm{r=500\,m}}_{\mathrm{h=500\,m}}$, $l^{\mathrm{r=750\,m}}_{\mathrm{h=750\,m}}$, and $l^{\mathrm{r=1000\,m}}_{\mathrm{h=1000\,m}}$. $l^{\mathrm{r=750\,m}}_{\mathrm{h=750\,m}}$ is already described in Section 7.2.4. $l^{\mathrm{r=500\,m}}_{\mathrm{h=500\,m}}$ and $l^{\mathrm{r=1000\,m}}_{\mathrm{h=1000\,m}}$ are calculated exactly the same way but are computed for different cylinder heights and radii (500 m and 1000 m). These are indicators of the length of a potential particle's track in different parts of the in-ice arrays of **IceCube**.

The last three features are indicators of the height of the event within the **IC**. The attributes of the detection medium of the **IC** depend on the height as described in Section 2.1.1 and hence may influence the quality and values of other features. All three features are based on the height of the hit **DOM**s used for the **GMF16TH**. With $\{z_1, \dots z_N\}$ being the set of all heights of all hit **DOM**s, the included features are

$$\bar{z} = \frac{\sum_{i=1}^{N} z_i}{N}, \tag{7.15}$$

the mean of all heights and

$$z_{\text{min/max}} = \min / \max z_i, \tag{7.16}$$

the minimal and maximal height. Note that these features are not weighted and thus only describe the height of the event, not the light deposit.

7.3.3 Boosted Decision Tree Training Configuration

All **BDT**s are trained with the same settings, derived from pre-studies, yielding **BDT**s rather in the slightly under-trained than over-trained regime. Following the nomenclature in Chapter 4, each **BDT** consists of 500 decision trees (**DT**s), each with a depth of 6. Sub-sampling with a factor of $S = 0.5$ is applied as well as dropout regularization, with a probability to apply the regularization of $D = 0.7$ and a fraction of predictors ignored in the prediction of $R = 0.3$. A learning rate of $\eta = 0.05$ is used. As binary regression is to be achieved, the cross entropy is chosen as the loss function.

7.3.3.1 Validation

Before the set of **BDT**s can be utilized in the final selection stage, their performance has to be validated. As **BDT**s are only applied and evaluated based on events disjunct from their training set, over-training is only an issue of sub-optimal overall performance but not of overestimated performance. Instead, overestimated performance may arise from potential background simulation and measured data mismatches albeit the precautions described in Section 7.3.2.
Comparing the importance of features in Figures E.1 and E.3 and Figures E.2 and E.4 does not indicate reliance on similar features and thus strong separation between simulated events and measured events.
Comparing the score distributions of the **BDT**s for simulated events and measured events, shown in Figure 7.12, does not indicate a mismatch between the two.

7.3.4 Bootstrap Aggregating

Bootstrap aggregating (**bagging**), introduced in reference [Bre96], is a method in which a set of homogeneous classifiers $\{C_0, \ldots, C_M\}$ are combined into a single classifier C.
In this analysis, **BDT**s are utilized as binary classifiers as described in Section 4.2. The training set is split into randomly sampled subsets. These subsets are in turn utilized to build the individual **BDT**s.

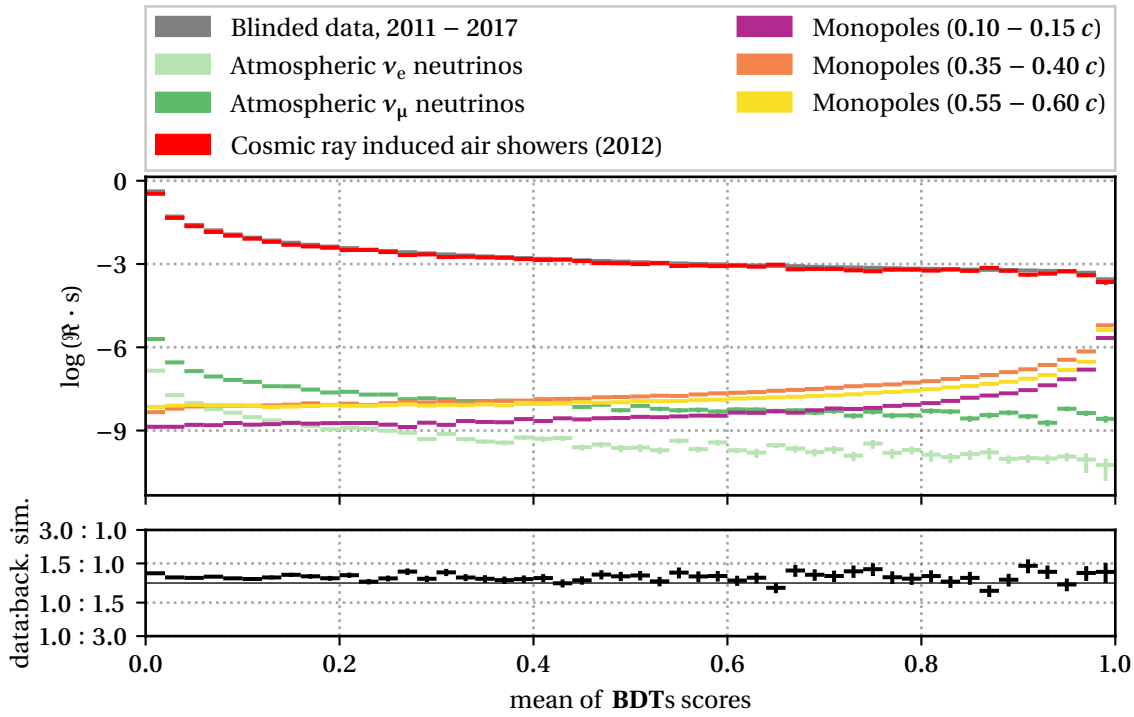


Figure 7.12: **Mean score distributions of BDTs**
*Two vertically stacked plots are depicted. On the shared horizontal axis, the mean of all **BDT**s is shown on a linear scale. The top vertical axis is used to show the natural rate $\mathscr{R}$ on a logarithmic scale while the bottom vertical axis is used as an indicator of the ratio between the simulated background and the measured blinded data. A close match between the simulated background events and the measured data is observed as expected before the last selection step.*

Consequently, each event e_j has a set of classifications $\{c_0, \ldots, c_M\}_{e_j}$. This set can be reduced to a single classification, here by calculating the mean $\bar{c}_{e_j}$. After defining the passing value to be classified as part of the positive class S, any event with $\bar{c}_{e_j} \geq S$ is identified as being signal-like.

7.3.4.1 Configuration

A set of 1000 **BDT**s, each with the same configuration, is trained on a random subsets of the total available training data described in Section 7.3.4. 1000 was chosen because the ratio between the average rate of events of the coursed simulated background set and the rate of a single signal event is $\mathscr{O}(1000)$. While more **BDT**s can be added, training time of the whole set also increases linearly with size.

Each event in the training set has a 10 % chance to get assigned to each subset. The prediction of a **BDT** on the event of the subsets used for training are masked from further usage to prevent biasing towards these subsets. Consequently, every event has an average of 900 classification predictions.

7.3.4.2 Probabilistic Interpretation of the Classification Predictions

Alternatively to selecting events as described in the previous sections, the distribution of classifications can be treated probabilistic. This is useful to derive information about the average distribution of any attribute of an event after a classification has been made in the low statistic limit. Let an event e_j have a weight of w_j. By interpreting the distribution of $\{c_0, \dots, c_M\}_{e_j}$ as the probability distribution of the event to pass the selection step $P\left(e_j|S\right)$, a probable distribution about any attribute of the events can be created by reweighting the events from w_j to $w_j \cdot P\left(e_j|S\right)$. The setup is sketched out in Figure 7.13. Consequently, more events contribute to the post selection distribution albeit with reduced weights creating a smooth approximation of the actual underlying distribution.

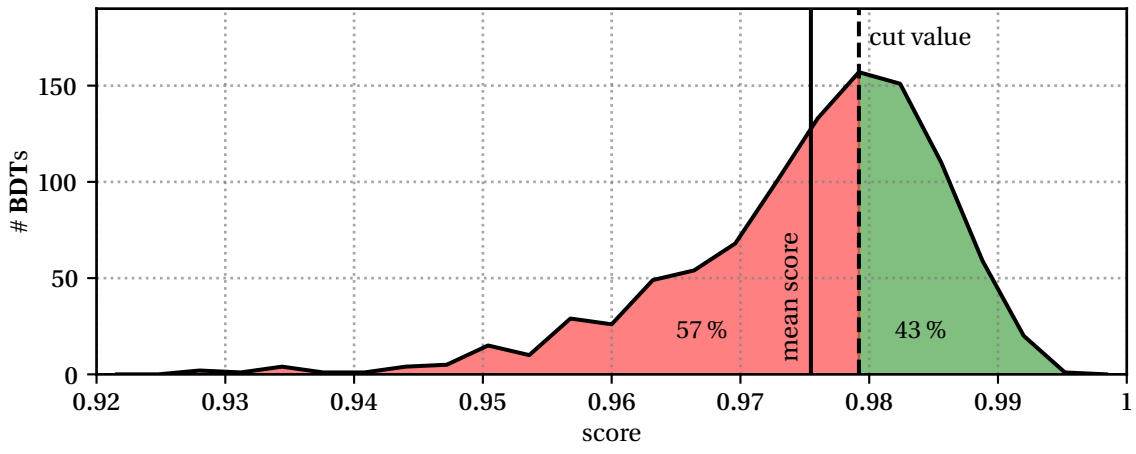


Figure 7.13: **Example of probabilistic treatment of a single event**
*A histogram of the number of **BDT**s predicting a certain score of a singular, exemplary event based on an **MM** is depicted. The mean score of all **BDT**s is highlighted by a black, vertical, solid line. Assuming an arbitrary cut value of 0.9792, the event would not be selected based on the mean value of all **BDT**s and consequently not contribute to any further distribution after the selection step. By interpreting the distribution probabilistic, the event still contributes but by a reduced weight of 43 % of its original weight.*

7.3.5 Feldman-Cousins Sensitivity

In event-counting-based experiments, the result of the i^{th} measurement is the number of observed events N_i^{O} after a predetermined time. N_i^{O} is the sum of the unknown number of observed background-like events during the measurement N_i^{B} and the unknown number of observed signal-like events during the measurement N_i^{S}. If the experiment is repeatable and the average number of background-like events $\overline{N}^{\text{B}}$ known, the average number of detected signal-like events $\overline{N}^{\text{S}}$ can be derived by

$$\overline{N}^{\text{S}} = \overline{N}^{\text{O}} - \overline{N}^{\text{B}}. \tag{7.17}$$

In the case of a single measurement with known $\overline{N}^{\text{B}}$, only an upper limit of signal-like events compatible with the measurement can be calculated. To derive this upper limit in this work, the Feldman-Cousins approach introduced in reference [FC98] is used. It can be applied in the small signal regime and is not biased towards the measured data. For this approach

$$\int_{N^{\text{S}}_{\text{min}}}^{N^{\text{S}}_{\text{max}}} P\left(N^{\text{O}} \middle| N^{\text{S}}, \overline{N}^{\text{B}}\right) dN^{\text{S}} = \alpha \tag{7.18}$$

needs to be solved where $N^{\text{S}}_{\text{min}}$ to $N^{\text{S}}_{\text{max}}$ is the interval the true N^{S} is situated in, with a probability of α. $P\left(N^{\text{O}} \middle| N^{\text{S}}, \overline{N}^{\text{B}}\right)$ is the probability to measure N^{O} if the true signal count is N^{S} with a mean, true background of $\overline{N}^{\text{B}}$. For this, a Poissonian probability density is used, implemented by a modified Gamma distribution and consequently defined for all real values.

In the subsequent analysis steps, only $N^{\text{S}}_{\text{max}}$ is of interest as $N^{\text{S}}_{\text{min}}$ will be 0. Furthermore, $\overline{N}^{\text{B}}$ has a noticeable uncertainty. This is taken into account by folding the maximal number of signal events $N^{\text{S}}_{\text{max}}$ consistent with N^{B} with the probability for N^{B} to be the true number of background events to be observed in the experiment or

$$\int_{0}^{\infty} P\left(N^{\text{B}} \middle| \overline{N}^{\text{B}}\right) N^{\text{S}}_{\text{max}}\left(N^{\text{O}}, N^{\text{B}}\right) dN^{\text{B}} = \mu^{\alpha}_{N}. \tag{7.19}$$

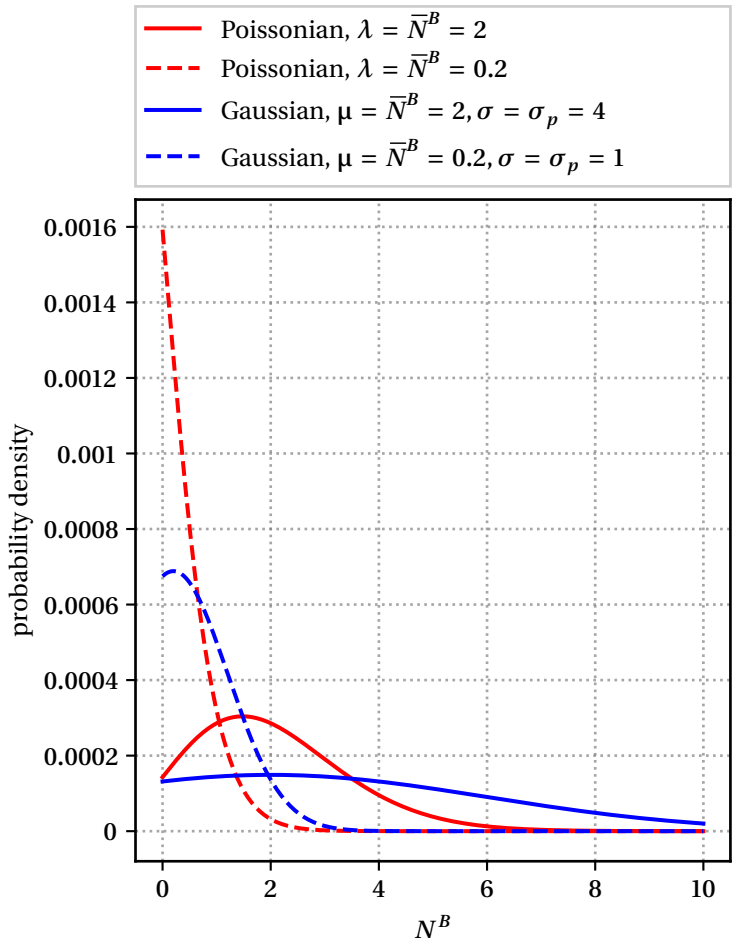


Figure 7.14: **Comparison of Poissonian and Gaussian approximation**
A plot of the probability density as a function of N^B for different $\overline{N}^B$ with associated uncertainty is depicted. Red is used for Poissonian distributions, blue for Gaussian distributions. As exemplary values, the expected number after the final selection step described in Section 7.3.7 are used. It can be observed that the Gaussian approximation predicts higher values of N^B.

In this thesis, α is chosen to be 90 %. While a Poissonian distribution was regarded to approximate the distribution of N^{B}, a Gaussian distribution is utilized instead to accommodate the uncertainty of $\overline{N}^{\mathrm{B}}$. Consequently, the expected number of background events is higher and the derived upper average number of compatible signal-like events smaller and thus conservative. A comparison between the probability densities for N^{B} assuming a Poissonian distribution and a Gaussian distribution is depicted in Figure 7.14.

$\mu_N^{90\,\%}$ is the average upper number of true signal counts compatible with the observed counts and the predicted average number of background events if the experiment was repeated an infinite amount of times.

While in this thesis, the experiment, the unblinding of the measured data, can only be done once, this parameter will be used as the upper number of signal-like events within the recorded data sample.

7.3.6 Model Rejection Factor

Separating events sampled from two different distribution can be accomplished by varying selection techniques. In case of an **SVD**, described in Section 4.3.1, the optimization of the separation is the same as optimizing the *cut* value c.
Deducing C, the optimal value of c, can be accomplished in multiple ways depending on the objectives of the selection. In the case that the underlying true distribution is not known and only sampled data points are available like in this thesis, the value of c should not be biased towards the sampled data points. On top of this, the value of c should be optimized to offer the best discovery potential for the proposed signal. This is achieved by optimizing the model rejection factor (**MRF**).
The **MRF**, introduced in reference [HR03], is the ratio between the average upper number of true signal counts $\mu_N^{90\,\%}$ as calculated by the Feldman-Cousins's unified approach, described in Section 7.3.5, and the number of remaining expected signal events N_S, both for a given value of c and thus

$$\textbf{MRF}\,(c) = \frac{\mu_N^{90\,\%}\,(c)}{N_S\,(c)}. \tag{7.20}$$

C is the value that minimizes **MRF** (c).

7.3.7 Optimized Final Selection Step

The final event selection is optimized by minimizing the **MRF** as described in Section 7.3.6. As the underlying distribution, the probabilistic predictions of all **BDT**s, described in Section 7.3.4, is used.
N^{O}, the number of remaining events, is an integer value in counting-based experiments. In this work, due to **bagging**, the projected number of remaining events can be a fraction. This is taken into account by linearly interpolating the between integer steps.
The aforementioned process is applied to both the simulated background data and the blinded, measured data set. The number of remaining events projected for the full measurement time of 7 yr is shown in Figure 7.15. The mean $\bar{p}$ and standard error (**SE**) σ_p is derived by assuming a Gaussian distribution in the predictions of all **BDT**s. As an approximation of the asymmetric errors once the predictions approach 0, σ_p is split into a positive bound σ_p^+ and negative bound σ_p^-. For a given number of **SE** intervals d, once

$$\bar{p} - d \cdot \sigma_p \leq 0, \text{ then} \tag{7.21}$$

$$d \cdot \sigma_p^- = 0 \tag{7.22}$$

and

$$d \cdot \sigma_p^+ = d \cdot \sigma_p - (\bar{p} - d \cdot \sigma_p) = 2 \cdot d \cdot \sigma_p - \bar{p}. \tag{7.23}$$

For the numbers given below, $d = 1$ is chosen.

A one dimensional grid search with a step size of 0.0001 of the cut value c is conducted to find the value C which minimizes the **MRF**. The **MRF** becomes minimized at $C = 0.9997$, both for the simulated data set and the blinded, measured data set. After this selection step, the expected mean number of background events based on the blinded, measured data set is 0 with an upper Gaussian approximated limit of 4. Conducting the same calculation based on the background simulation set results in an expected number of 2 background events with an upper Gaussian approximated limit of 10 events.

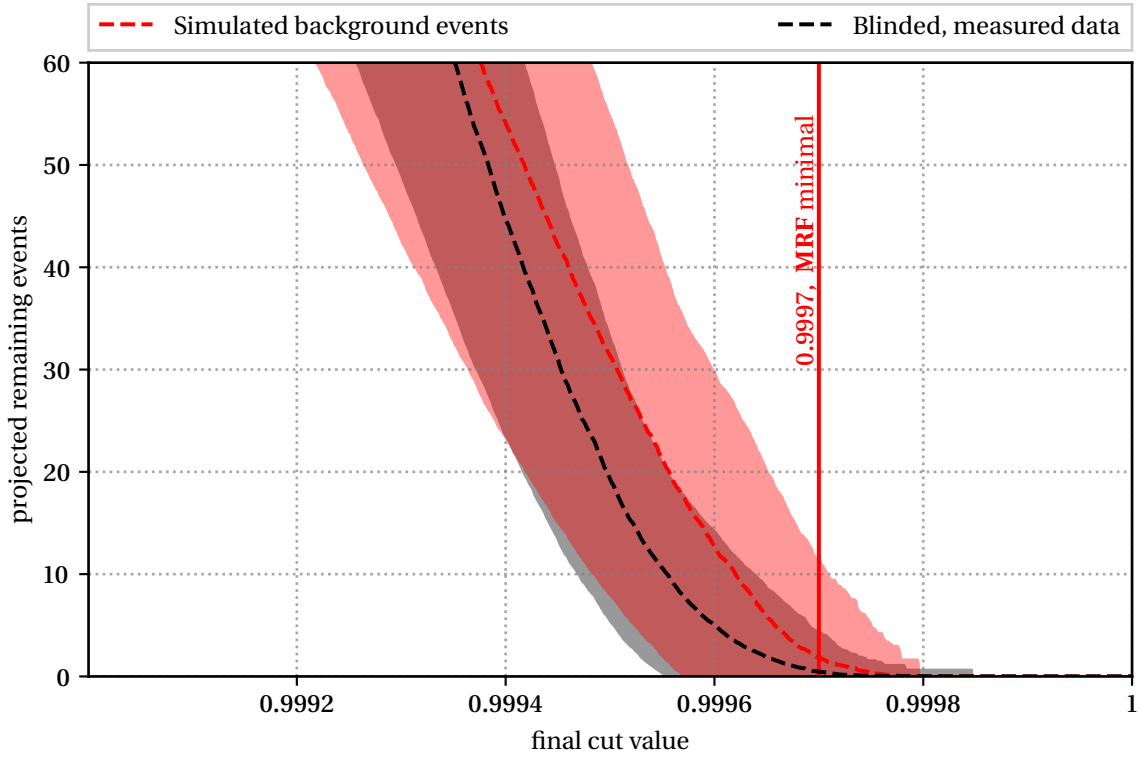


Figure 7.15: **Estimated number of background events**
*A plot of the expected number of remaining background events on the vertical, linear axis after the unblinding process is presented as a function of the final cut value on the horizontal, linear axis. Expected numbers are derived from either simulated background events or measured, blinded data indicated by red and black respectively. The mean expected values are indicated by a dashed line while a colored contour indicates the statistical uncertainty around the mean. A vertical line indicates the points where the **MRF** becomes minimal. This happens in both cases at 0.9997. The magnitudes of the upper bounds are dominated by the underlying statistics and can only be improved by adding more events. These numbers are the integrated amount of expected events over the whole velocity space.*

7.3.8 Reconstructed Kinematic Attributes

To this point, the *reconstructed* kinematic attributes of events are treated as being synonymous with the actual or *true* kinematic attributes. While this was possible as only kinematic distributions of many events have been examined, a more careful treatment is required for singular events.

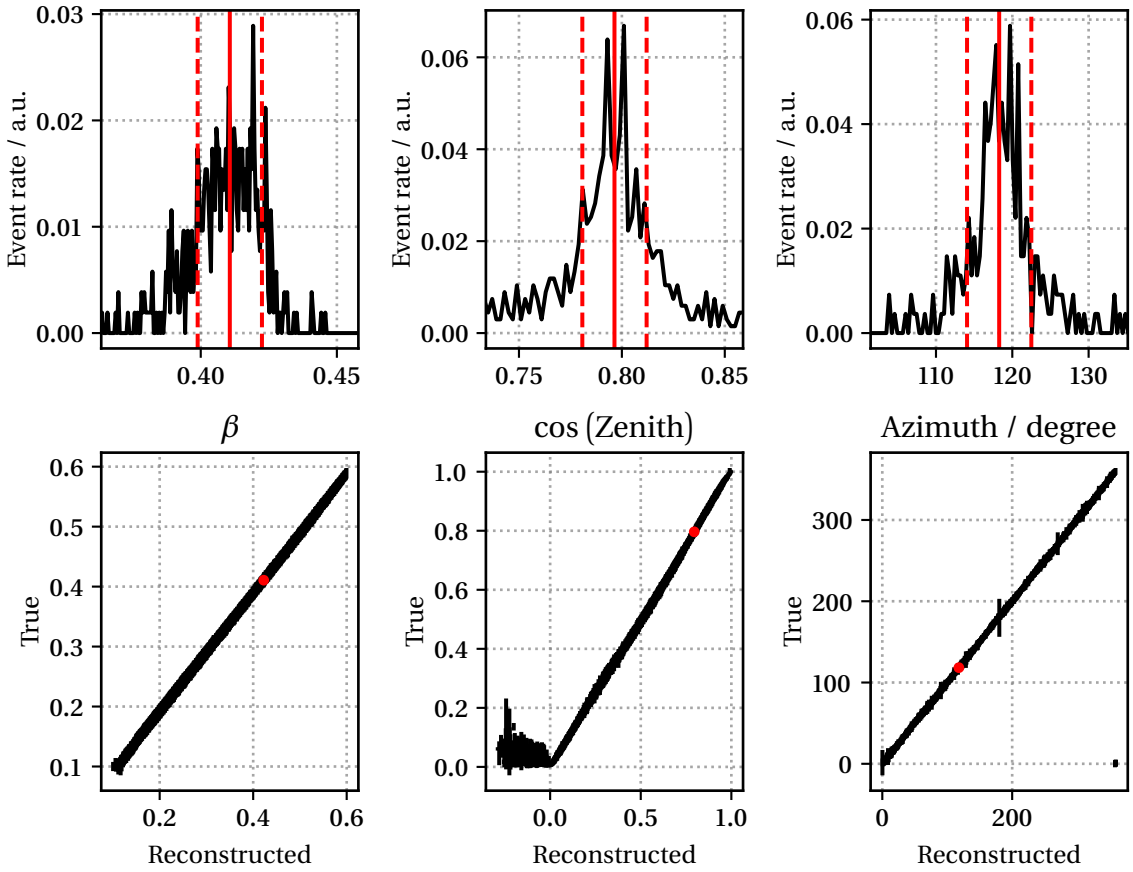


Figure 7.16: **Mapping of reconstructed to true kinematic attributes**
*Six plots in a two to three grid are shown. From left to right, figures related to the kinematic attributes β, zenith, and azimuth are positioned. At the bottom row, the true mean values and **SE** as a function of the reconstructed values of all three kinematic attributes are illustrated. For each kinematic attribute, one example distribution is shown in the top row. There, the distribution of the true kinematic attributes for a specific reconstructed value, indicated by a red in the bottom plot, is illustrated. The horizontal axis indicates the reconstructed value while the vertical axis represents the rate of events. The red, solid, vertical line corresponds to the mean, the red, dashed, vertical lines to the **SE**. The relationship is approximately linear for all kinematic attributes.*

The kinematic attributes are the velocity and direction of the candidate events. The direction is described by two angles, the zenith and the azimuth, depicted in Figure 6.3. To derive the most probable true values with an associated standard deviation, the reconstructed kinematic attributes of simulated signal events and their standard deviation are calculated from simulated events in dependence of the true kinematic attributes. Afterwards, the relation is inverted.

All three parameters are treated independently. Velocity is assumed to be Gaussian shaped with appropriate limits based on the available signal simulation. Same is applied to the cos (zenith) distribution. For the azimuth, a circular Gaussian distribution is applied. Examples for single reconstructed values for all three parameters as well as the total reconstruction are depicted in Figure 7.16.

7.3.9 Systematic Shifts of Detector Response

The simulation of events depends on the correct modeling of the interaction of an **MM** with the detection medium and the corresponding detector response. As these models are the same for all simulated events, a shift from the true values will result in a systematic shift between the signature of simulated events and the signature of measured events.

For this analysis, four systematic effects are investigated and included in the final signal detection efficiency: variations in the detection mediums scattering and absorption, variations in the photon scattering due to uneven freezing after the deployment of the detector, a shift in the photon acceptation rate for all photomultiplier tubes (**PMT**s), and a systematic shift in the luminescence light yield efficiency.

For each signal dataset, an extra dataset with one of the aforementioned systematic shifts is simulated albeit with reduced statistics. For each velocity, the minimal efficiency dataset for each variation is used as the assumed actual efficiency of the detector. Afterwards, the four efficiency are combined into a total effective efficiency shift for signal events which is applied to the final expected number of signal events.

The efficiency is calculated for each individual speed dependent dataset. In between, linear interpolation is applied to derive a systematic efficiency function independent of the underlying binning. The resulting combined systematic efficiency shift is depicted in Figure 7.17.

For simulated background events, no systematic shifted datasets are simulated for the final selection step. The expected number of remaining background events after the final selection step is limited by the available statistics of simulated background events.

Any systematically shifted dataset would encompass a smaller statistic as multiple systematically shifted dataset have to be computed within a reasonable amount of processing time. Consequently, no prediction could be made with such hypothetical systematically shifted dataset. Furthermore, the expected number of background events after the final selection step is close to zero limiting the impact of potential systematic shifts in the number of predicted events.

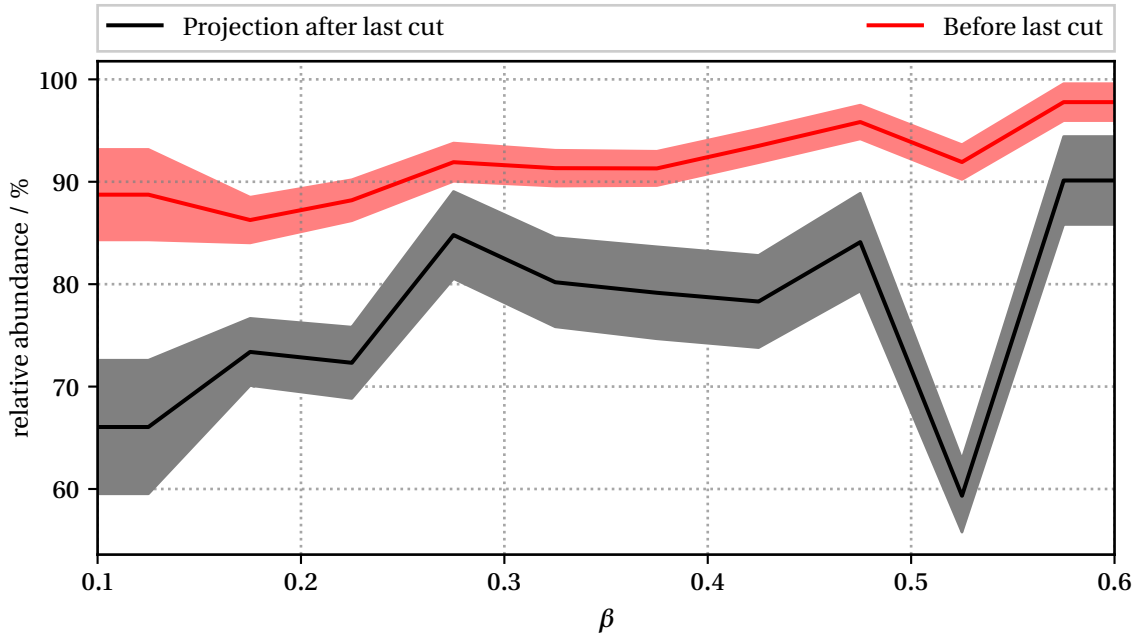


Figure 7.17: **Combined systematic signal efficiency**
*The combined systematic efficiency to detect signal-like events is depicted. On the horizontal axis, the incident **MM** velocity is drawn in fractions of c. On the vertical axis, the relative abundance of **MM**s in % relative to the normal dataset parameters is plotted. The final projected systematic shift is drawn in black while the systematic shift before the final selection step is drawn in red. The mean and **SE** are calculated for the mean velocity of each dataset with linear interpolation between points. The individual behavior of the four underlying examined systematic variations are depicted in Figures C.1 to C.4. A higher systematic shift is observed at lower velocities which is expected due to the brightness of the event having a stronger impact in this region. An additional reduction can be seen in the velocity region 0.38 c to 0.55 c. It can be observed that this reduction is not present before this selection stage. This reduction starts to appear at an unoptimized cut value of 0.996 and gradually becomes more pronounced approaching the optimized cut value of 0.9997.*

However, checks with already existing, low statistic systematic datasets have been performed for lower levels of the selection process. This limits the testable systematic to a change in the **PMT** efficiency. No relative systematic shift in the number of background events greater than the relative shift for the number of signal events could be observed.

7.3.10 Projected Sensitivity

After the final event selection, the expected sensitivity of this analysis in the velocity space can be derived. First, the velocity spectrum of the expected background events is calculated. This is done following Section 7.3.4.2. The reconstructed velocity of each event is folded and shifted to the most probable truth velocity of the event assuming they are actual signal events as described in Section 7.3.8.
Afterwards, the effective signal detection area for each velocity is calculated by the fraction of signal events selected from the total simulated signal events. The amount of predicted signal events is multiplied with the derived systematic efficiency described in Section 7.3.9.
As a last step, the average upper number of signal events compatible with the expected number of background events at each velocity is computed as described in Section 7.3.5. This is done for both simulated background and the remaining events from the blinded data, 2011-2017 set. The resulting diagram is depicted in Figure 7.18.

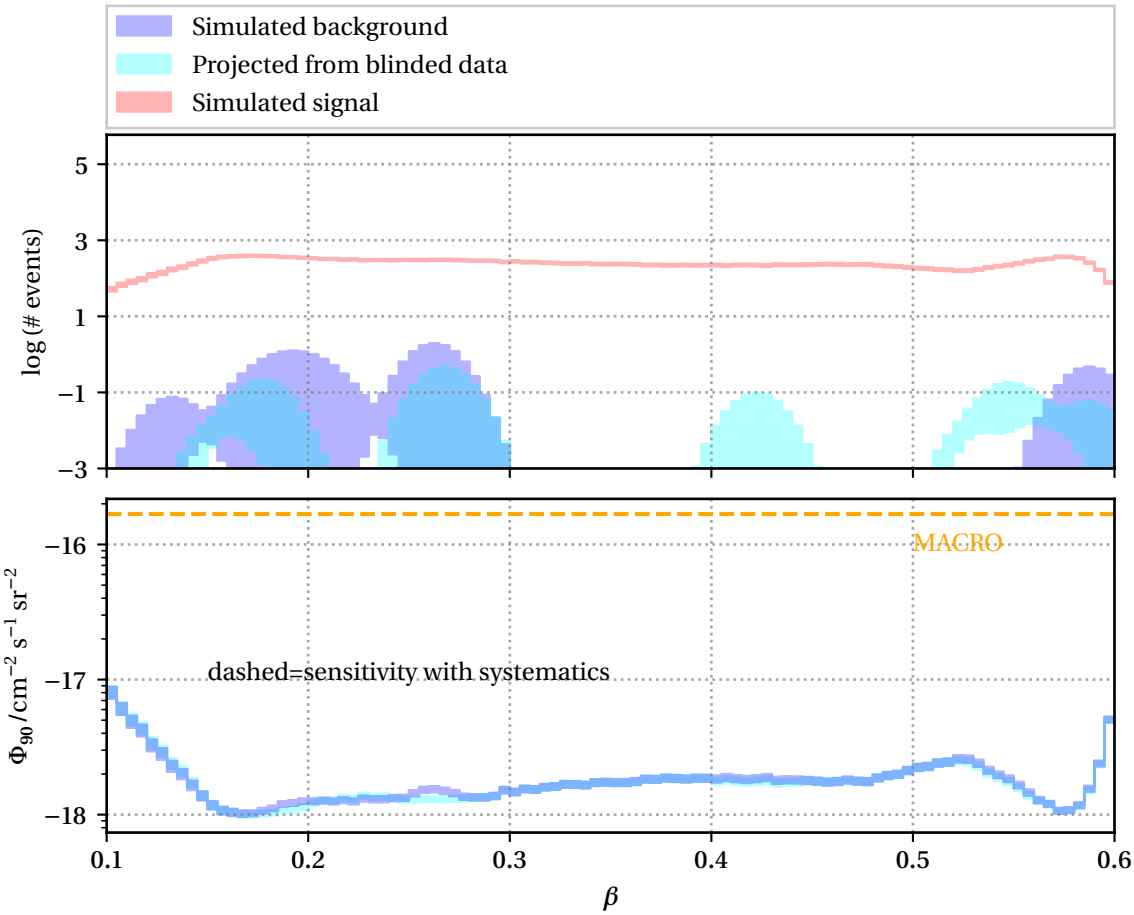


Figure 7.18: **Projected sensitivity before the unblinding process**
Two plots are presented. The horizontal axis is utilized to indicate the incident particles velocity in units of c in both plots. The vertical axis of the top plot is used to show the projected number of events after the final selection step for three different event groups: the signal events, the simulated background events, and the blinded data, 2011-2017 set. In the bottom plot, the resulting sensitivity based on both the blinded data, 2011-2017 set and the simulated backgrounds is shown. The same signal simulation is used for both calculations. The shaded regions indicate the uncertainty on the derived numbers. For the signal simulation, this consists of both the statistical uncertainty from statistics as well as the aforementioned systematic shift in signal detection efficiency described in Section 7.3.9. For the blinded data, 2011-2017 set and simulated background, only the statistical error is included. The dashed orange line marks the previous best flux limit in this velocity range published by the MACRO collaboration in reference [Amb+97].

Chapter 8

Results

The weight of evidence for an extraordinary claim must be proportioned to its strangeness.

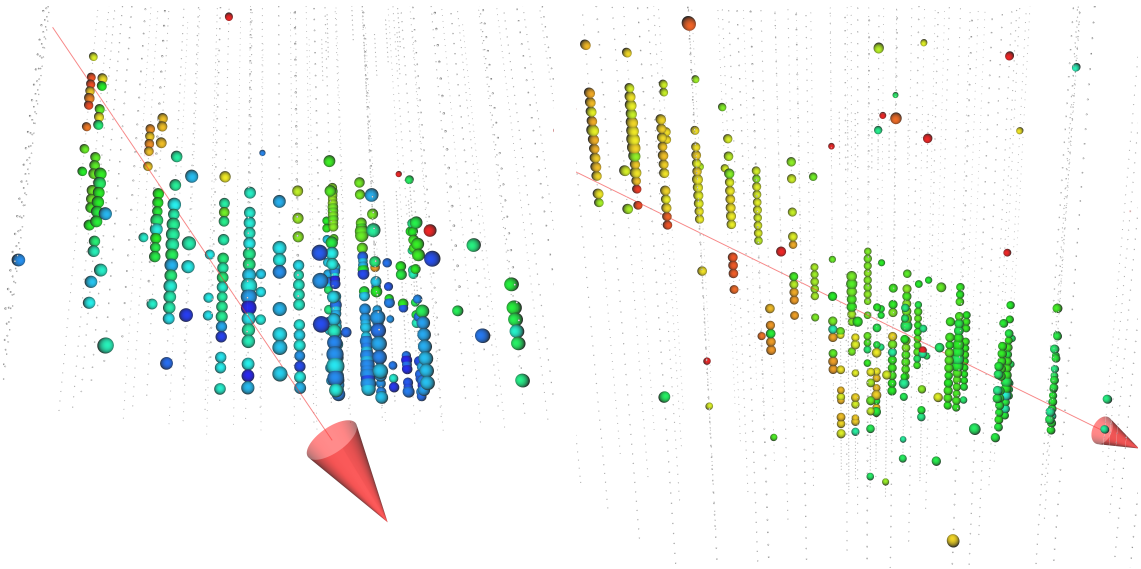
T. Flournoy [Flo99]

In this chapter, the results of the *unblinding* process are presented. Unblinding means that the analysis presented in Chapter 7 is applied to all available measured data of the IceCube Neutrino Observatory (**IceCube**) instead of applying it to a smaller and thus statistically blinded sample of the measured data. This process is implemented as a precaution against biasing the analysis against the measured data. All predictions postulated in Section 7.3.7 and Section 7.3.10 are compared against the actual results after the unblinding process. Tension between the predictions and the results after the unblinding process could indicate such biases.

In total, the analysis is applied to 2524.6 days of measured data taken at **IceCube** between May 13, 2011 and July 10, 2018. The number of remaining events is compared to the pre-unblinding projections, the remaining events themselves are presented, analyzed, and interpreted. An upper flux limit compatible with the observed remaining events is calculated and compared against previous searches.

8.1 Remaining Events

After unblinding, two measured events remained after all selection steps which is within the projected number of background events. As described in Section 7.3.7, the mean expected value before unblinding was 0 with an upper Gaussian approximated limit of 4 derived from the blinded, measured data set and 2 events with upper Gaussian approximated limit of 10 events derived from the background simulation set. Consequently, no detection of magnetic monopoles (**MM**s) can be claimed. These events will be called Hubert and Staller in this thesis. In any graphic, Hubert will be drawn in blue and Staller will be drawn in red. Both events are illustrated in Figure 8.1. The reconstructed kinematic attributes in the reference frame of the in-ice arrays of **IceCube** are shown in Table 8.1 with a translation of the arrival direction in galactic coordinates shown in Figure 8.3.



(a) *Hubert* (b) *Staller*

Figure 8.1: **Renders of Hubert and Staller**
*Renders of both Hubert and Staller are presented. Spheres indicate the location of **DOM**s, their size indicates the amount of light detected at each **DOM**. The color of the sphere indicates the timing of the first light being detected, from early in red to late in blue. Additionally, a red arrow is drawn indicating the reconstructed global **MPF16** track hypotheses as described in Section 7.1.2.*

Name	Velocity / c	Azimuth / °	Zenith / °	Date
Hubert	0.564 ± 0.009	38 ± 6	38 ± 1	September 19, 2013
Staller	0.54 ± 0.01	246 ± 5	62 ± 1	March 14, 2018

Table 8.1: **Reconstructed kinematic attributes at IceCube**
*A table containing the reconstructed kinematic attributes of Hubert and Staller is shown. Values are given in the coordinate system of **IceCube** as detected by the in-ice arrays. Additionally, the day the events took place is included.*

8.2 Event Origin

In order to determine the most probable origin of these two remaining events, the variables used in the final selection steps boosted decission treess (**BDT**s) are utilized. All variables which could not be reconstructed are removed and the remaining set of variables, 25 in total, are used to build multivariate kernel density estimators (**KDE**s). A multivariate **KDE** $f(\boldsymbol{x})$ is constructed by applying a kernel function to each sampled data point. A kernel function maps a given sampled data point to an assumed probability density around the location of the data point. In this work, a Gaussian kernel is used and the **KDE** becomes

$$f(\boldsymbol{x}) = \frac{1}{n}\sum_{j=1}^{n}(2\pi)^{-\frac{d}{2}}\,|\mathbf{H}|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{x}_j\right)^{\intercal}\mathbf{H}^{-1}\left(\boldsymbol{x}-\boldsymbol{x}_j\right)\right) \tag{8.1}$$

where d is the number of multivariate variables, n is the number of sampled data points, and $\mathbf{H}$ is the bandwidth matrix. For this section, $\mathbf{H}$ was chosen following Scott's rule, described in reference [Sco79], and thus

$$\mathbf{H}_{ii} = n^{-\frac{2}{d+4}} \cdot \vartheta_i \tag{8.2}$$

where ϑ_i is the variance of the i^{th} variable within the sampled data points.
As there is no event from the background simulation remaining after the final selection step, the multivariate **KDE**s have to be built based on events before the final event selection. While just removing the final selection step would be enough to supply ample statistics to calculate the multivariate **KDE**s, it introduces the risk of the multivariate **KDE**s not being representative of the events after the final selection step due to the high background rejection. An intermediate selection step between the two is used. Any event which has at least one **BDT** with a score greater than 0.996 is used to build the multivariate **KDE**s. With this intermediate selection step, there is still a reasonable high number of events left to build the multivariate **KDE**s from simulated background events ($\mathcal{O}(1000)$) while still being as close to the final event selection as is possible with the available simulated background datasets.

Five **KDE**s are build, one is based on simulated cosmic ray induced air showers, the other four are based on **MM** simulation. At the aforementioned intermediate selection stage, only $\mathcal{O}(10)$ simulated neutrinos are available with an integrated expected natural rate $\mathcal{O}(10^4)$ below the integrated natural rate of the remaining cosmic ray induced air shower simulation. Consequently, no **KDE** was built for neutrino based background simulation.
Two velocity ranges from the **MM** simulation are chosen as they are the closest to the reconstructed velocities of the remaining events, each with two luminescence light yields which were used for the **MM** simulation: 0.2 γ/MeV and 1 γ/MeV. The resulting probability density estimates are logarithmized and afterwards normalized between the five **KDE**s. The resulting relative probabilities are depicted in Figure 8.2 pointing to coincident muons as the most probable origin from the tested hypotheses for both remaining events. It is crucial to understand that these **KDE**s are only useful to compare the relative likelihood between the tested models and also do not make a statement about the absolute likelihood of the model. Any added model may change the relative probabilities.

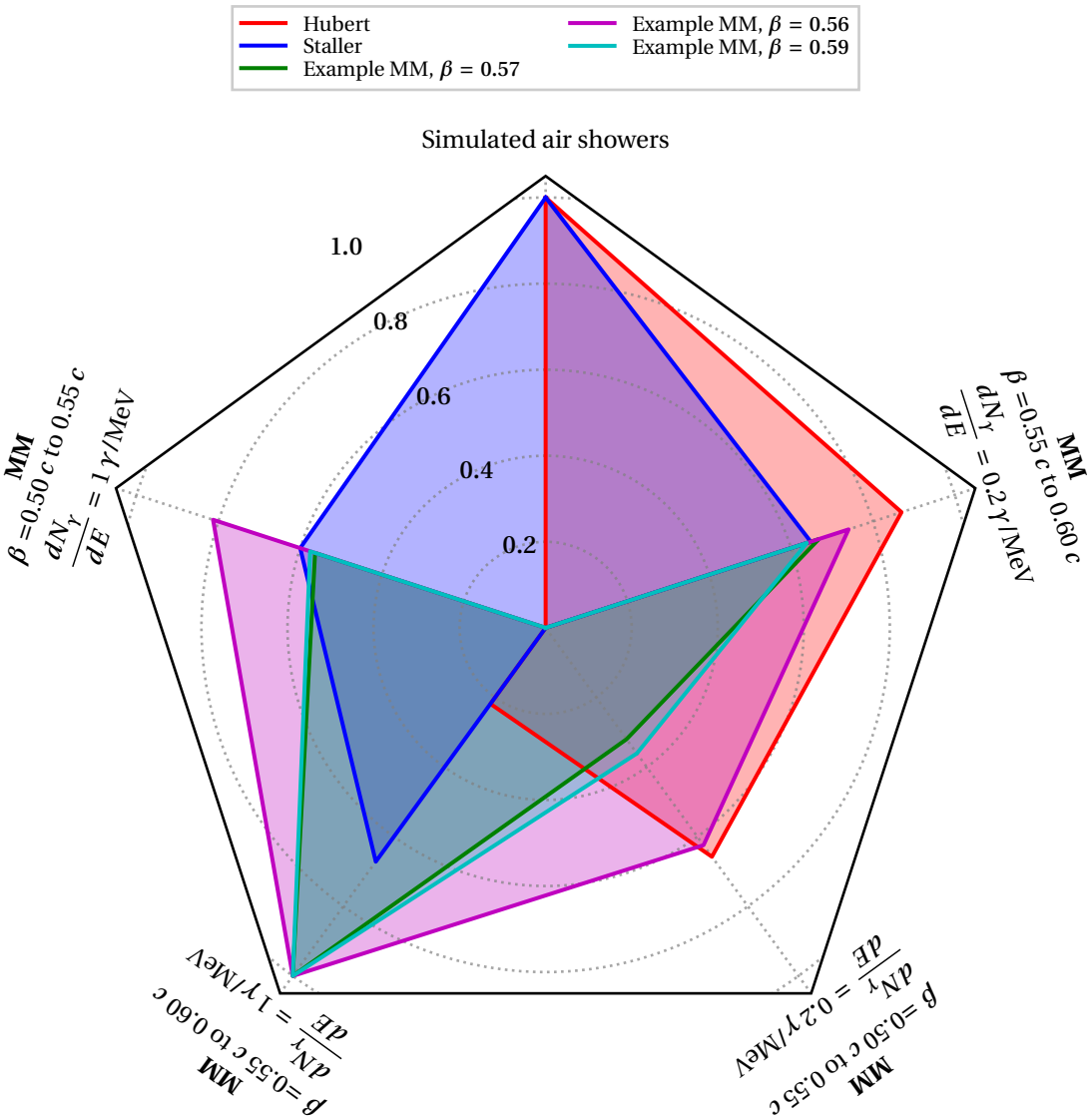


Figure 8.2: **KDEs for remaining events**
*A radar chart with five corners, each representing the relative probability derived from a single multivariant **KDE** is depicted. The probabilities of five events, Hubert, Staller, and three exemplary **MM** events, randomly selected from the **MM** β =0.55 c to 0.60 c, 1 γ/MeV subdataset, are shown. The simulated **MM**s all have their maximal probability towards the data set they originate from while Hubert and Staller have their maximal probability towards the **CORSIKA** based **KDE** indicating a non-**MM** origin of the two events.*

8.3 Internal Consistency of Remaining Events

Even if no **MM** detection can be claimed, if the remaining two events exhibit similar characteristics, this might indicate a common origin. This could be used as a hint for future analysis as either a region of interest or a specific **MM**-like background signature to be rejected.

8.3.1 Arrival Direction

A similar arrival direction for the remaining events might indicate a similar acceleration source, i. e., if both are aligning with the direction of the galactic magnetic field lines, this might hint at a low mass relativistic **MM** being accelerated following these field lines. In this analysis, reconstructions for the velocity and arrival direction of **MM** exist and can be tested immediately. Direction-wise, the remaining events are not aligned in galactic coordinates as illustrated in Figure 8.3.

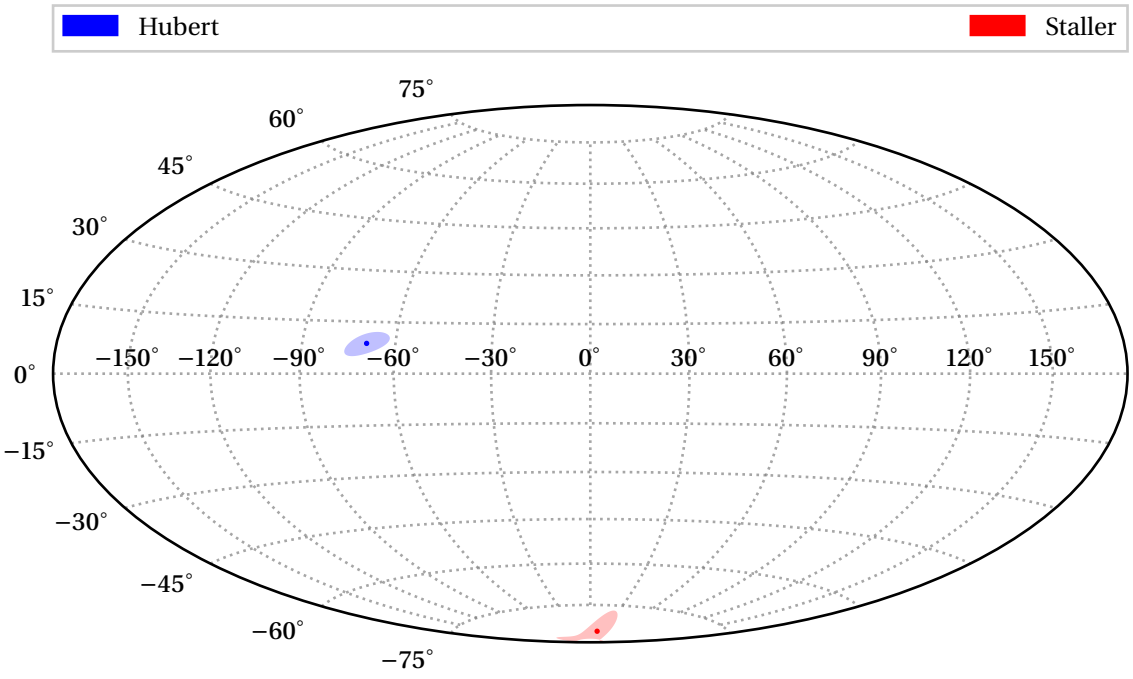


Figure 8.3: **Reconstructed galactic coordinates of Hubert and Staller**
*A map of the sky as seen on Earth in a Hammer projection, described in reference [Ham92], is presented. The Hammer projection is chosen as it is an equal-area map. Consequently, it does not introduce distortions to the confidence interval of the reconstructed directions based on their longitudinal position. The reconstructed arrival direction of Hubert and Staller, assuming they are **MM**, with the 90 % confidence interval of the reconstruction, has been drawn in blue and red. The arrival directions of the events do not coincide.*

8.3.2 Magnetic Charge

Assuming the remaining events are signal signatures of the same kind of **MM**, the magnetic charge and rest mass should be the same. As an indicator of the magnetic charge, the light production can be used as a metric. This analysis is designed to be insensitive to the light yield of an event due to the uncertainties in the luminescence light yield. Consequently, no metric to measure the light yields of events is included. Instead, an estimator for the light yield post-unblinding is constructed.
Specialized **MM** simulation is conducted, illuminating the in-ice arrays of **IceCube** with **MM**s. Instead of an isotropic illumination with different velocity ranges, the illumination is limited to the coincident interval of the reconstructed kinematic attributes of the remaining events.
Different luminescence light yields are used ranging from 0.2 γ/MeV to 2.4 γ/MeV to emulate a change in Dirac charge and/or luminescence light yield. Mirroring Section 8.1, multivariant **KDE**s are build based on this signal simulation. Enough statistic is available, so the final selection step can be conducted instead of the intermediate selection step needed to remedy low statistics as described in Section 8.1.
The resulting relative probabilities are depicted in Figure 8.4. While both events favor their arrival directions, which is expected, they do not favor the same light yield indicating that if both are **MM**s, they cannot exhibit the same Dirac charge.

8.3.3 Rest Mass

A direct indicator of the rest mass would be the change of velocity within the in-ice arrays of **IceCube**. This change is too small to be detected in this analysis. Consequently, an indirect indicator of the rest mass has to be used. Given a certain acceleration scenario for **MM**, the rest mass will influence the terminal velocity of an **MM** and thus the observed velocity of the **MM** at the in-ice arrays of **IceCube**. As both remaining events are down going events, shielding from the Earth and subsequent possible changes in velocity can be neglected.
Both remaining events exhibit very similar velocities, shown in Table 8.1, which might indicate similar rest masses. The reconstructed velocities are also close to the upper end of the low relativistic regime. It can be observed that there are two competing challenges for separating background events from signal events in this thesis. At low velocity, event selection is limited by event brightness. At high velocity, event selection is limited by mis-reconstructed coincident muons. While the similar velocity points to the same origin for both events, it cannot be determined if this origin is **MM**s or a geometric effect from coincident muons.

8.3.4 Discussion

It seems unlikely that these two events can be **MM**s as they do not favor the same luminescence light yield. Assuming one event was an **MM** and the other was a background event, the other consistency checks also disfavor this as they point either to the same origin or favor coincident muons outright. From the tested scenarios, it is most likely that these events are both coincident muons and result from a geometric effect. To remain conservative in flux limit calculations, the events will still be treated as if they were **MM**.

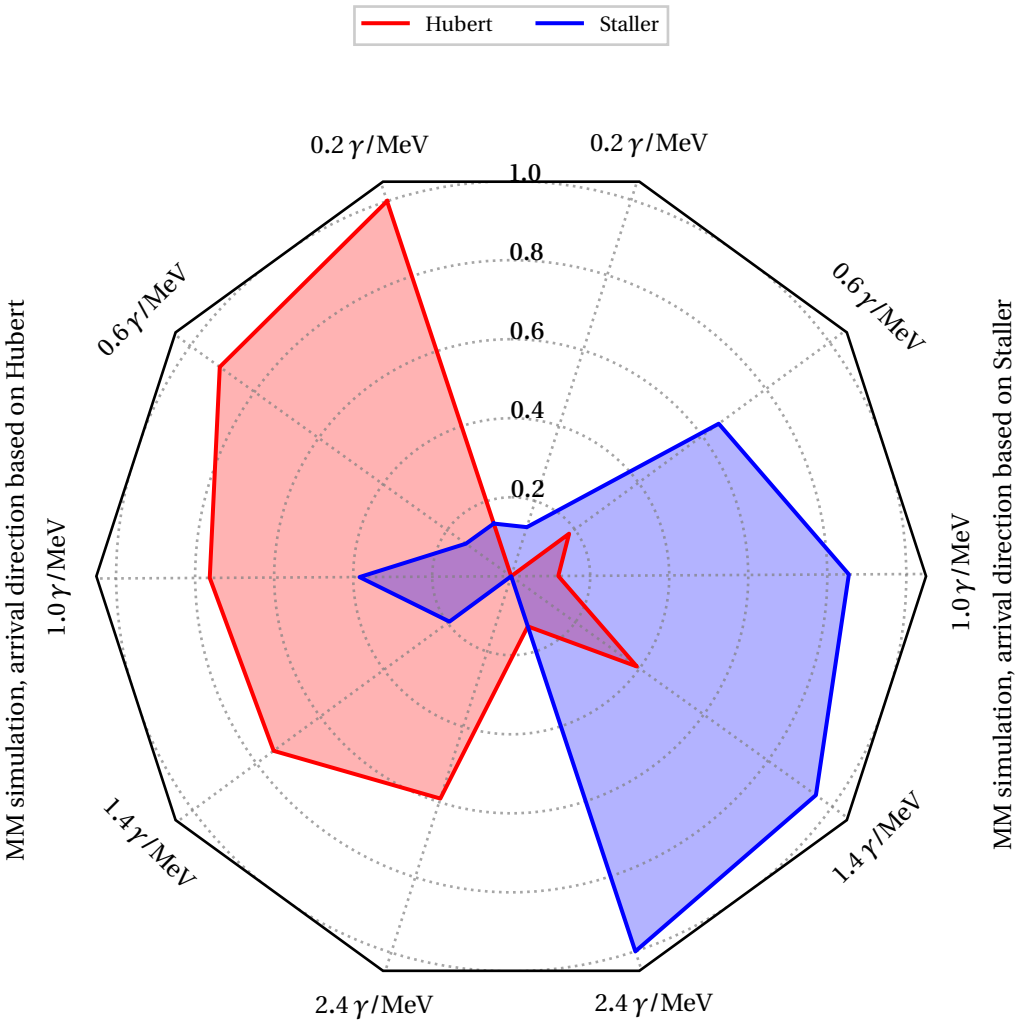


Figure 8.4: **KDEs for remaining event, light yield comparison**
*A radar chart with ten corners, each representing the relative probability derived from a single multivariant **KDE**, is depicted. The **KDE**s on the left are derived from specialized simulation with kinematic attributes taken from Hubert, the **KDE**s on the right are derived from specialized simulation with kinematic attributes taken from Staller.*

8.4 Low Relativistic Magnetic Monopole Flux Limit

No **MM** discovery can be claimed as no excess of candidate events after the unblinding of the measured data was observed.
Hints on the origin of the remaining events point toward coincident muons originating from cosmic ray induced air showers. This limits the possible flux of **MM**s in the low relativistic regime.
For the flux calculation, the two remaining events are treated as possible signal events to be conservative. The flux calculation is described in Section 7.3.10, but no projection based on bootstrap aggregating (**bagging**), illustrated in Section 7.3.4, is conducted as events are either classified as candidates or removed from the sample based on the mean of all **BDT**s. The calculated **MM** flux limit, shown in Figure 8.5, is mostly the same as the projected sensitivity depicted in Figure 7.18. Differences are due to the number of observed events. These reduce the flux limit in the velocity region of the observed remaining events in respect to the projected sensitivity by roughly a factor of two.
In contrast to previous reported flux limits in the low relativistic regime, the new flux limit is two orders of magnitude stronger and roughly as strong as the best flux limit in the low end of the mildly relativistic regime. Note that the flux limit is calculated assuming total shielding by the Earth for any **MM** with an incident angle below the horizon and thus can be applied to **MM** of higher magnetic charge than the Dirac charge.

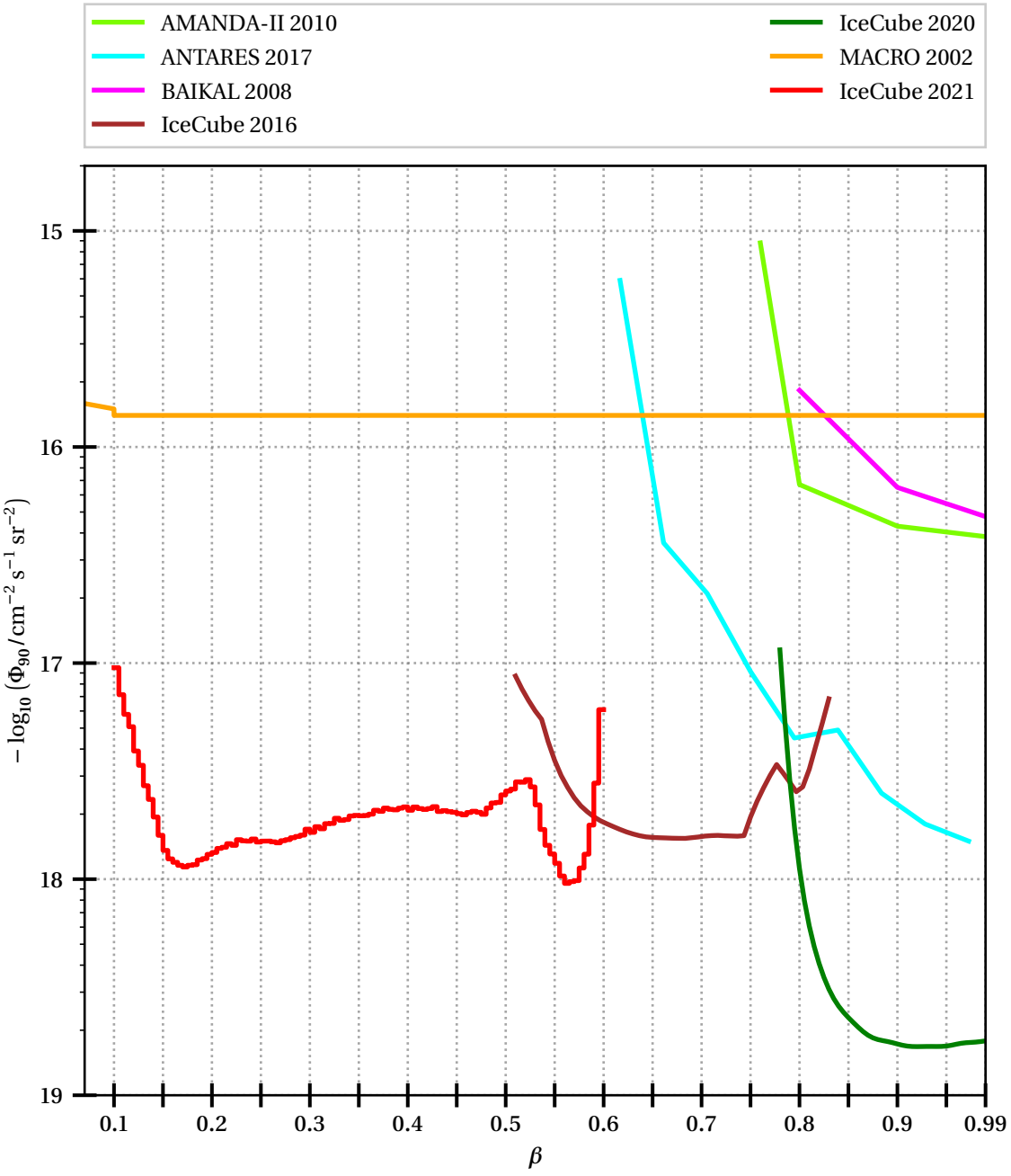


Figure 8.5: **Flux limit in contrast to previous searches**
*Current best flux limits for **MM**s from the low relativistic to the relativistic regime as a function of velocity are illustrated. The horizontal, linear axis represents the initial velocity of an **MM** between 0.07 c to 0.99 c. The vertical, logarithmic axis is the level of flux in units of $cm^{-2}\ s^{-1}\ sr^{-2}$. Previous flux limits are taken from different experiments and collaborations, sorted in the legend by alphabetical order. The corresponding references, in the same order, are [Abb+10b; Alb+17; Ayn+08; Aar+16; Bur20; Amb+02]. The flux limit calculated in Section 8.4 is drawn as a solid, red line. The data for this line can be found in Appendix F.*

Chapter 9

Conclusion and Outlook

If in this I have been tedious, it may be some excuse, I had not time to make it shorter.

W. Cowper [SC05]

A search for magnetic monopoles in the low relativistic regime at the IceCube Neutrino Observatory (**IceCube**) is described in this work. No such particle could be detected in the examined data. As a light production mechanism, luminescence was utilized for the first time at **IceCube**. The modeling of luminescence light is examined, implemented into the simulation framework of **IceCube**, and validated against theoretical expectation. Potential shielding of magnetic monopoles by the Earth is discussed, limiting the search for magnetic monopoles to down-going events instead of excluding higher charged magnetic monopoles.

The data acquisition system of **IceCube** is found to be able to detect magnetic monopoles in the low relativistic regime. Without any changes to the data acquisition system, about 25 % of isotropically passing magnetic monopoles can be detected. Selection criteria are presented allowing for up to 20 % of the detected magnetic monopoles to be isolated while suppressing $\mathcal{O}(1\,\mathrm{kHz})$ of background events recorded by **IceCube**'s data acquisition system. 2524.6 days of recorded data are examined for magnetic monopoles. Two events passed all selection criteria which is within the background estimate. After examination both visually and by statistical methods, these events are found to be most likely background-like. The first exclusion limit on the flux of magnetic monopoles in the low relativistic regime at **IceCube** is derived superseding previous best limits by up to two orders of magnitude.

To surpass this exclusion limit in the near future, an increase in the statistics of background simulation by an estimated $\mathcal{O}(10)$ to $\mathcal{O}(100)$ will be required. This would alleviate the statistically limited predictions on the expected background event distributions at the final selection stage of this analysis. Additionally, this would facilitate the usage of more advanced algorithms to be deployed potentially resulting in a higher percentage of magnetic monopoles retained while suppressing similar orders of magnitude of background. Neural Networks are one of such more advanced algorithms. While they have been explored as a promising, albeit unproved algorithm for the final selection step for this thesis in prestudies, they could not be utilized for this thesis due to the lack of available and reliable background simulation.
Seven seasons of recorded data by **IceCube** have been used in this thesis. Consequently, any superseding of the aforementioned exclusion limit by an order of magnitude solely due to increased statistics is infeasible in the near future. Instead, a focus should be placed on the data acquisition system of the **IceCube**-Gen2 upgrade with its current design described in reference [Aar+21]. The upgrade is designed to increase the instrumented volume of ice by about a factor of 10. This comes at the cost of instrumentation density which could pose a challenge for relatively dim events like Dirac charged magnetic monopoles in the low relativistic regime. Light sensors with improved sensitivity could be utilized to alleviated this. Proposed high efficiency air shower vetoes could suppress the dominant class of background events observed in this analysis.

Appendix A

Improved LineFit

Many events of interest recorded at the IceCube Neutrino Observatory (**IceCube**) originate from particles traversing the in-ice arrays of **IceCube** at a constant velocity. Their trajectory can be described by

$$\boldsymbol{R}(t) = \boldsymbol{V} \cdot (t - T) + \boldsymbol{R}(T) \tag{A.1}$$

where $\boldsymbol{R}(t)$ is the particle position at time t, $\boldsymbol{V}$ is a constant velocity, and T is an arbitrary reference point in time. As a first guess, analytical, fast algorithm for direction, timing, and velocity of such events, the improved LineFit (**iLF**) algorithm was designed by the **IceCube** collaboration (**ICC**) and published in reference [Aar+14b]. Below, the algorithm is summarized to facilitate the understanding for this thesis.

An **iLF** is applied to a set of hits where $\boldsymbol{r_i}$ is the position and t_i is the time of the i^{th} hit. As a first step, any *late* hits are removed. Here, late means that the hit occurred without another hit occurring within a configurable radius d and time window of t. In this thesis, the default values described in the aforementioned reference are used namely $d = 156\,\text{m}$ and $t = 778\,\text{ns}$.

Afterwards, a *Hubert fit*, described in reference [BV04] is calculated. A Hubert fit is a common multiple linear regression with an added penalty function, the *Hubert penalty*, to reduce the impact of outliers. Let $\boldsymbol{\Delta r_i} = \boldsymbol{r_i} - \boldsymbol{R'}$ and $\Delta t_i = t_i - T'$ where T', V', and $\boldsymbol{R'}$ follow the naming convention of Equation A.1. Then,

$$p_i = \left| \boldsymbol{V'} \cdot \Delta t_i - \boldsymbol{\Delta r_i} \right| \tag{A.2}$$

is the residual of the i^{th} hit and the reconstructed trajectory. For residuals smaller than a configurable threshold μ, the common linear regression penalty

$$P_i = p_i^2 \tag{A.3}$$

is used. Residuals bigger than the threshold are assigned a penalty of

$$P_i = \mu \cdot (2 \cdot p_i - \mu). \tag{A.4}$$

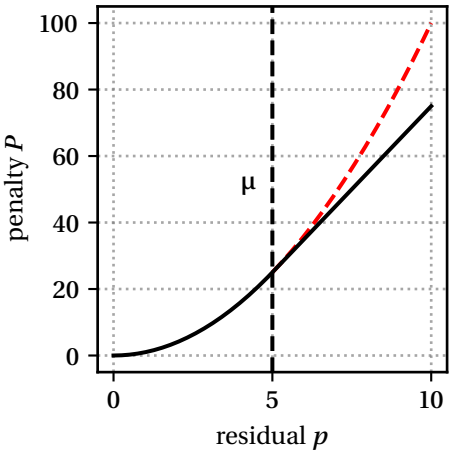


Figure A.1: **Penalty P as a function of residual p** *The relationship between the value of a residual p and the value of the applied penalty P is drawn as a black solid line for an exemplary case of* $\mu = 5$. *To contrast the Hubert penalty, a regular penalty of* $P = p^2$ *is drawn as a red, dashed line.*

This is called the Hubert penalty. In Figure A.1, the relation between the penalty and the residual is drawn. $\boldsymbol{R}'$, T', and $\boldsymbol{V}'$ are the values which minimize the sum of all penalties $\sum P_i$.

Once $\boldsymbol{R}'$, T', and $\boldsymbol{V}'$ are determined, all hits with a residual greater than μ are removed. Here, the default values described in the aforementioned reference of $\mu = 116\,\mathrm{m}$ is used.

If the number of remaining hits is lower than a configured number of hits, no further processing is applied and no **iLF** can be derived for the given hit list. The **iLF** is said to have not converged. Here, the default value of minimal required hits, 2, is used.

As a last step, a linear regression is applied to the remaining hits. Following the naming convention in Equation A.1, the analytical solution is

$$\boldsymbol{R} = \langle \boldsymbol{r} \rangle - \boldsymbol{V} \langle t \rangle = \frac{\langle \boldsymbol{r} \cdot t \rangle - \langle \boldsymbol{r} \rangle \cdot \langle t \rangle}{\langle t^2 \rangle - \langle t \rangle^2} \tag{A.5}$$

where $\langle \ldots \rangle$ is the mean value operator. $|\boldsymbol{V}|$ is the reconstructed velocity, $\frac{\boldsymbol{V}}{|\boldsymbol{V}|}$ the reconstructed direction.

Appendix B

Earth Shielding for Intermediate Magnetic Charges

Additional graphics, calculated for intermediate magnetic charges between $N_{g_D} = 1$ and $N_{g_D} = 9$ which have already been presented in Section 3.8, are depicted. Each figure is calculated assuming a different magnetic charge, each a multiple of the Dirac charge g_{MM}. The horizontal axes are indicative of the rest mass of an incident **MM** in logarithmic presentation while the vertical axes are used to indicate different zenith angles. An initial velocity as described in Figure 3.5 is assumed. The **MM** is propagated through the Earth following the density and composition model described in Figure 3.7. Its arrival velocity at the in-ice arrays of **IceCube** is indicated by color corresponding to a velocity as noted at the right hand color bar.

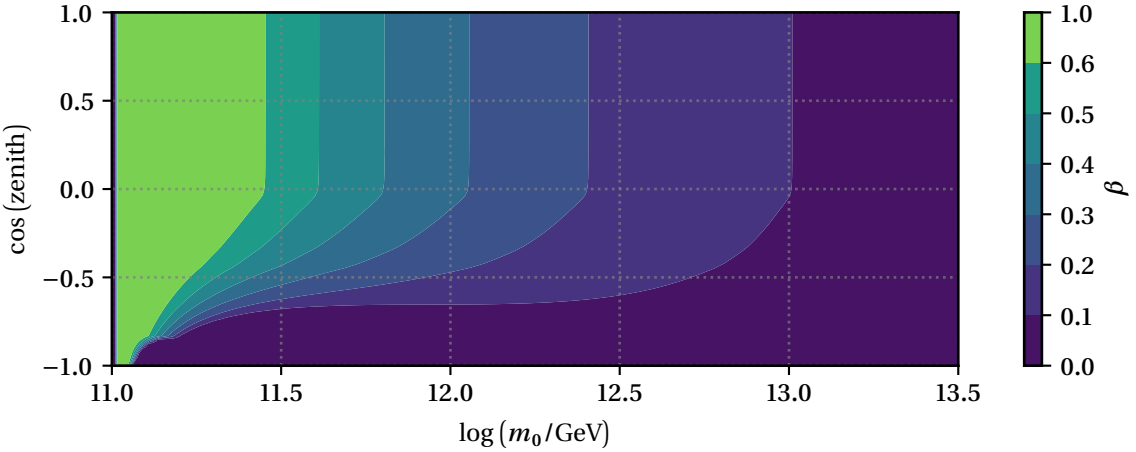


Figure B.1: **IMMs velocity at IceCube for $N_{g_D} = 2$**

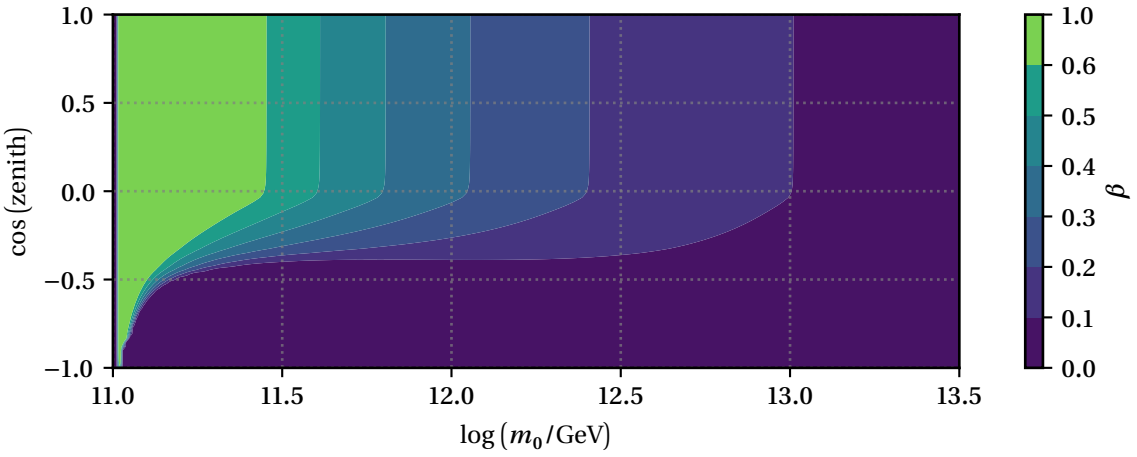


Figure B.2: **IMMs velocity at IceCube for $N_{g_D} = 3$**

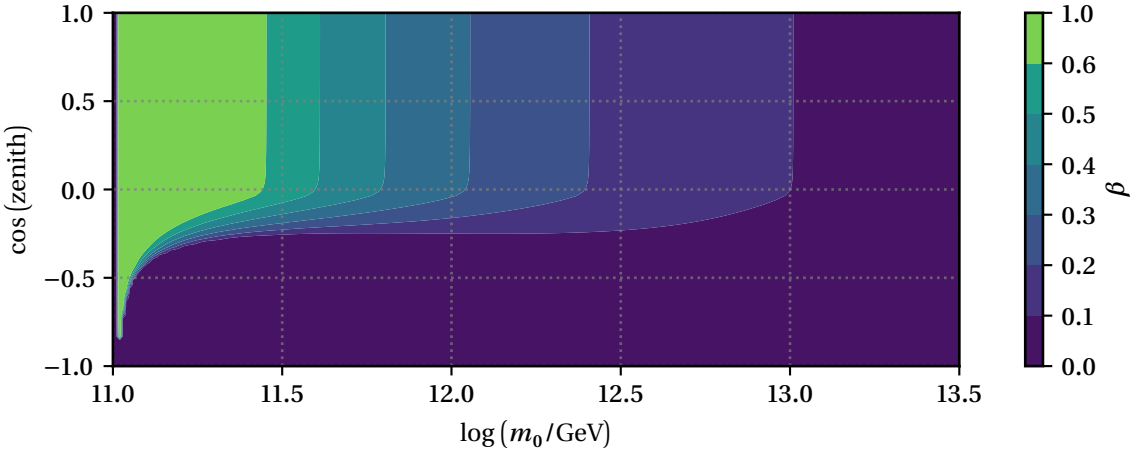


Figure B.3: **IMMs velocity at IceCube for $N_{g_D} = 4$**

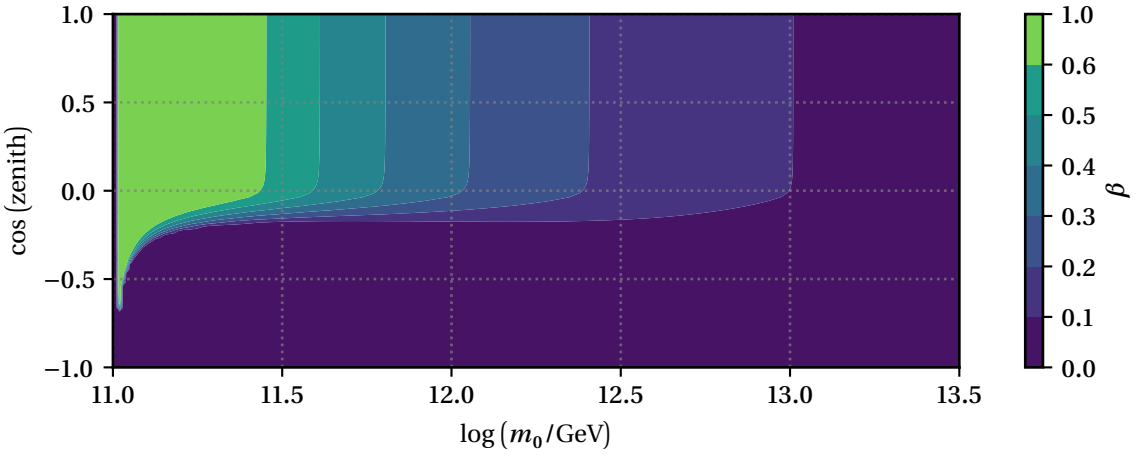


Figure B.4: **IMMs velocity at IceCube for $N_{g_D} = 5$**

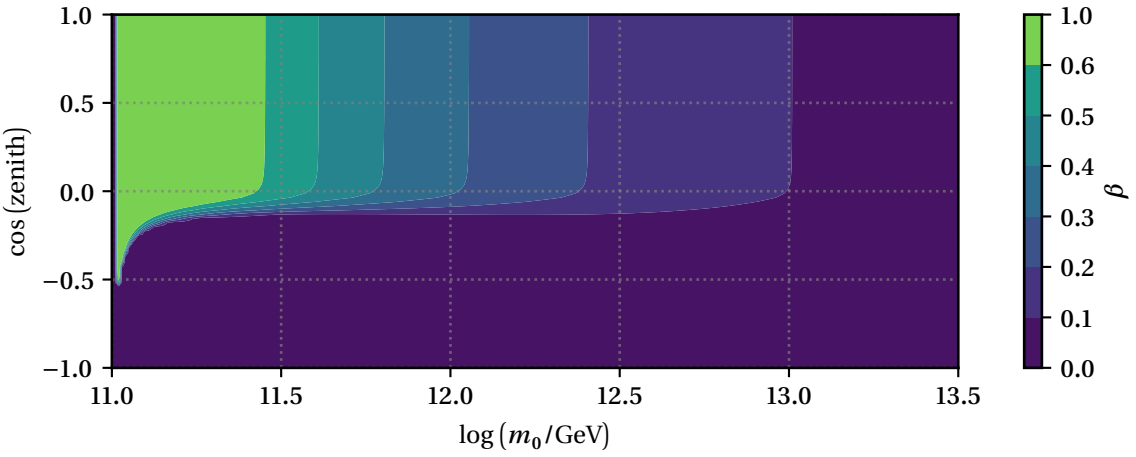


Figure B.5: **IMMs velocity at IceCube for $N_{g_D} = 6$**

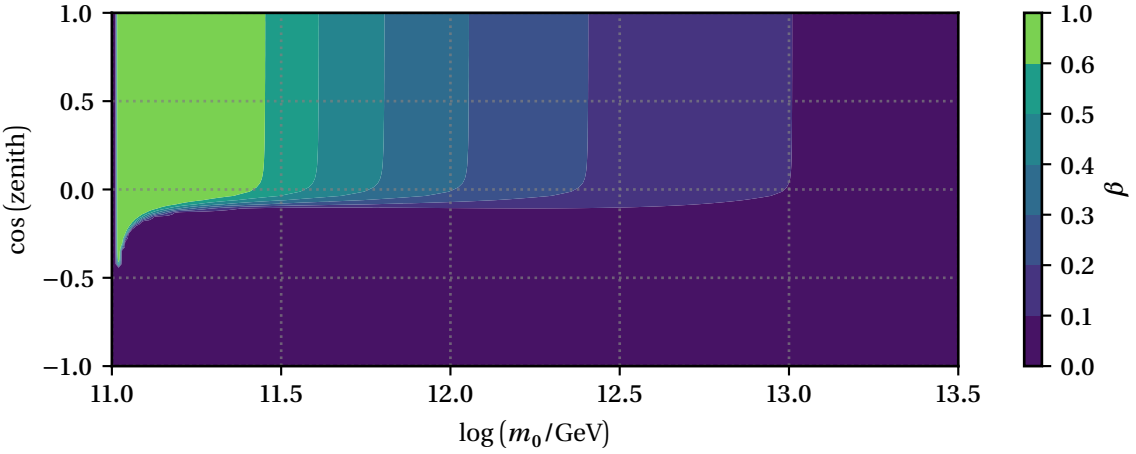


Figure B.6: **IMMs velocity at IceCube for $N_{g_D} = 7$**

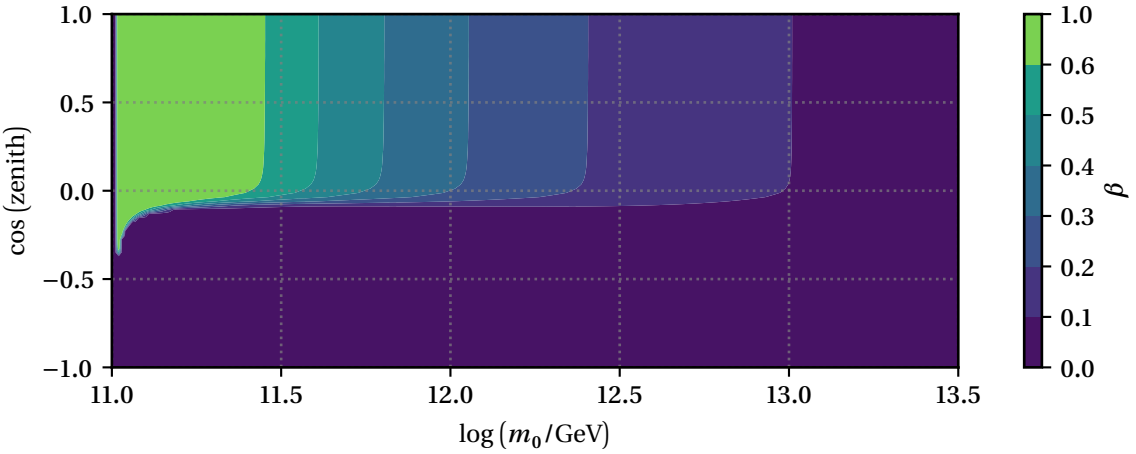


Figure B.7: **IMMs velocity at IceCube for $N_{g_D} = 8$**

Appendix C

Individual Systematic Effects

While in Section 7.3.9, the total systematic variation due to all examined effects is presented, here, the individual behavior of singular effects are depicted. Effects expected due to the nature of **IceCube** are systematic shifts in the light detection efficiencies of the deployed Digital Optical Modules (**DOM**s) and changes in the angular acceptance of light for each **DOM** due to melted and refrozen ice around the **DOM**s during deployment. Potential mismatches between the utilized ice model and the real existing ice are taken into account by varying the expected scattering and absorption of photons inside the bulk ice of the in-ice arrays of **IceCube**. These effects are also investigated in similar analyses performed on **IceCube** data. Unique to this work, an additional systematic effect is investigated: a systematic shift of the luminescence light yield. While this effect is virtually the same as lowering the light detection efficiencies of each **DOM**s in the low relativistic regime, this is not true once indirect Cherenkov light is produced by an **MM**. All effects are treated as independent of each other and are varied within the expected uncertainties of the parameters.
In each graphic, the horizontal axis represents the velocity of a simulated **MM**. On the vertical axis, the observed ratio of natural rates between a systematically shifted dataset and an unshifted dataset is illustrated.

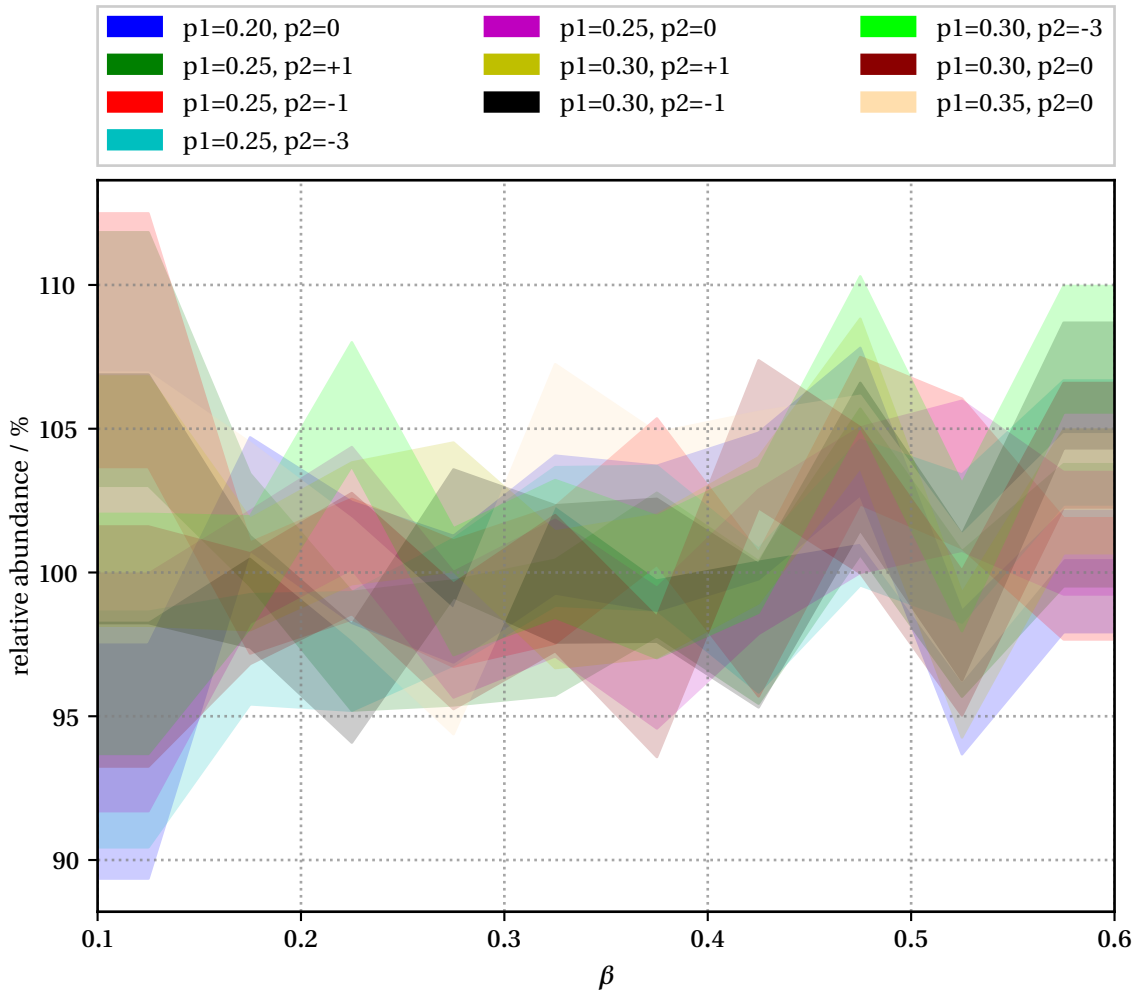


Figure C.1: **Deployment systematic efficiency**
*Due to the deployment of the **DOM**s of the in-ice arrays of **IceCube**, holes had to be melted into the original or bulk ice of Antarctica. After the melted water froze around the positioned **DOM**s, the properties of the hole ice differed from the properties of the bulk ice. This is taken into account by altering the acceptance of photons for each **DOM** depending on the arrival angle η with $\cos(\eta) = 1$ being straight up. The acceptance is then modeled by $0.34 \cdot \left(1 + 1.5\cos\eta - \cos^3\frac{\eta}{2}\right) + p1 \cdot \cos\eta \cdot \left(\cos^2\eta - 1\right)^3 + p2 \cdot \exp\left(10\left(\cos\eta - 1.2\right)\right)$ where p1 and p2 are fitted parameters from specialized studies within the **ICC** as described in reference [Chi17]. No dependence on these parameters is expected, neither for the analysis overall nor velocity wise. For very slow velocities, some dependence is observed, most likely due to the low brightness of events. Otherwise, no dependence is observed.*

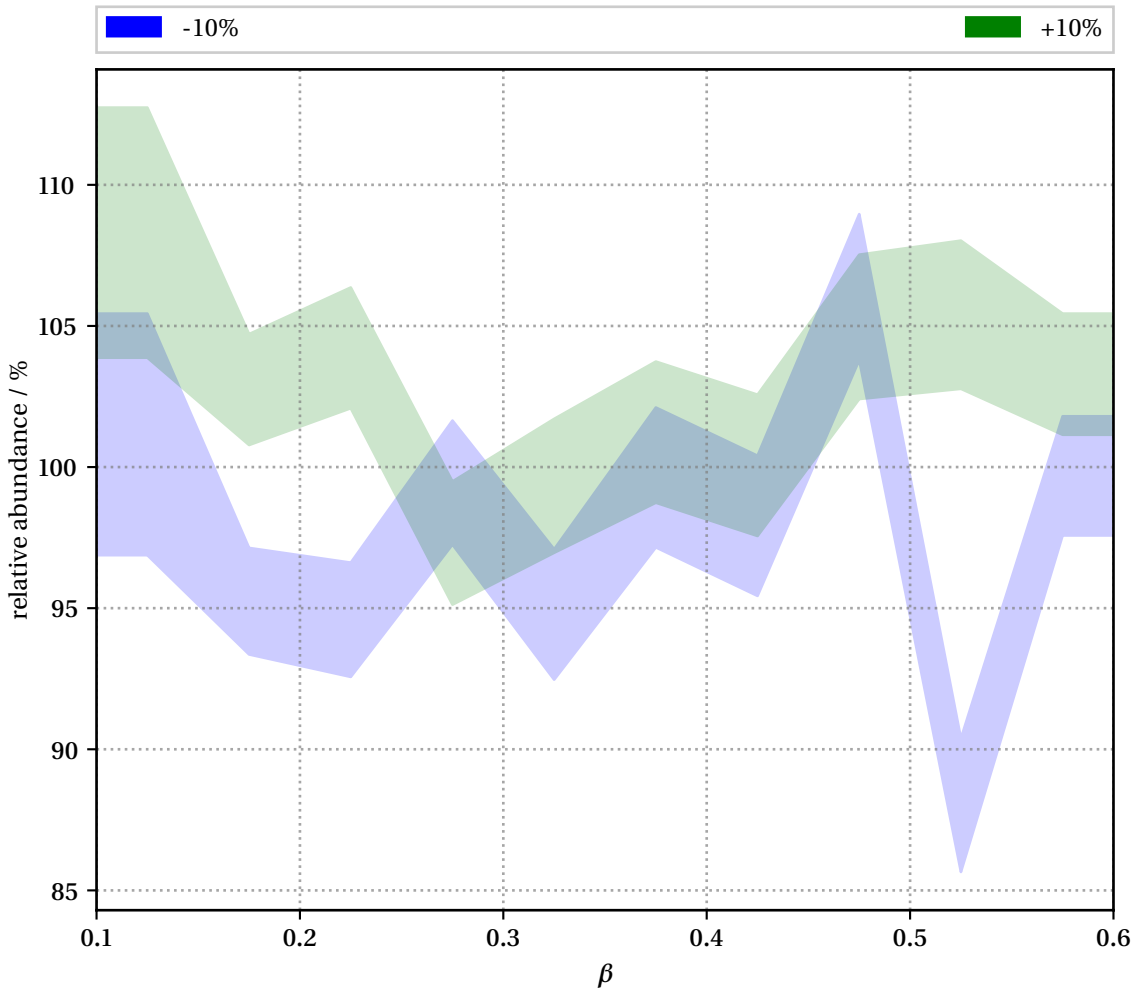


Figure C.2: **DOM efficiency systematic efficiency**
*The detected amount of light by the in-ice arrays of **IceCube** depends on the light detection efficiency of each **DOM**. A relative shift of efficiency between −10 % to 10 % is tested for in line with the current uncertainty on the absolute calibration conducted by the **ICC** described in reference [Aar+17]. For this analysis, it is expected to have a higher signal retention for higher **DOM** efficiencies as event detection is limited by the amount of light in the first place. As at lower velocities, even less is produced, a stronger effect could be assumed. In the figure, the relative abundance of signal events as a function of the velocity is depicted. The shaded region indicates the statistical contour around the mean values drawn as lines. Ignoring the velocity dependence, a roughly ±2 % point shift in the relative abundance is observed. If there is any velocity dependent effect is inconclusive due to the statistical limitations of the simulated systematic sets.*

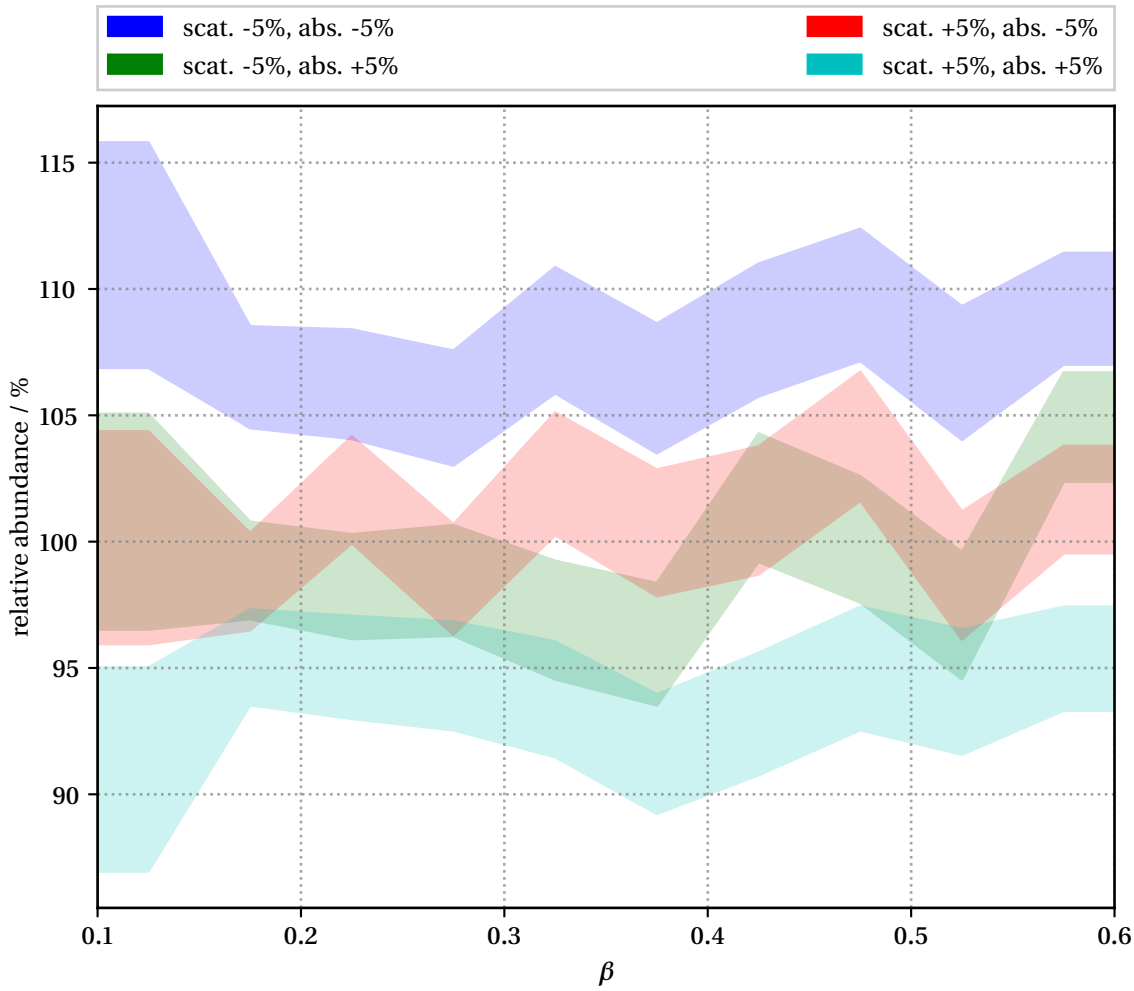


Figure C.3: **Ice property systematic efficiency**
*The detected amount of light by the in-ice arrays of **IceCube** depends on light transmission attributes of the detection medium. This can be described by two parameters: the scattering of photons inside the ice and the absorption of photons inside the ice. These two parameters are correlated and thus systematic shifts of one needs to be tested with a systematic shift of the other. In line with the current estimates on the systematic shift of these parameters described in reference [Aar+13], a systematic shift of ±5 % for each parameter is regarded. It is expected that an increase in scattering and absorption should decrease the number of recorded signal events while decreasing scattering and absorption should increase the number of recorded signal events. Besides a general trend of the detection of very slow events (0.1 c) to be more sensitive to the total amount of recorded light and thus scattering and absorption, no influence of velocity is expected. A priori, no prediction on the severeness of scattering over absorption has been made. The observed relative shift of signal detection efficiency as a function of velocity is depicted in the figure above. As expected, the scenario with higher scattering and absorption has the biggest impact on the number of recorded signal events. Roughly 93 % of all baseline events are still recorded. No major dependency on the velocity is observed.*

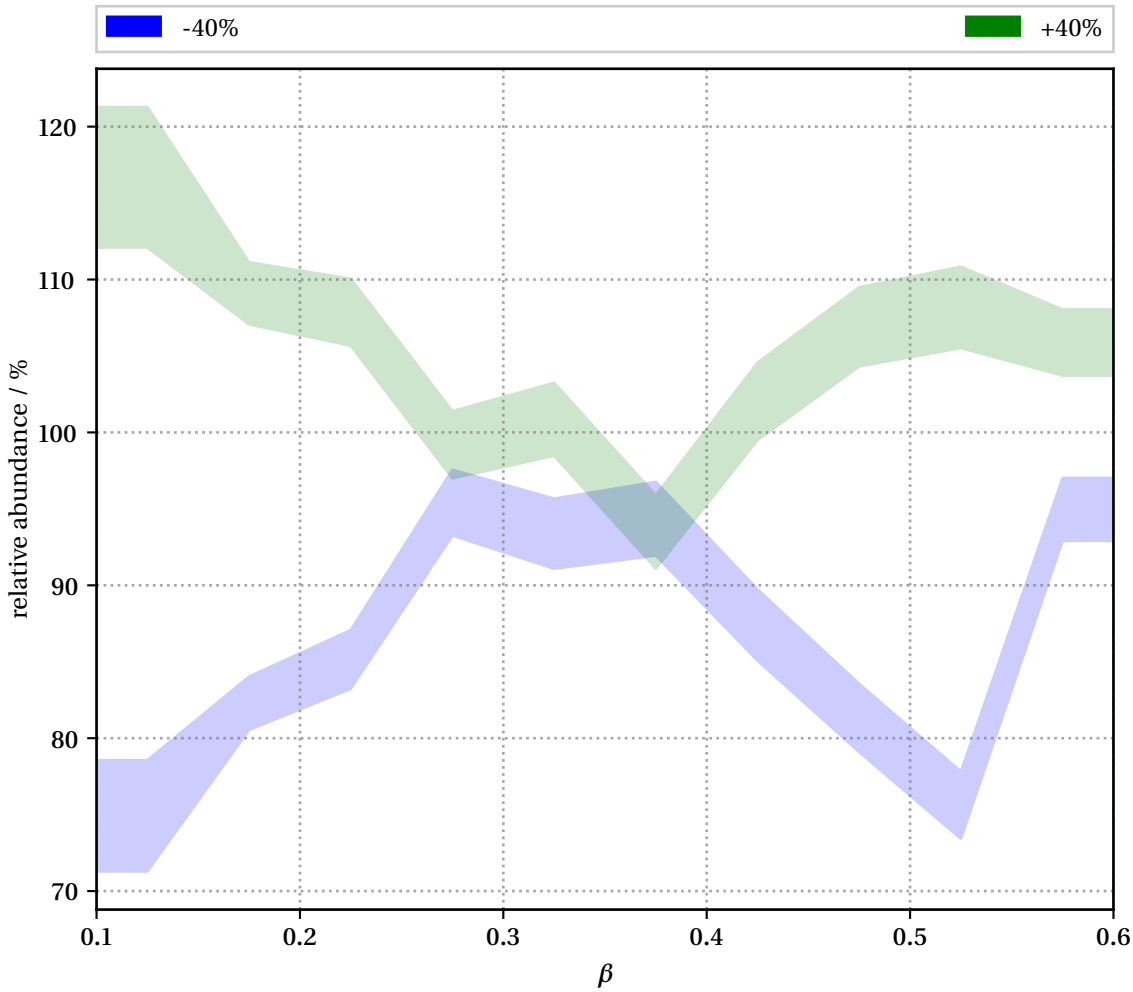


Figure C.4: **Luminescence light yield systematic efficiency**
*The detected amount of light by the in-ice arrays of **IceCube** depends on the light induced by **MM**s. In this thesis, a previously unused light production mechanism, luminescence of water based ice, described in Section 5.3, is utilized for the detection of **MM**s. In-situ measurements of the luminescence light yield of the Antarctic ice have been described in reference [Pol19]. Following reference [Pol20], a uncertainty of ±40 % for this measurement is estimated. As this is by far the strongest systematic effect, a strong impact on this analysis is expected overall. Additionally, a strong velocity dependence is predicted as at 0.5 c a second, luminescence light yield independent light production channel becomes effective: indirect Cherenkov light, described in Section 5.2. The observed relative shift of detection efficiency of signal efficiency as a function of velocity is depicted in the figure above. As expected, a strong influence is seen at low velocities where a 40 % reduction in the produced light results in only 80 % of the baseline events to be detected while a 40 % increase leads to 110 % of the baseline events to pass through all selection steps. With increasing velocity, this effect becomes less pronounced until the region between 0.4 c to 0.55 c where a strong impact can be seen. This effect becomes less pronounced for weaker cuts in the final selection step indicating a reliance of the set of **BDT**s on the proper amount of light in this region for unknown reasons.*

Appendix D

Input Variables for the Last Analysis Selection Step

In this section, all input variables utilized in the machine learning (**ML**) based selection step described in Section 7.3 are illustrated. A short description of each variable is included in all figures. For each variable, two vertically stacked plots are depicted. On the shared horizontal axis, the variable in question is shown on a linear scale. The upper vertical axis is used to display the natural rate $\Re$ on a logarithmic scale while the lower vertical axis is used as an indicator of the ratio between the simulated background and the measured blinded data. The simulated background is the sum of the cosmic ray induced air showers (2012) set, the atmospheric ν_μ set, and the atmospheric ν_e set. A description of each variable and the exact equation to calculate the variable can be found in Section 7.3.2. To keep repetition of graph description to a minimum, the information presented in this section will be omitted from the graph description. All error bars are 1 σ error bars. In the ideal case, a flat line is expected in the second plots which indicates a perfect match between background simulation and measured events as the measured data at this selection stage is dominated by background events.

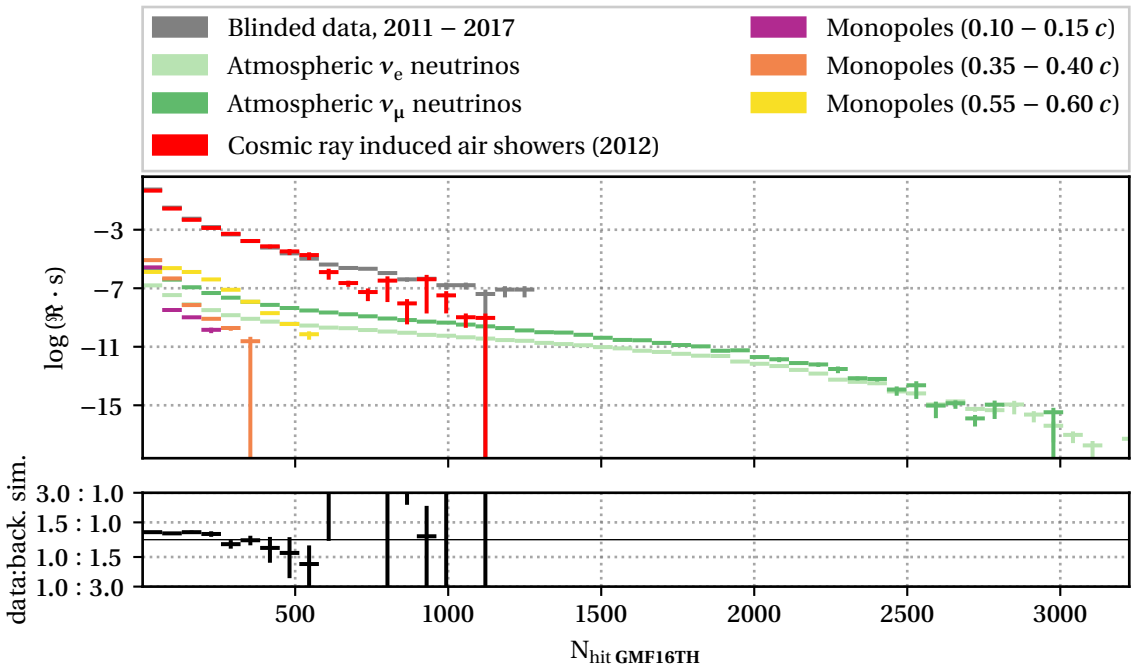


Figure D.1: **Input variable: $N_{hit\ GMF16TH}$**
*Number of hits utilized in the global **MPF16** track hypotheses (**GMF16TH**).*

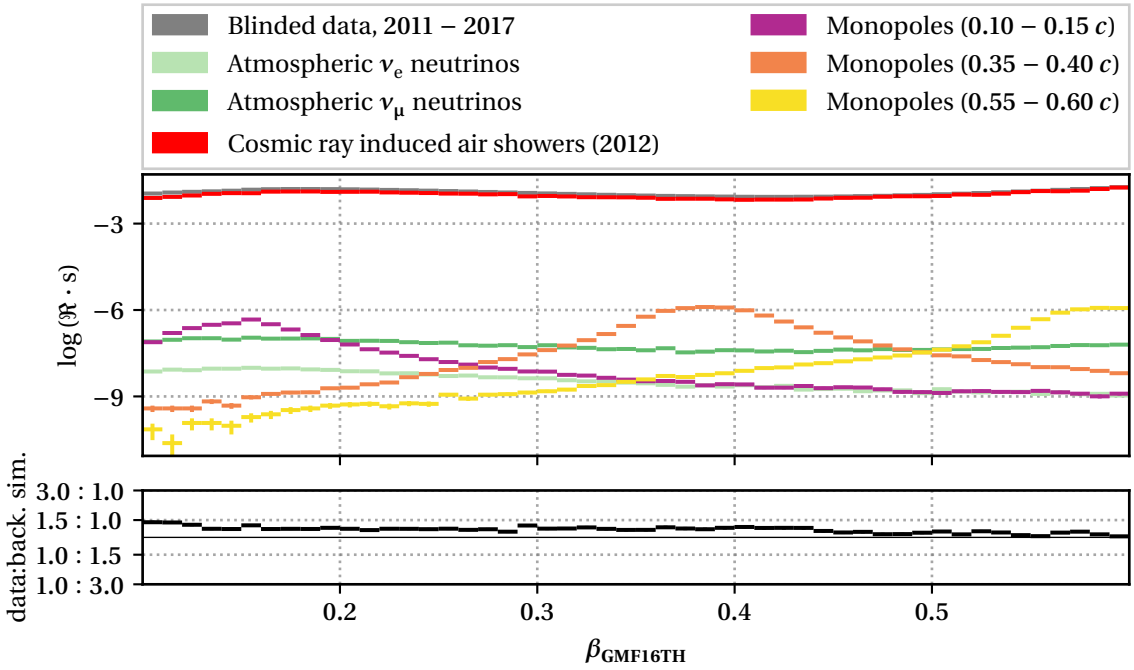


Figure D.2: **Input variable: $\beta_{GMF16TH}$**
*Reconstructed velocity of the global **MPF16** track hypotheses (**GMF16TH**).*

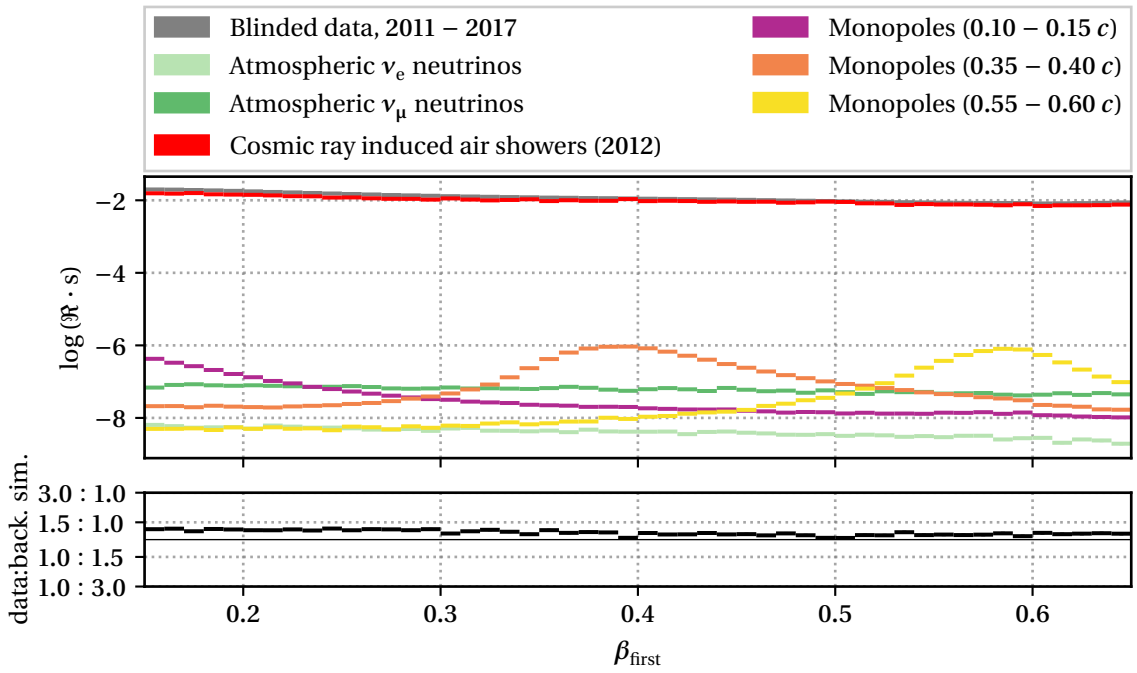


Figure D.3: **Input variable: β_{first}**
Reconstructed velocity based on the temporal first half of hits of the event.

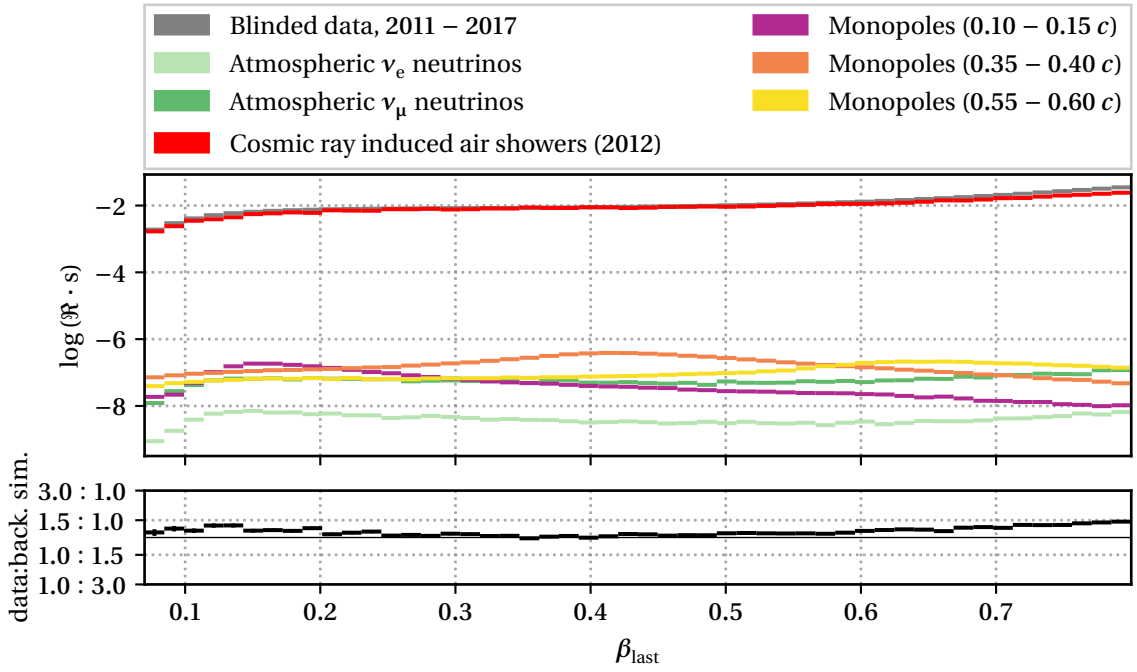


Figure D.4: **Input variable: β_{last}**
Reconstructed velocity based on the temporal second half of hits of the event.

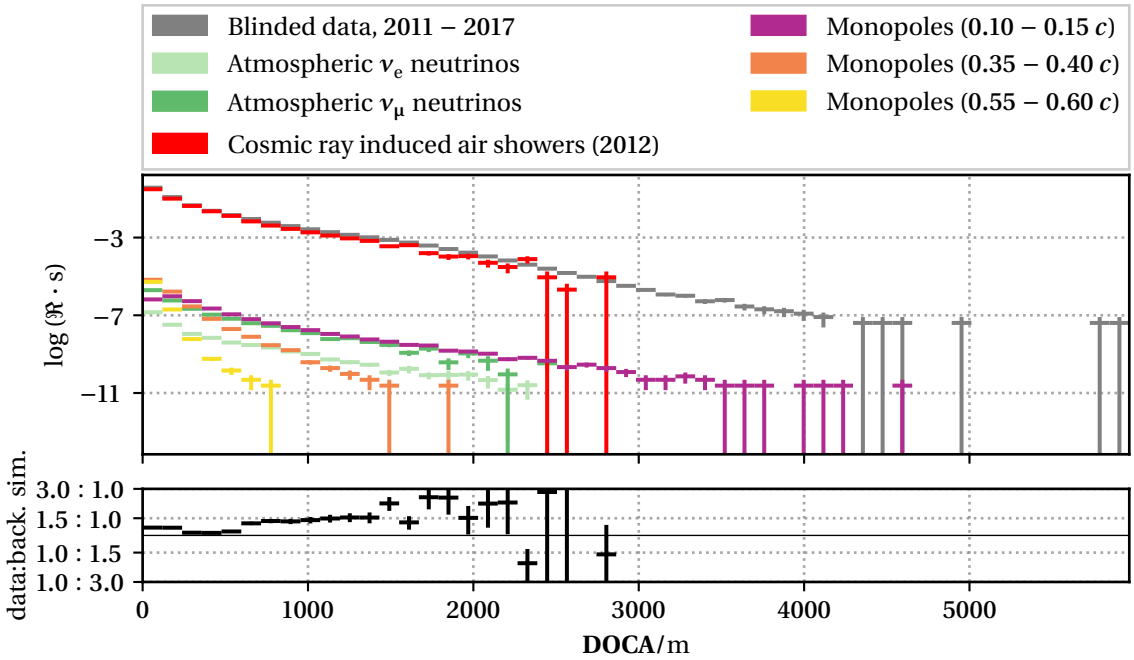


Figure D.5: **Input variable: DOCA**
*Distance of closest approach (**DOCA**) between the reconstructed particle paths of the first and second half of temporal split hits.*

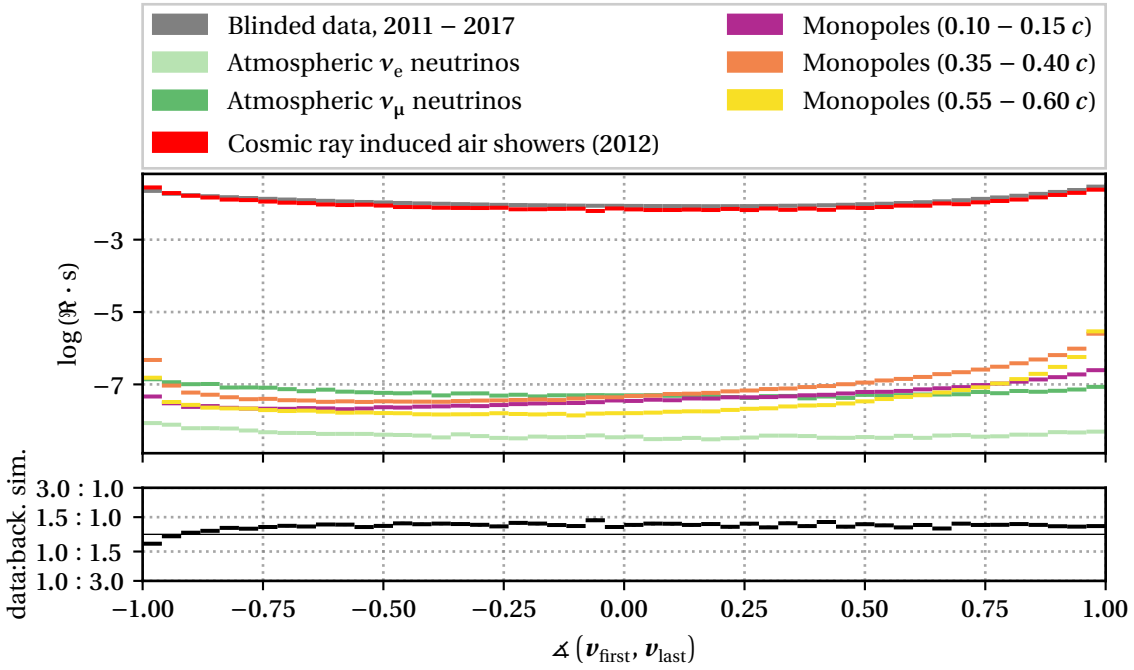


Figure D.6: **Input variable:** $\measuredangle(\boldsymbol{v}_{\text{first}}, \boldsymbol{v}_{\text{last}})$
Angle between the between the reconstructed particle paths of the first and second half of temporal split hits.

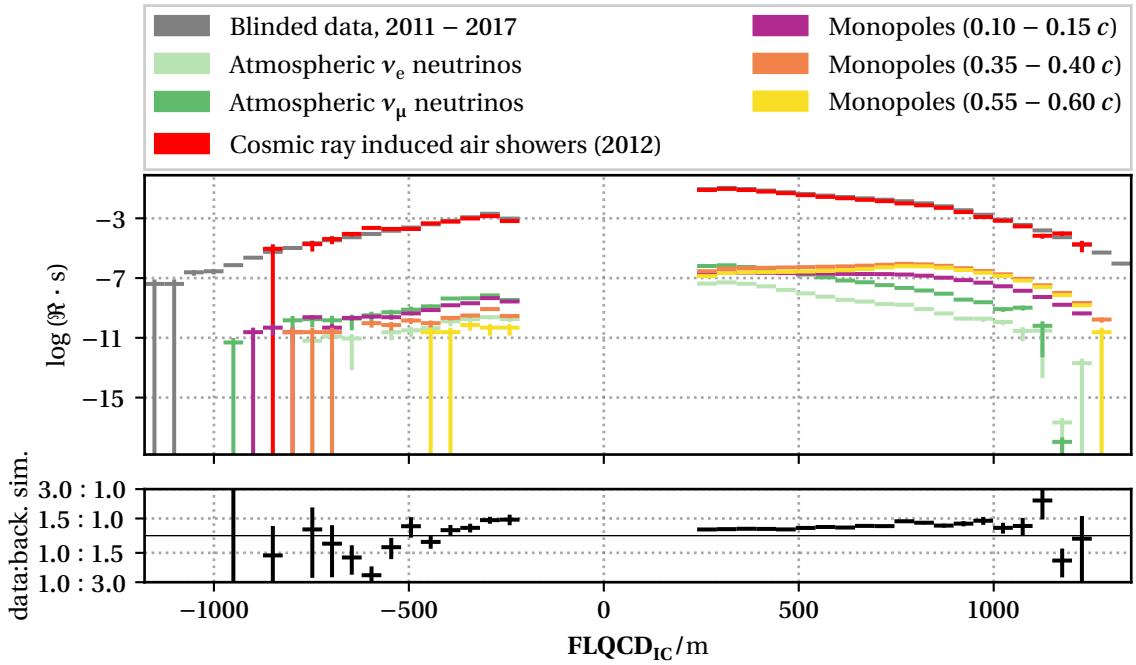


Figure D.7: **Input variable: FLQCD$_{IC}$**
*First to last quartile **COG** distance (**FLQCD**) of the hits in the **IceCube** in-ice array (**IC**) projected on the global **MPF16** track hypotheses (**GMF16TH**).*

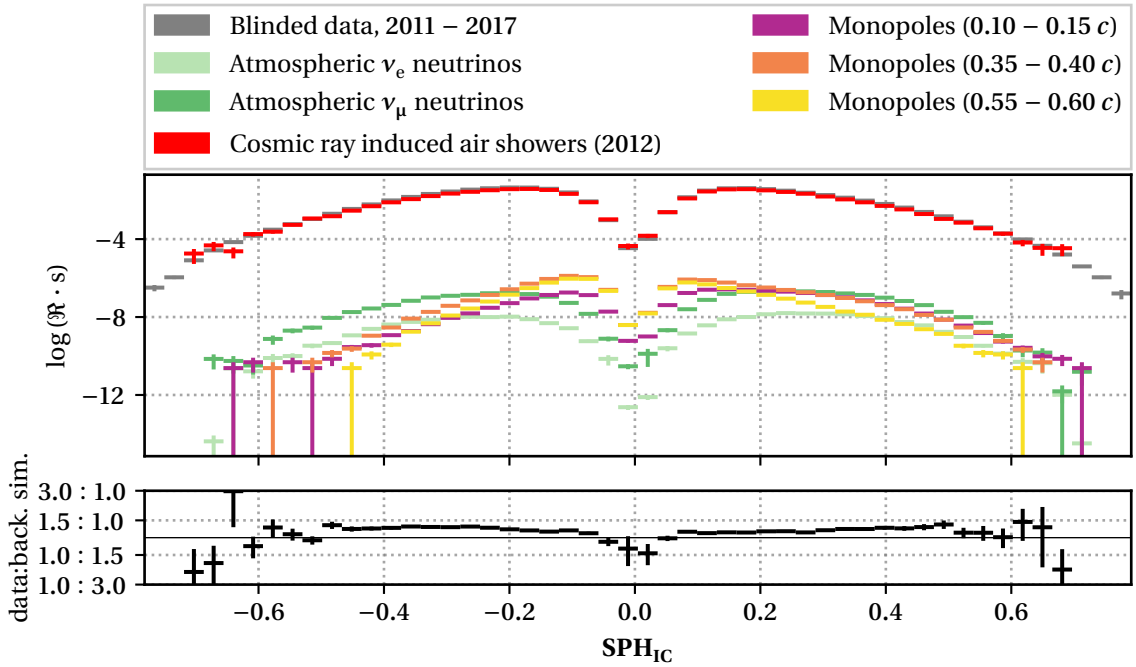


Figure D.8: **Input variable: SPH$_{IC}$**
*Smoothness of projected hits (**SPH**) on the global **MPF16** track hypotheses (**GMF16TH**) in the **IceCube** in-ice array (**IC**).*

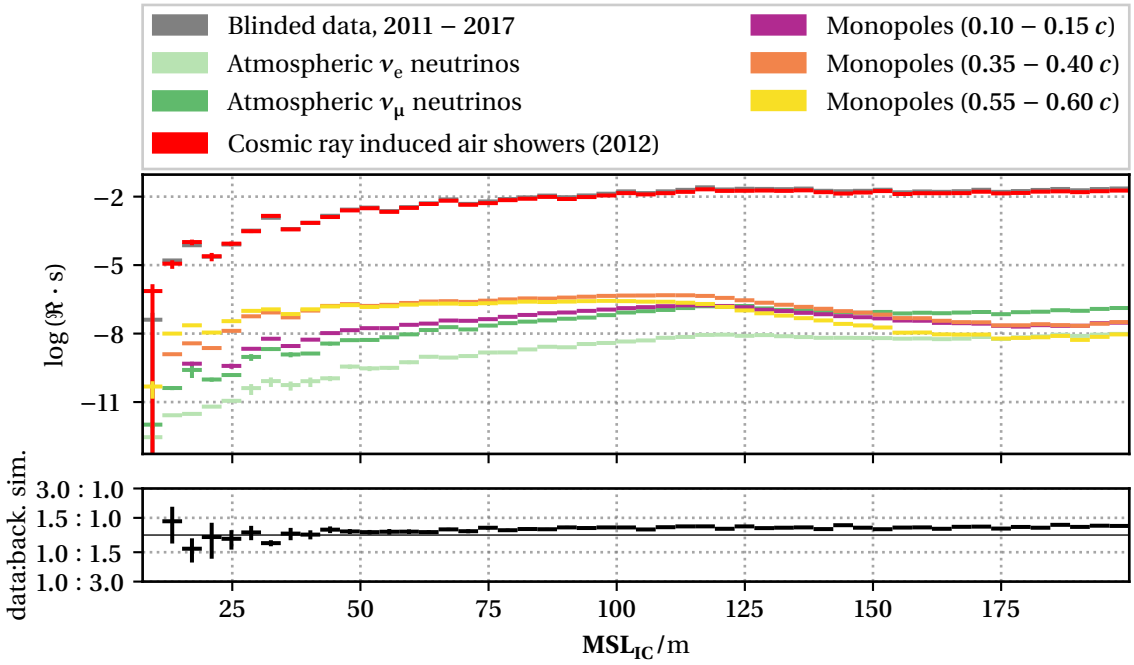


Figure D.9: **Input variable: MSL$_{IC}$**
*Maximal separation length of hits (**MSL**) projected onto the global **MPF16** track hypotheses (**GMF16TH**) in the **IceCube** in-ice array (**IC**).*

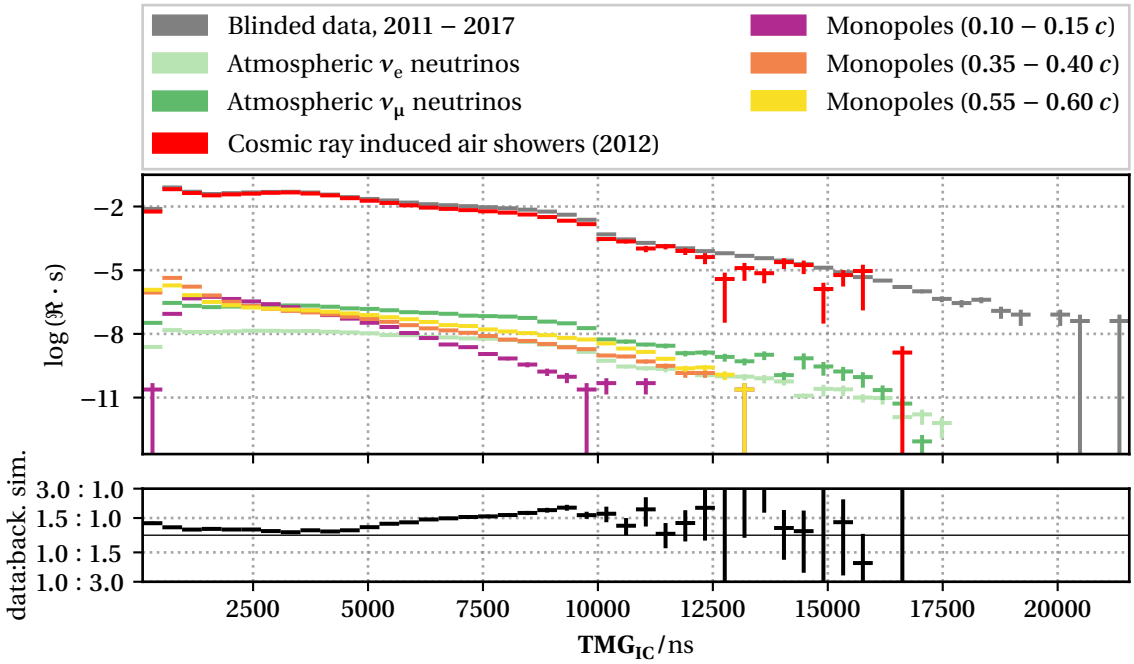


Figure D.10: **Input variable: TMG$_{IC}$**
*Temporal maximal gap (**TMG**) between hits in the **IceCube** DeepCore array (**DC**).*

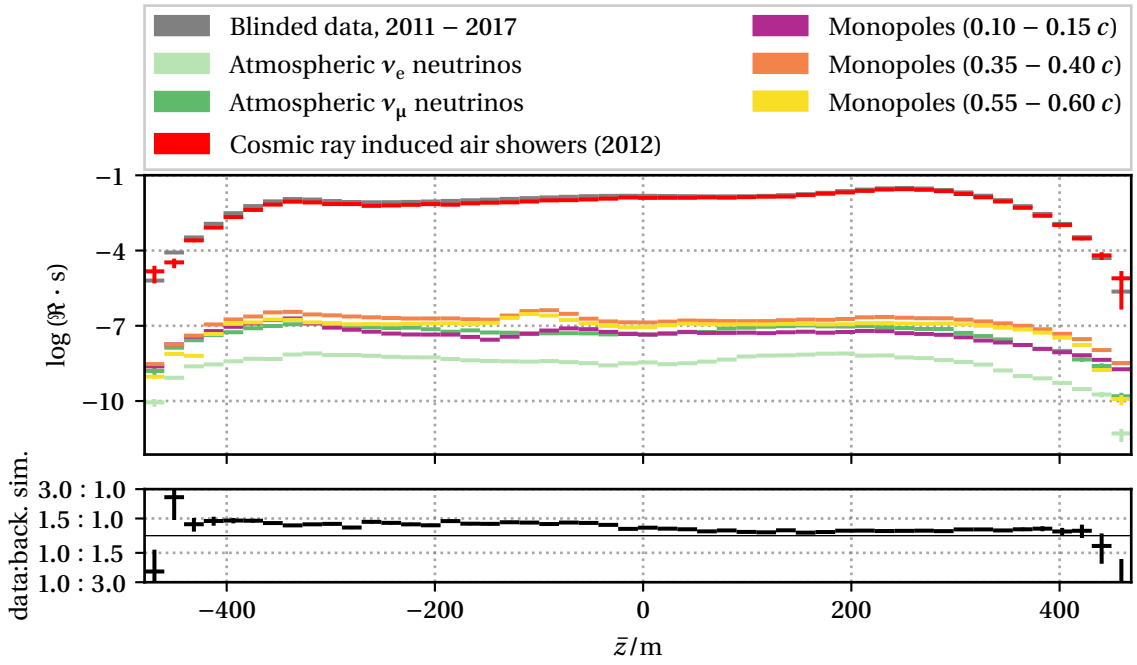


Figure D.11: **Input variable:** $\bar{z}$
*Mean height of all hit Digital Optical Modules (**DOM**s) in the event.*

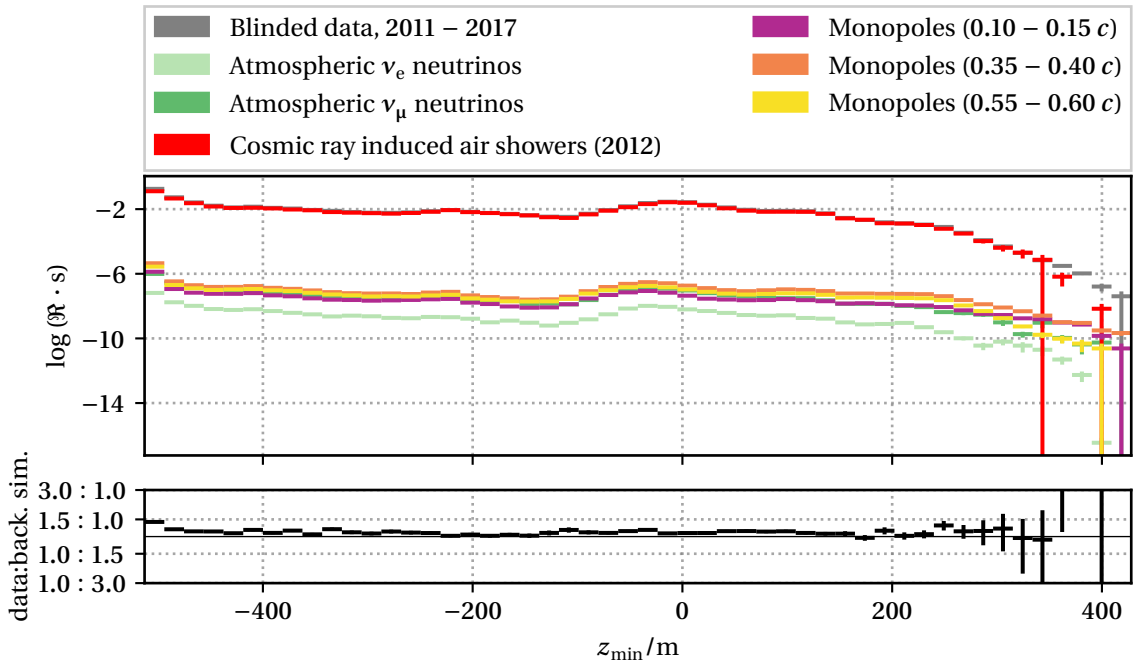


Figure D.12: **Input variable:** $z_{\mathbf{min}}$
*Minimal height of any hit Digital Optical Module (**DOM**) in the event.*

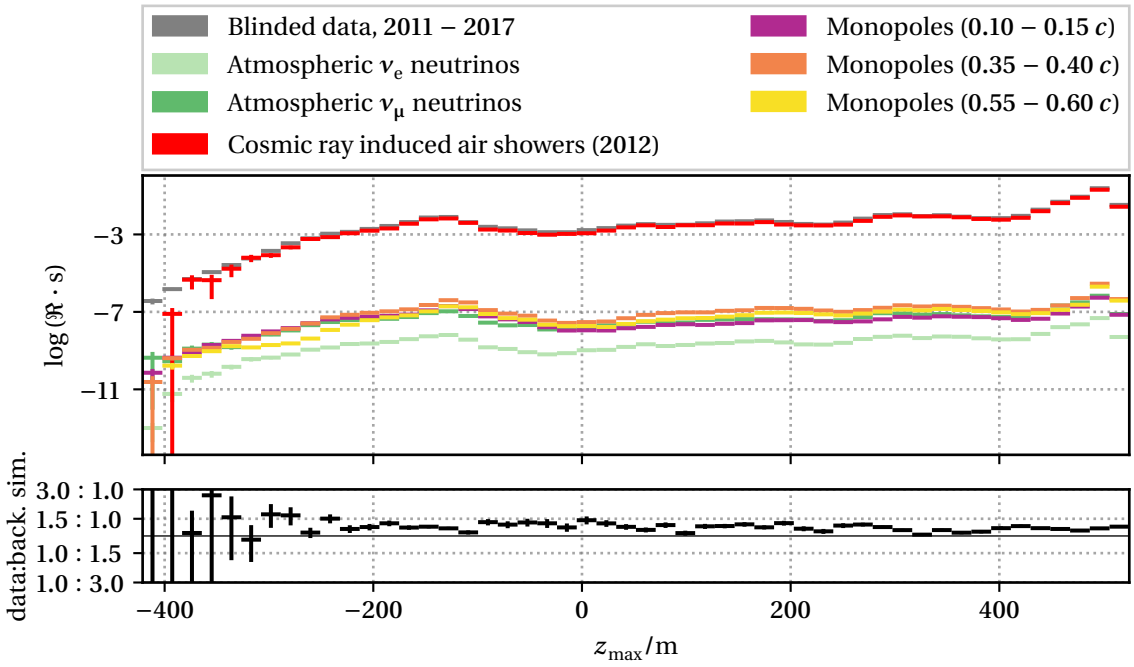


Figure D.13: **Input variable:** z_{max}
*Maximal height of any hit Digital Optical Module (**DOM**) in the event.*

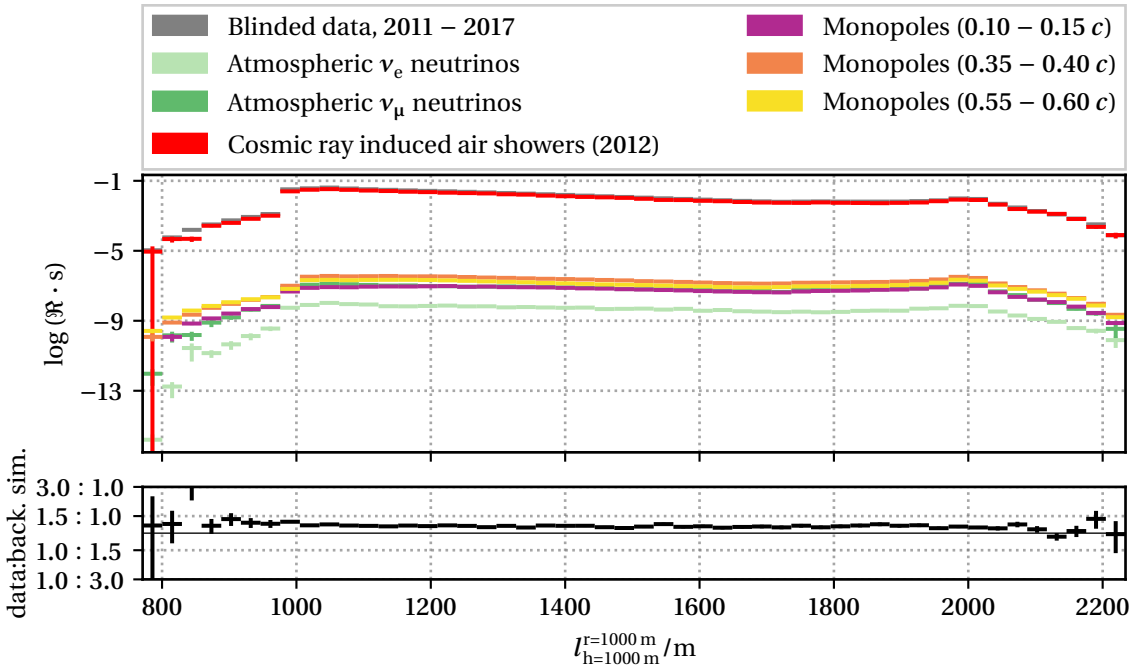


Figure D.14: **Input variable:** $l^{\text{r=1000 m}}_{\text{h=1000 m}}$
*Length of the global **MPF16** track hypotheses (**GMF16TH**) through a cylinder with a height and a radius of 1000 m at the center of the **IceCube** in-ice array (**IC**).*

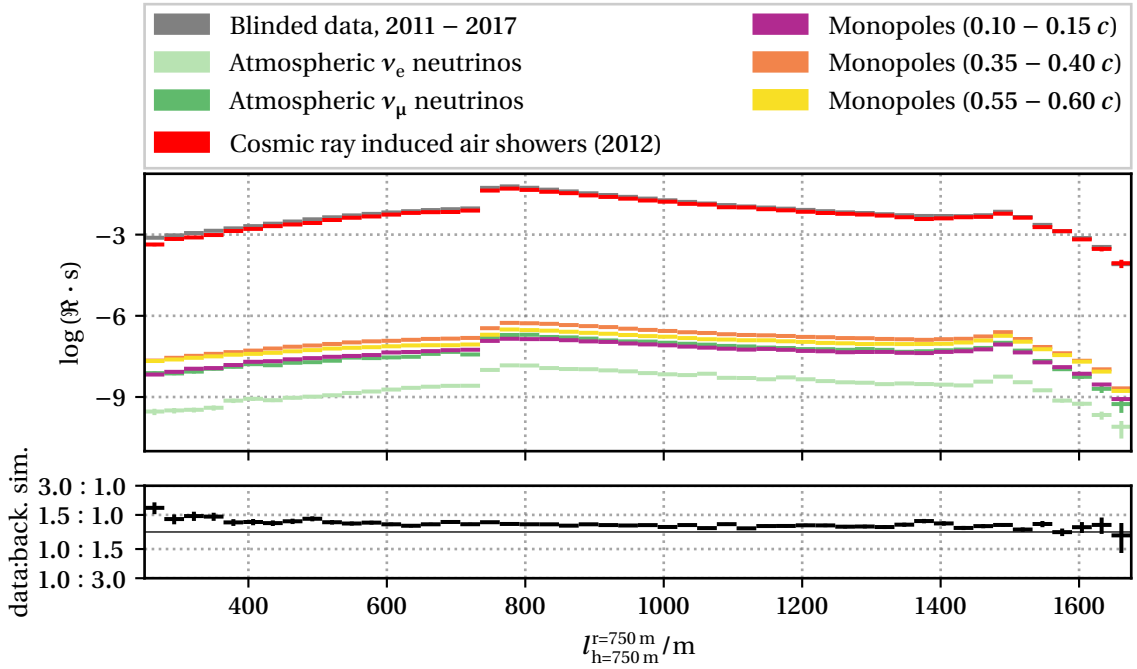


Figure D.15: **Input variable:** $l^{\mathbf{r=750\,m}}_{\mathbf{h=750\,m}}$
*Length of the global **MPF16** track hypotheses (**GMF16TH**) through a cylinder with a height and a radius of 750 m at the center of the **IceCube** in-ice array (**IC**).*

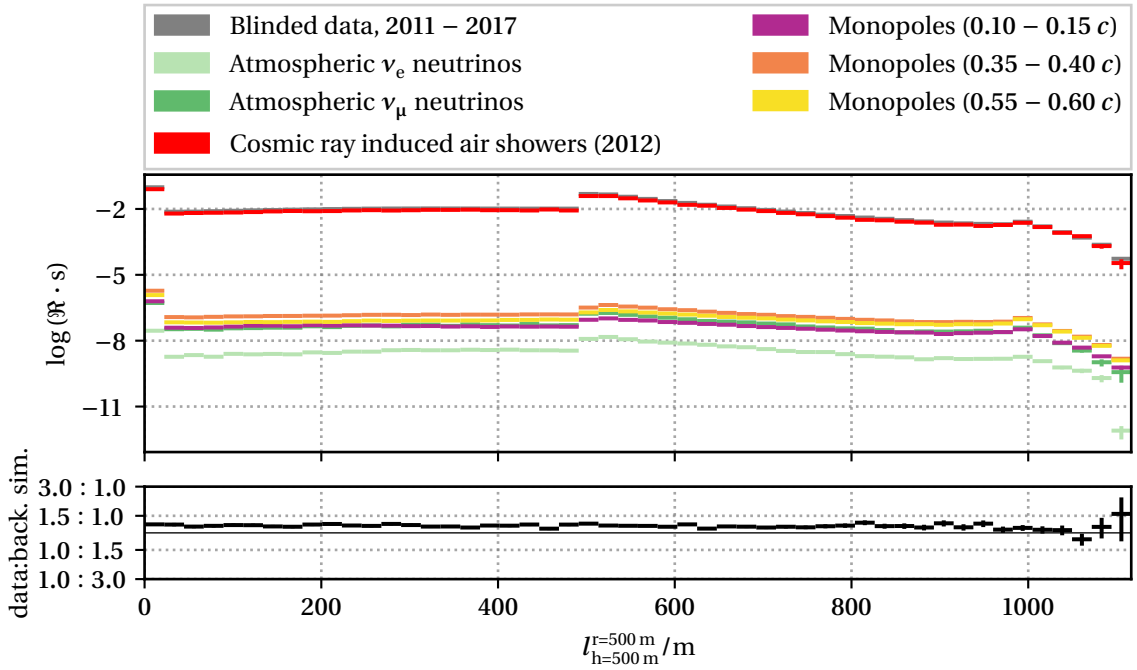


Figure D.16: **Input variable:** $l^{\mathbf{r=500\,m}}_{\mathbf{h=500\,m}}$
*Length of the global **MPF16** track hypotheses (**GMF16TH**) through a cylinder with a height and a radius of 500 m at the center of the **IceCube** in-ice array (**IC**).*

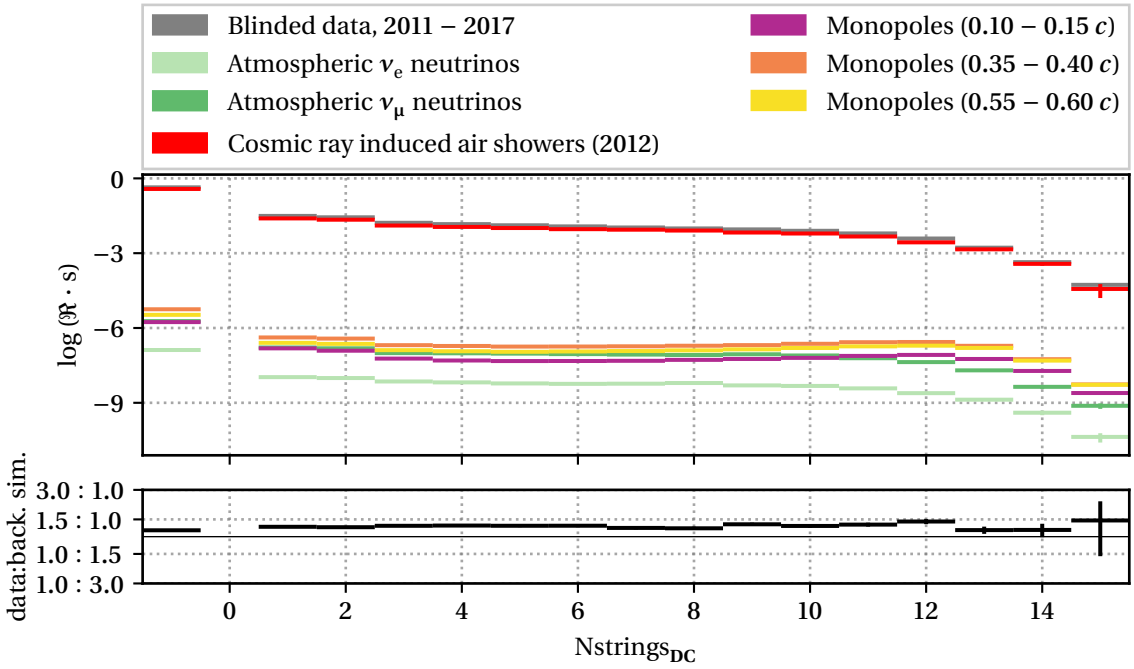


Figure D.17: **Input variable: *N*strings$_{DC}$**
*Number of strings which had at least one Digital Optical Module (**DOM**) with a hit in the **IceCube** DeepCore array (**DC**). Note: a value of -1 indicates the value was missing i. e. no strings detected any hits.*

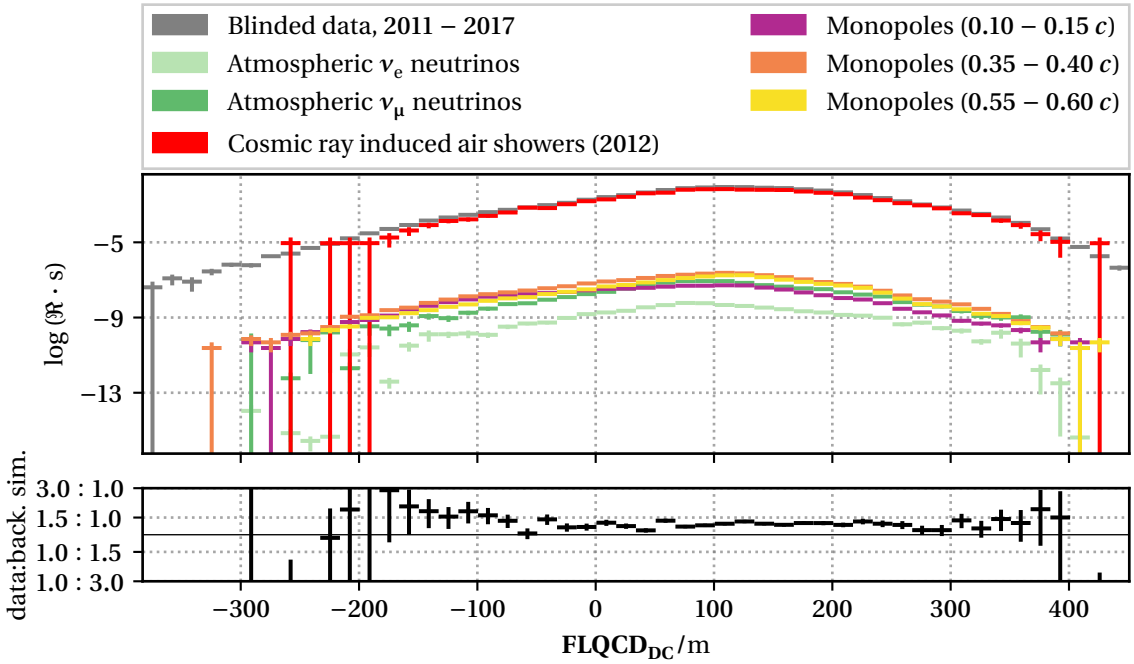


Figure D.18: **Input variable: FLQCD$_{DC}$**
*First to last quartile **COG** distance (**FLQCD**) of the hits in the **IceCube** DeepCore array (**DC**) projected on the global **MPF16** track hypotheses (**GMF16TH**).*

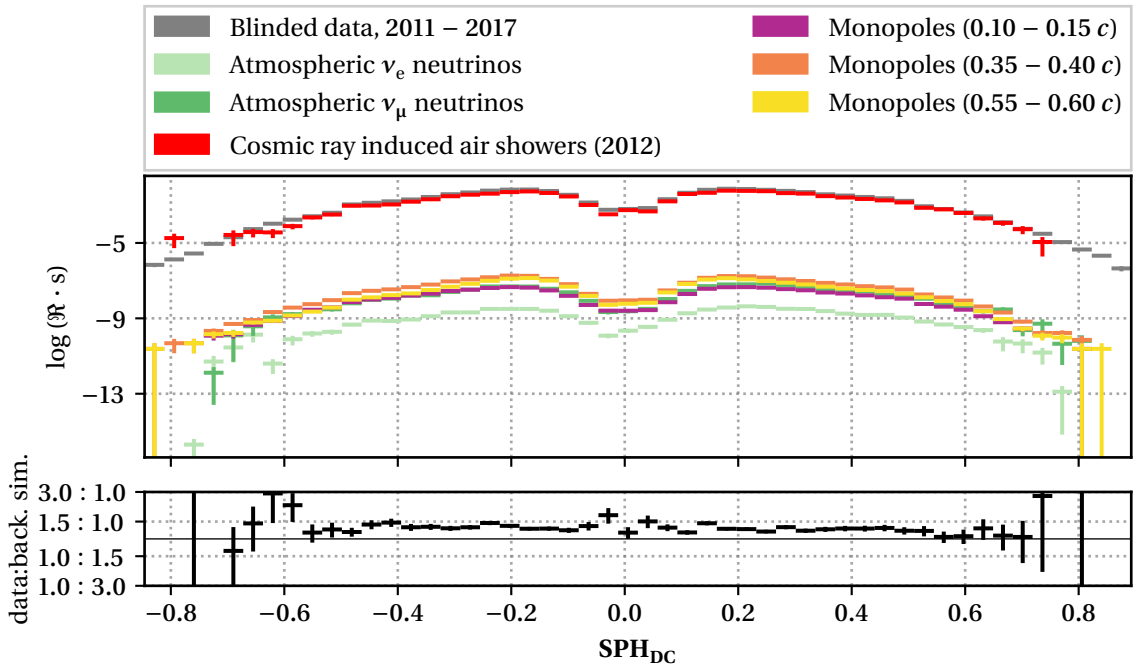


Figure D.19: **Input variable: SPH$_{DC}$**
*Smoothness of projected hits (**SPH**) on the global **MPF16** track hypotheses (**GMF16TH**) in the **IceCube** DeepCore array (**DC**).*

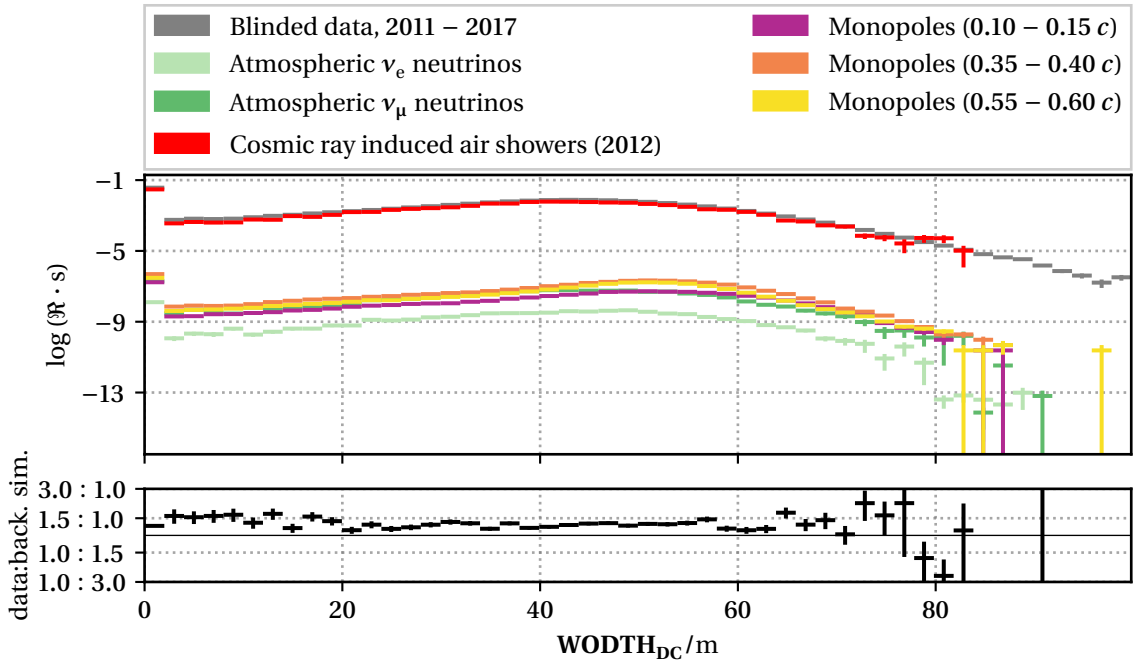


Figure D.20: **Input variable: WODTH$_{DC}$**
*Weighted orthogonal distance from the track hypothesis (**WODTH**) in the **IceCube** DeepCore array (**DC**).*

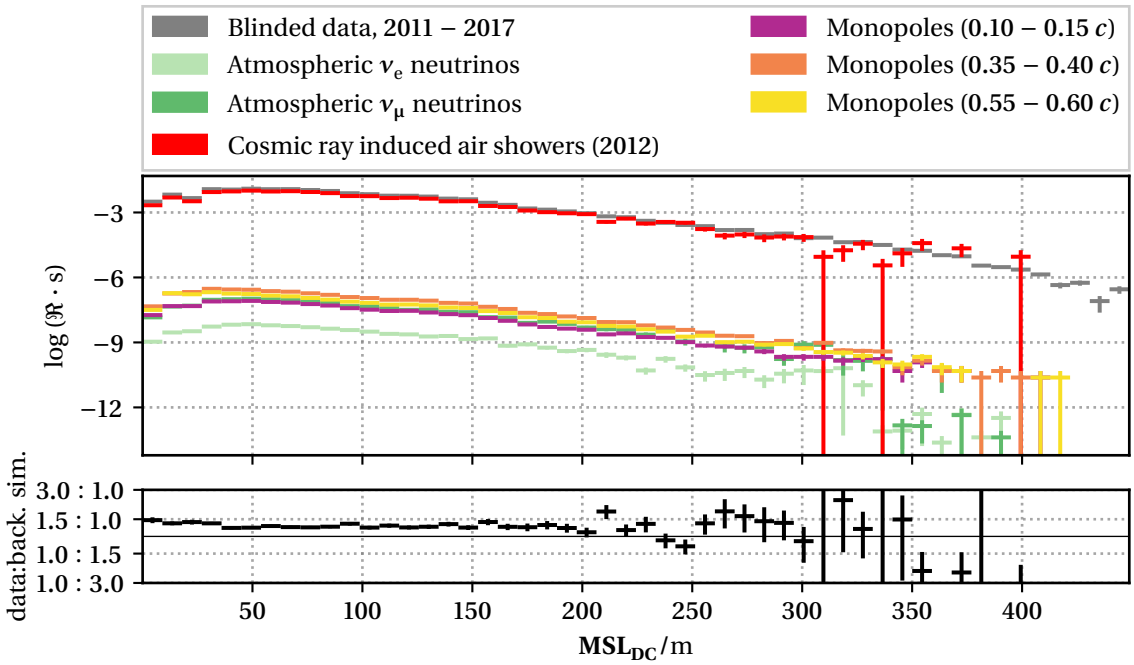


Figure D.21: **Input variable: MSL$_{DC}$**
*Maximal separation length of hits (**MSL**) projected onto the global **MPF16** track hypotheses (**GMF16TH**) in the **IceCube** DeepCore array (**DC**).*

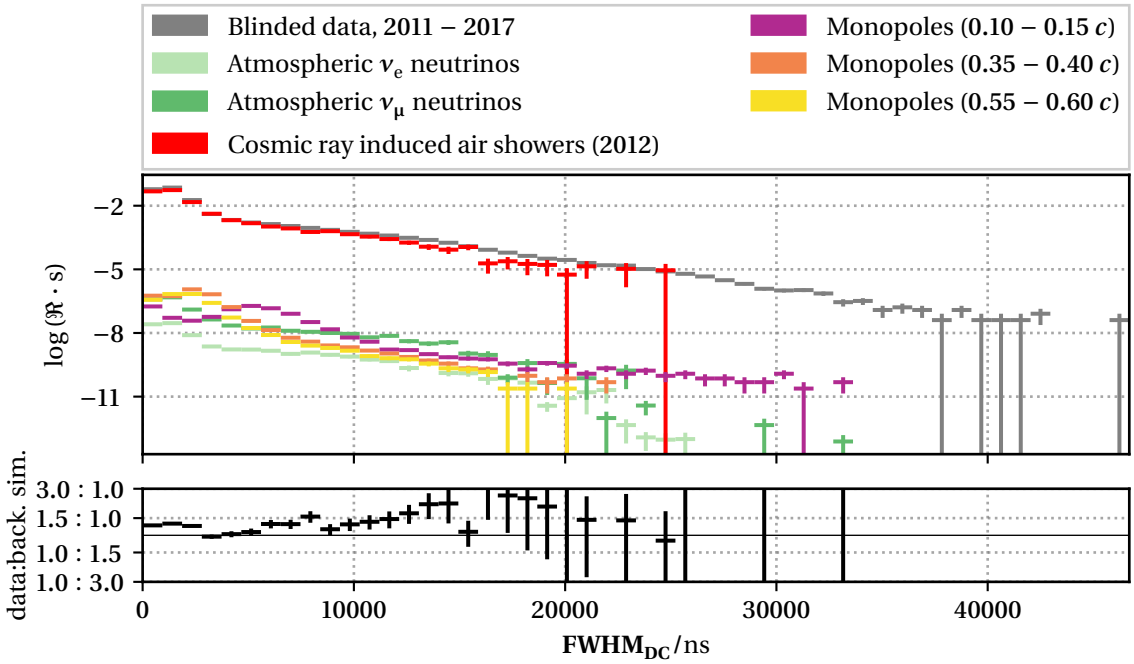


Figure D.22: **Input variable: FWHM$_{DC}$**
*Full width half maximum (**FWHM**) of the timing distribution of hits in the **IceCube** DeepCore array (**DC**).*

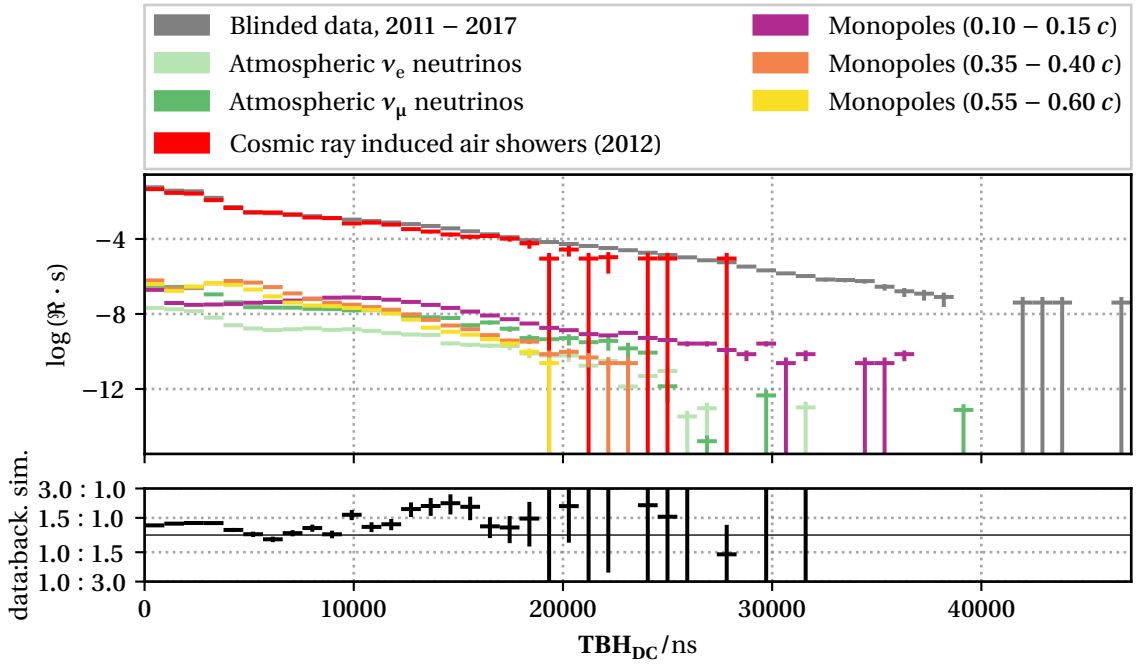


Figure D.23: **Input variable: TBH$_{DC}$**
*Maximal temporal distance between first hits (**TBH**) in the **IceCube** DeepCore array (**DC**).*

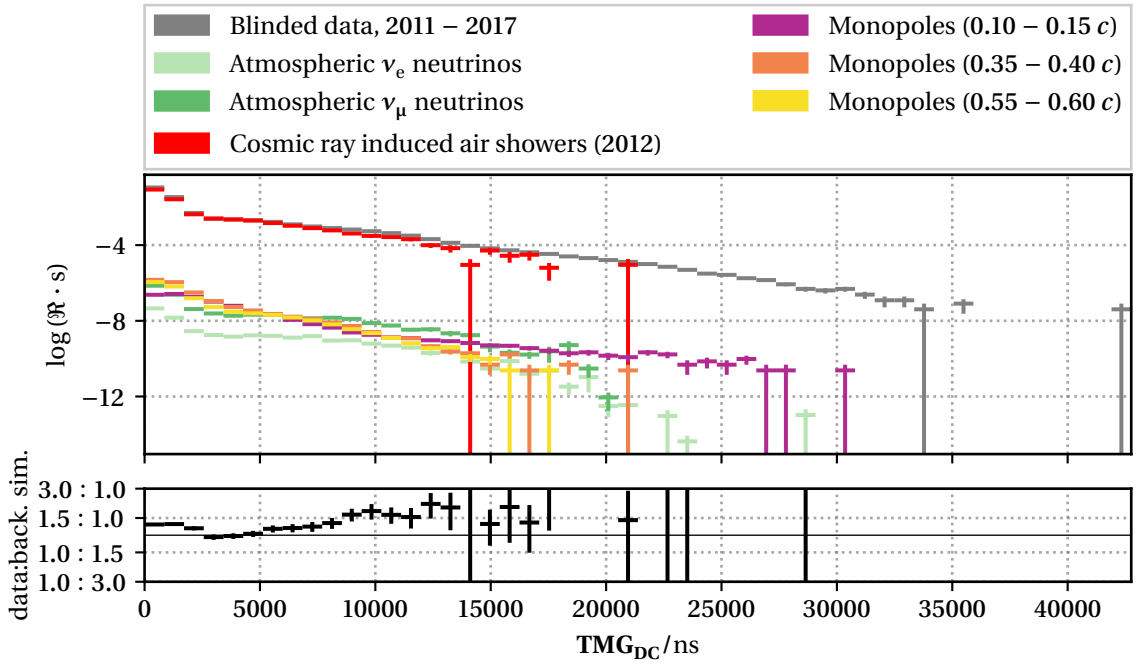


Figure D.24: **Input variable: TMG$_{DC}$**
*Temporal maximal gap (**TMG**) between hits in the **IceCube** DeepCore array (**DC**).*

Appendix E

Feature Importance of Different Boosted Decision Trees During the Last Selection Step

In this section, feature importance metrics are shown for different boosted decission treess (**BDT**s) during the last selection step described in Section 7.3. Two metrics are utilized, the *weight* metric indicating how often a feature is used by a singular variate decision (**SVD**) within the **BDT**s and the *gain* metric indicating the increase of accuracy of the **BDT**s due to adding the feature. Figures E.1 and E.2 are indicative for a specialized **BDT** to separate simulated events from measured events described in Section 7.3.2. Figures E.2 and E.4 are indicative for the set of 1000 **BDT**s utilized in the final selection step. To combine the feature importance of all **BDT**s, violin plots are used. It can be observed that features which have the strongest separation strength between simulated background events and measured events are not strong features for the **BDT**s separating signal-like and background-like events.

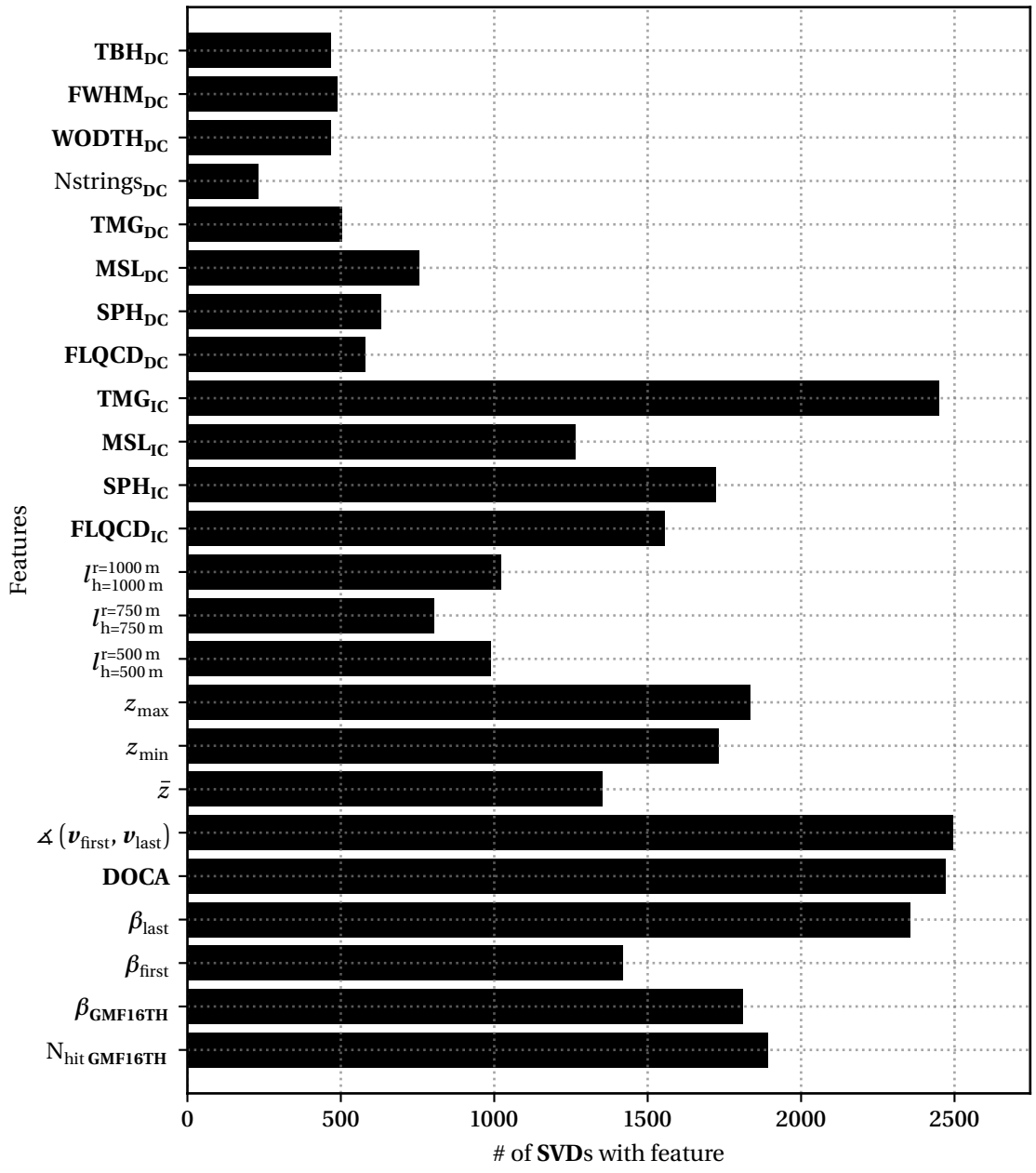


Figure E.1: **Feature weight of last BDT of feature selection step**
*A bar plot of the distribution of feature importance by weight for the last **BDT** of the iterative input feature selection process described in Section 7.3.2 is drawn. On the vertical axis, all features utilized as input features are shown, the linear, horizontal axis represents the number of **SVD**s utilizing said features within the **BDT**s.*

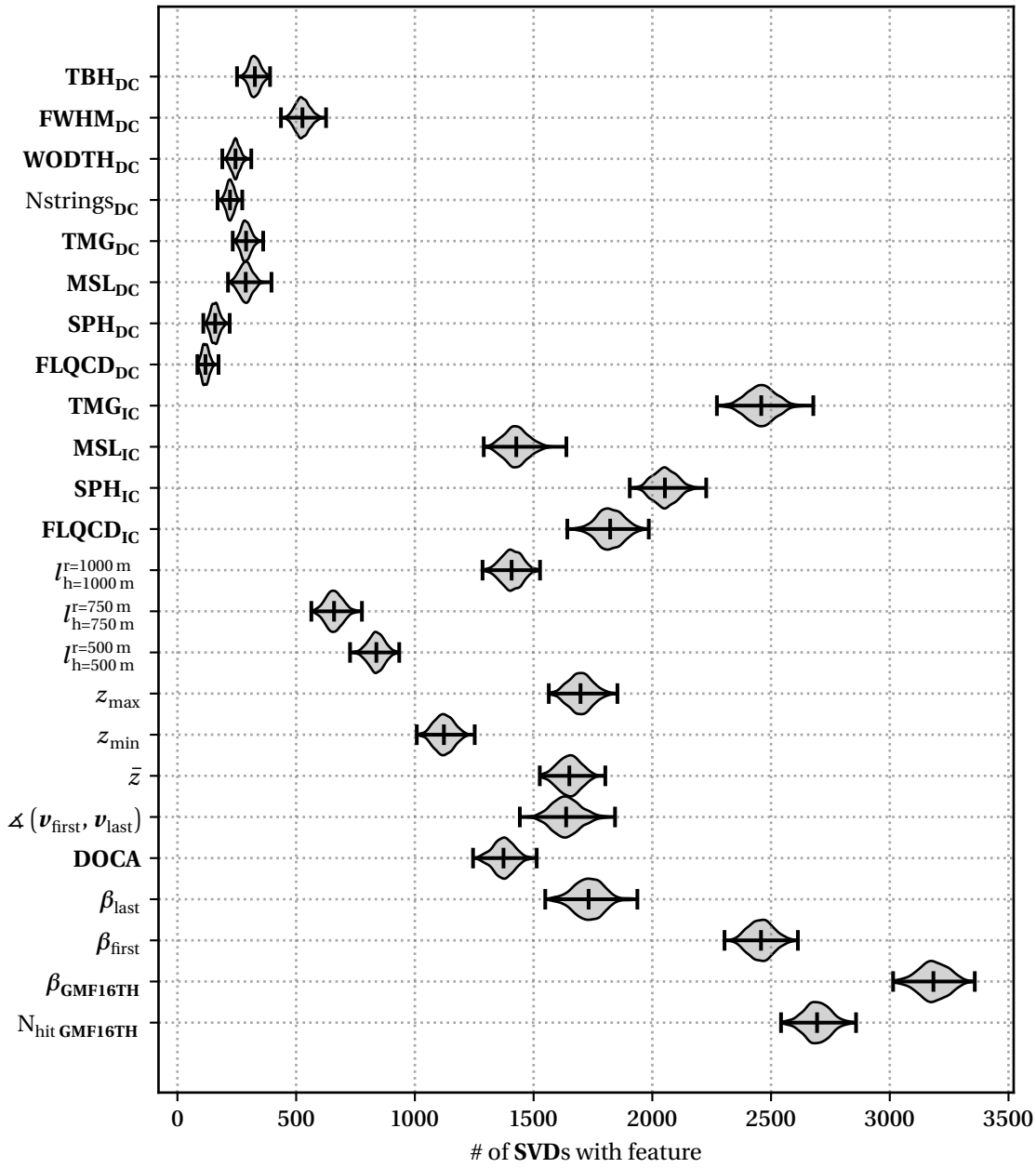


Figure E.2: **Feature weight of all BDTs of last selection step**
*A violin plot of the distribution of feature importance by weight for the 1000 trained **BDT**s is presented. On the vertical axis, all features utilized as input features are shown, the linear, horizontal axis represents the number of **SVD**s utilizing said features within the **BDT**s. For each feature, the mean, minimal, and maximal importance observed in the set of 1000 **BDT**s is drawn. Additionally, a Gaussian kernel density estimate of the distribution of values for each feature is added, drawn as a shaded region.*

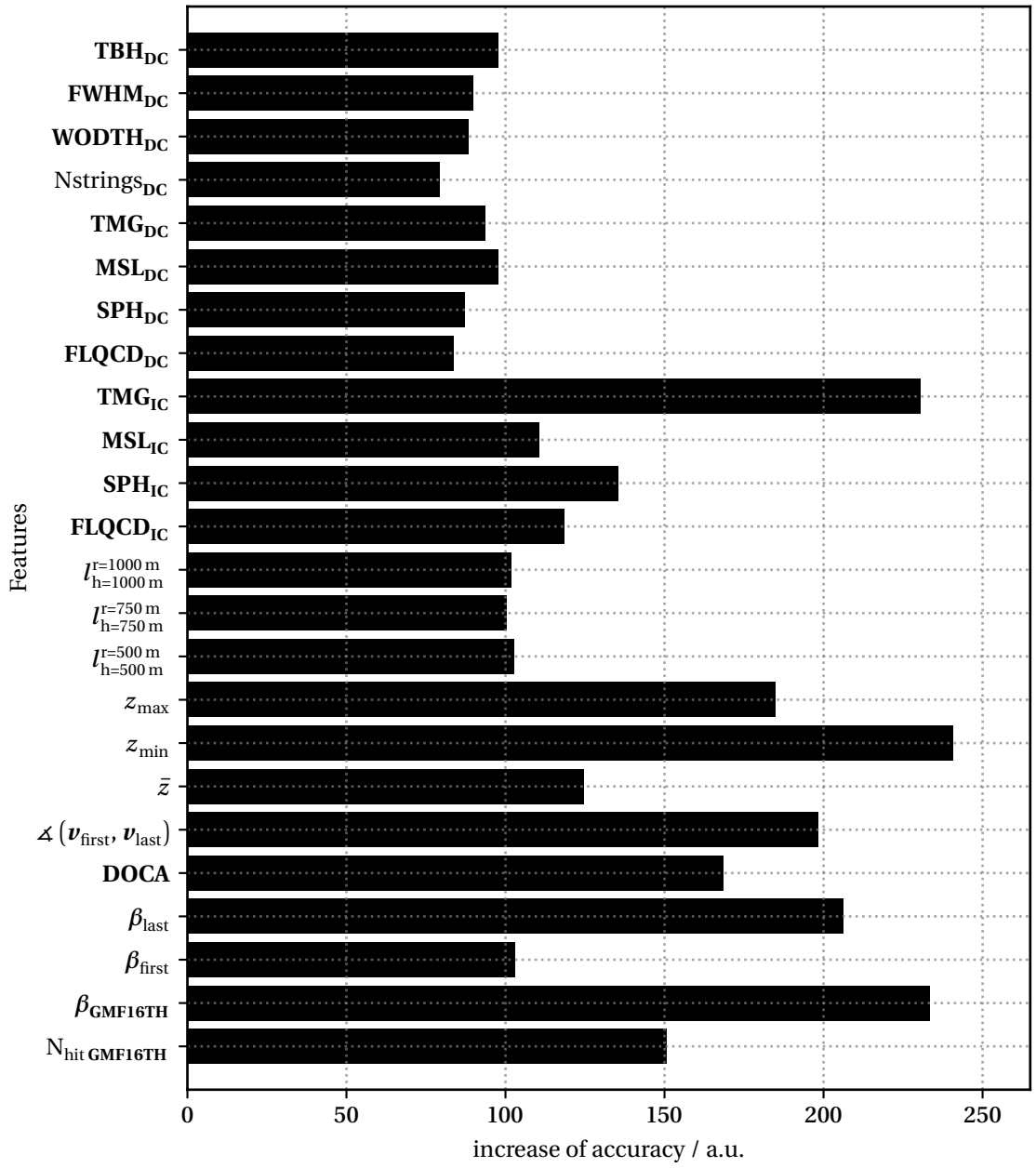


Figure E.3: **Feature gain of last BDT of feature selection step**
*A bar plot of the distribution of feature importance by gain for the last **BDT** of the iterative input feature selection process described in Section 7.3.2 is drawn. On the vertical axis, all features utilized as input features are shown, the linear, horizontal axis represents the increase of accuracy due to said features.*

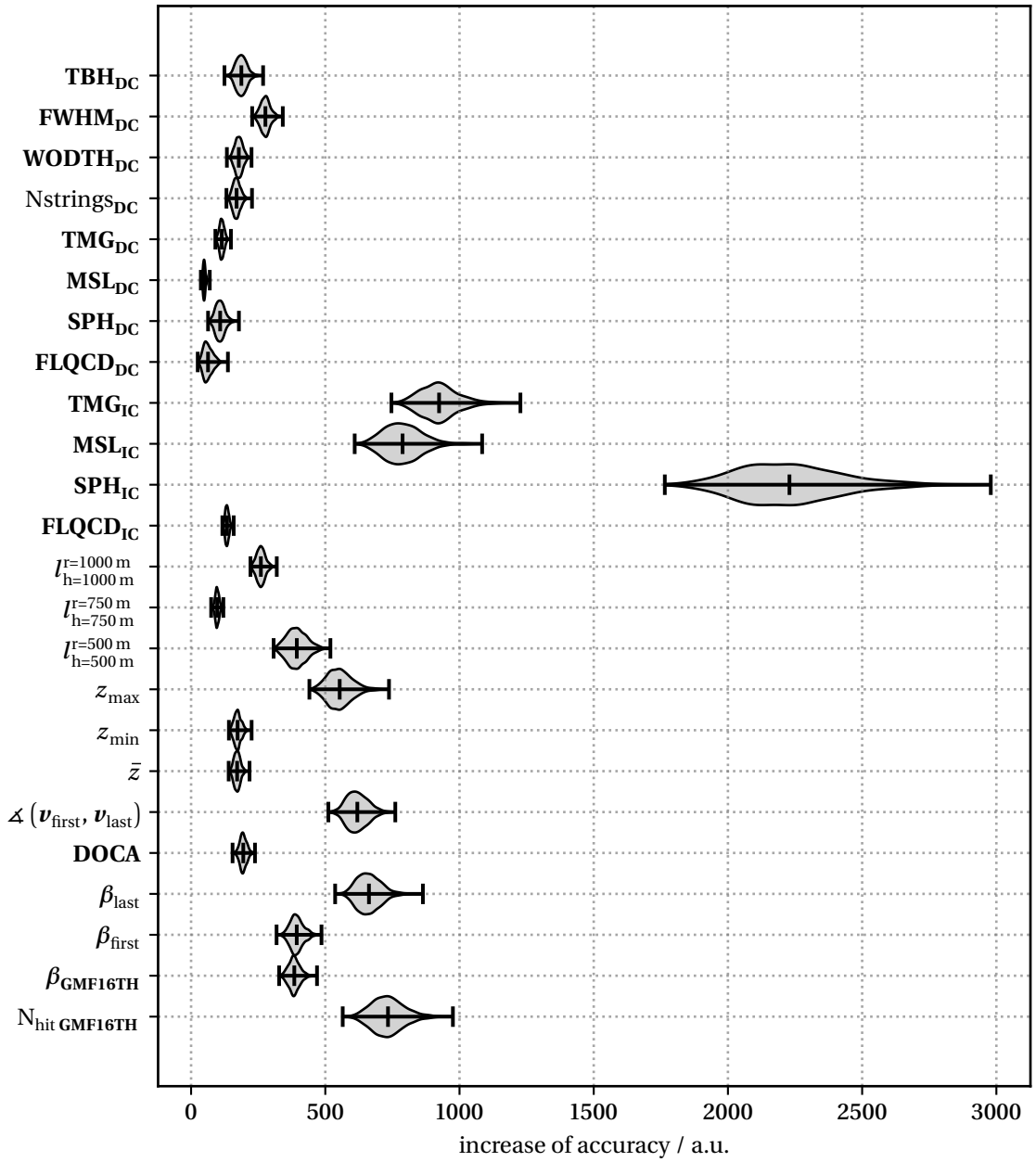


Figure E.4: **Feature gain of all BDTs of last selection step**
*A violin plot of the distribution of feature importance by gain for the 1000 trained **BDT**s is presented. On the vertical axis, all features utilized as input features are shown, the linear, horizontal axis represents the increase of accuracy due to said features. For each feature, the mean, minimal, and maximal importance observed in the set of 1000 **BDT**s is drawn. Additionally, a Gaussian kernel density estimate of the distribution of values for each feature is added.*

Appendix F

Data of Derived Flux Limit

The derived flux limit, depicted in Figure 8.5, is printed as a velocity dependent, equi-spaced histogram with the left hand sides of the bins β_l and the deduced upper flux bound Φ_{90} divided by a constant factor $W = 10^{-18}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{sr}^{-2}$.

β_l	Φ_{90}/W	β_l	Φ_{90}/W	β_l	Φ_{90}/W	β_l	Φ_{90}/W
0.100	9.521	0.225	1.523	0.350	1.965	0.475	1.996
0.105	7.139	0.230	1.505	0.355	1.971	0.480	2.136
0.110	5.790	0.235	1.499	0.360	1.999	0.485	2.252
0.115	5.076	0.240	1.534	0.365	2.085	0.490	2.265
0.120	3.917	0.245	1.486	0.370	2.060	0.495	2.455
0.125	3.365	0.250	1.499	0.375	2.132	0.500	2.558
0.130	2.706	0.255	1.501	0.380	2.104	0.505	2.612
0.135	2.339	0.260	1.489	0.385	2.093	0.510	2.820
0.140	1.939	0.265	1.473	0.390	2.131	0.515	2.810
0.145	1.595	0.270	1.507	0.395	2.160	0.520	2.888
0.150	1.359	0.275	1.523	0.400	2.084	0.525	2.673
0.155	1.241	0.280	1.555	0.405	2.158	0.530	2.207
0.160	1.197	0.285	1.576	0.410	2.121	0.535	1.698
0.165	1.159	0.290	1.599	0.415	2.102	0.540	1.435
0.170	1.139	0.295	1.703	0.420	2.122	0.545	1.308
0.175	1.159	0.300	1.646	0.425	2.166	0.550	1.184
0.180	1.167	0.305	1.750	0.430	2.056	0.555	1.034
0.185	1.226	0.310	1.710	0.435	2.077	0.560	0.956
0.190	1.244	0.315	1.807	0.440	2.053	0.565	0.974
0.195	1.303	0.320	1.811	0.445	2.039	0.570	0.985
0.200	1.329	0.325	1.914	0.450	2.010	0.575	1.126
0.205	1.388	0.330	1.871	0.455	1.984	0.580	1.305
0.210	1.405	0.335	1.887	0.460	2.035	0.585	1.779
0.215	1.460	0.340	1.961	0.465	2.062	0.590	2.783
0.220	1.434	0.345	1.974	0.470	2.038	0.595	6.089

Table F.1: **Table with final flux limit**

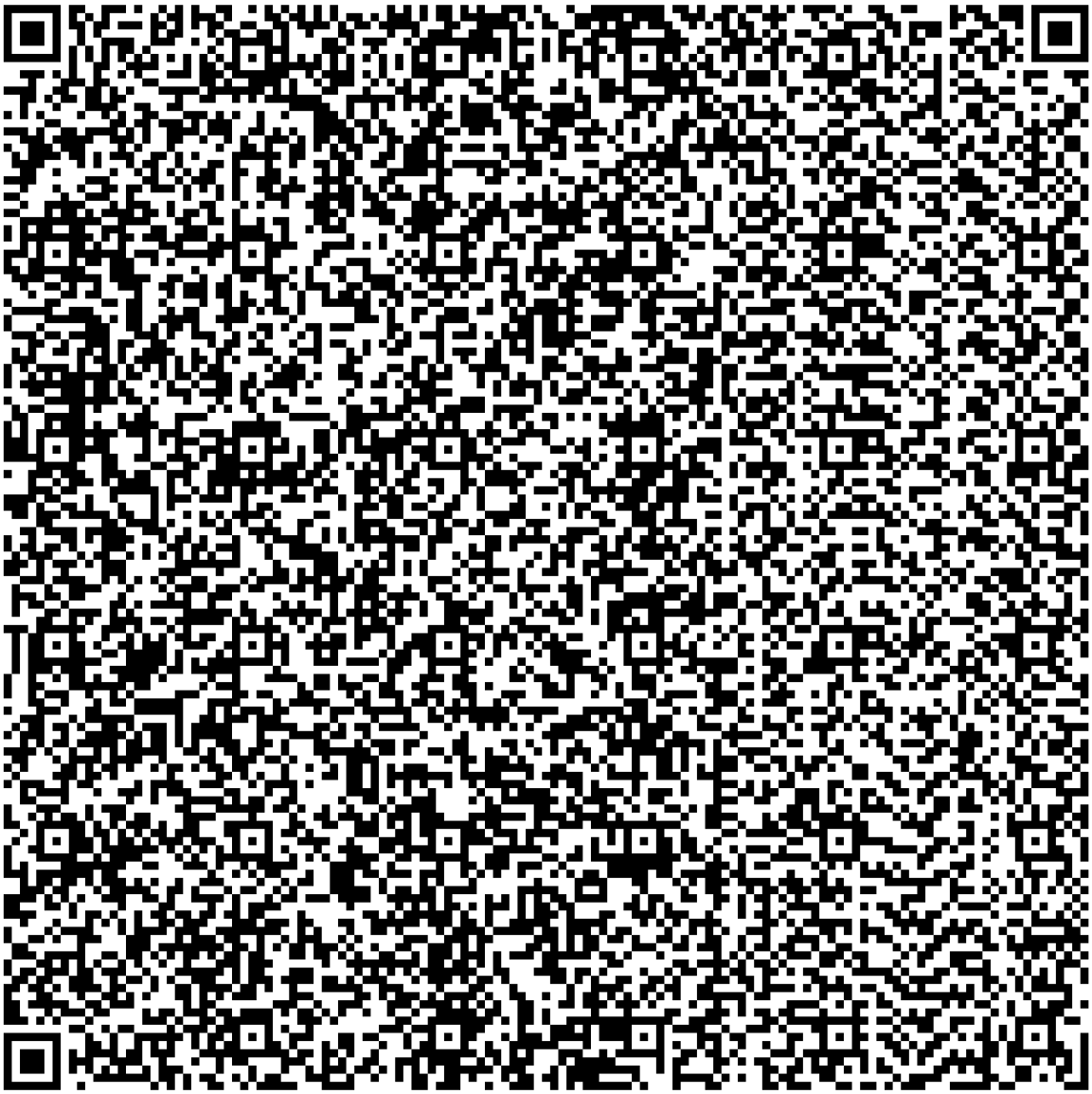


Figure F.1: **Quick response encoded final flux limit**
The same data as in Table F.1 is presented in a machine readable quick response encoding defined in reference [ISO15]. Recovery of the data was tested on multiple devices, both from paper versions printed at a resolution of 600 dots per inch and digitally displayed ones.

Acronyms

ATWD	Analog Transient Waveform Digitizer
bagging	bootstrap aggregating
BDT	boosted decission trees
COG	center of gravity
CORSIKA	COsmic Ray SImulations for **KASCADE**
DAQ	data acquisition
DC	**IceCube** DeepCore array
DOCA	distance of closest approach
DOM	Digital Optical Module
DT	decision tree
fADC	Fast Analog to Digital Converter
FLQCD	first to last quartile **COG** distance
FWHM	full width half maximum
GMF16TH	global **MPF16** track hypotheses
GUT	Grand Unified Theory
HLC	hard local coincidence
IC	**IceCube** in-ice array
ICC	**IceCube** collaboration
IceCube	IceCube Neutrino Observatory
ICL	**IceCube** laboratory
iLF	improved LineFit
IMM	intermediate mass magnetic monopole
KASCADE	KArlsruhe Shower Core and Array DEtector
KDE	kernel density estimator
KYG	Kazama Yang Goldhaber
MCS	Monte Carlo Simulation
ML	machine learning

MM	magnetic monopole
MPF16	MonopoleFilter_16
MRF	model rejection factor
MSL	maximal separation length of hits
PMT	photomultiplier tube
PnF	online Processing and Filtering
PROPOSAL	PROpagator with optimal Precision and Optimized Speed for All Leptons
SE	standard error
SLC	soft local coincidence
SMP	highly ionizing, stable, massive particle
SMT	simple multiplicity trigger
SPH	smoothness of projected hits
SPICE	South Pole ICE
SVD	singular variate decision
TBH	maximal temporal distance between first hits
TMG	temporal maximal gap
TWDSiLF	time-wise double split **iLF**s
WODTH	weighted orthogonal distance from the track hypothesis
XGBoost	eXtreme Gradient Boosting

Bibliography

[Aar+13] M. Aartsen et al. “Measurement of South Pole ice transparency with the IceCube LED calibration system”. In: *Nucl. Instrum. Methods Phys. Res. A* 711 (2013), pp. 73–89. DOI: 10.1016/j.nima.2013.01.054.

[Aar+14a] M. G. Aartsen et al. “Energy reconstruction methods in the IceCube neutrino telescope”. In: *J. Instrum.* 9.03 (2014), pp. 3009-1–3009-21. DOI: 10.1088/1748-0221/9/03/p03009.

[Aar+14b] M. Aartsen et al. “Improvement in fast particle track reconstruction with robust statistics”. In: *Nucl. Instrum. Methods Phys. Res. A* 736 (2014), pp. 143–149. DOI: 10.1016/j.nima.2013.10.074.

[Aar+16] M. G. Aartsen et al. “Searches for relativistic magnetic monopoles in IceCube”. In: *Eur. Phys. J. C* 76 (2016), pp. 133-1–133-16. DOI: 10.1140/epjc/s10052-016-3953-8.

[Aar+17] M. G. Aartsen et al. “The IceCube Neutrino Observatory: Instrumentation and online systems”. In: *J. Instrum.* 12.03 (2017), pp. 03012-1–03012-83. DOI: 10.1088/1748-0221/12/03/p03012.

[Aar+20] M. G. Aartsen et al. “In-situ calibration of the single-photoelectron charge response of the IceCube photomultiplier tubes”. In: *J. Instrum.* 15.06 (2020), pp. 6032-1–6032-27. DOI: 10.1088/1748-0221/15/06/p06032.

[Aar+21] M. G. Aartsen et al. “IceCube-Gen2: The window to the extreme Universe”. In: *J. Phys. G* 48.6 (2021), pp. 060501-1–060501-93. DOI: 10.1088/1361-6471/abbd48.

[AB59] Y. Aharonov and D. Bohm. “Significance of electromagnetic potentials in the quantum theory”. In: *Phys. Rev.* 115.3 (1959), pp. 485–491. DOI: 10.1103/PhysRev.115.485.

[Abb+10a] R. Abbasi et al. “Calibration and characterization of the IceCube photomultiplier tube”. In: *Nucl. Instrum. Methods Phys. Res. A* 618 (2010), pp. 139–152. DOI: 10.1016/J.NIMA.2010.03.102.

[Abb+10b] R. Abbasi et al. “Search for relativistic magnetic monopoles with the AMANDA-II neutrino telescope”. In: *Eur. Phys. J. C* 69.3 (2010), pp. 361–378. DOI: 10.1140/epjc/s10052-010-1411-6.

[Ach+06] A. Achterberg et al. “First year performance of the IceCube neutrino telescope”. In: *Astropart. Phys.* 26.3 (2006), pp. 155–173. DOI: 10.1016/j.astropartphys.2006.06.007.

[Ada80] A. Adams. *The eloquent light*. Aperture, 1980, p. 86.

[Agr20] V. K. Agrawal. *Mastering machine learning with Core ML and Python*. AppCoda, 2020, p. 16.

[Ahl78] S. P. Ahlen. “Stopping-power formula for magnetic monopoles”. In: *Phys. Rev. D* 17.1 (1978), pp. 229–233. DOI: 10.1103/PhysRevD.17.229.

[AK82] S. P. Ahlen and K. Kinoshita. “Calculation of the stopping power of very-low-velocity magnetic monopoles”. In: *Phys. Rev. D* 26.9 (1982), pp. 2347–2363. DOI: 10.1103/PhysRevD.26.2347.

[Alb+17] A. Albert et al. “Search for relativistic magnetic monopoles with five years of the ANTARES detector data”. In: *J. High. Energ. Phys.* 2017.7 (2017), pp. 54-1–54-19. DOI: 10.1007/JHEP07(2017)054.

[Amb+02] M. Ambrosio et al. “Final results of magnetic monopole searches with the MACRO experiment”. In: *Eur. Phys. J. C* 25 (2002), pp. 511–522. DOI: 10.1140/epjc/s2002-01046-9.

[Amb+97] M. Ambrosio et al. “Magnetic monopole search with the MACRO detector at Gran Sasso”. In: *Phys. Lett. B* 406.3 (1997), pp. 249–255. DOI: 10.1016/S0370-2693(97)00684-9.

[Ayn+08] V. Aynutdinov et al. “Search for relativistic magnetic monopoles with the Baikal neutrino telescope”. In: *Astropart. Phys.* 29.6 (2008), pp. 366–372. DOI: 10.1016/j.astropartphys.2008.03.006.

[Bak+88] M. G. Bakker et al. “The effect of crystal fragmentation on the luminescence from pulse-irradiated H_2O ice”. In: *Int. J. Radiat. Appl. Instrum. C: Radiat. Phys. Chem.* 32.6 (1988), pp. 767–772. DOI: 10.1016/1359-0197(88)90166-X.

[Bel53] E. H. Belcher. “The luminescence of irradiated transparent media and the Cerenkov Effect. I. The luminescence of aqueous solutions of radioactive isotopes”. In: *Proc. Roy. Soc. Lond. Math. Phys. Sci.* 216.1124 (1953), pp. 90–102. DOI: 10.1098/rspa.1953.0009.

[Blo33] F. Bloch. “Bremsvermögen von Atomen mit mehreren Elektronen”. In: *Z. Phys.* 81.5 (1933), pp. 363–376. DOI: 10.1007/BF01344553.

[Bra13] S. Bravo. *Neutrino telescope shines light on the last glaciation*. The IceCube Colloboration. 2013.

[Bre96] L. Breiman. “Bagging predictors”. In: *Mach. Learn.* 24.2 (1996), pp. 123–140. DOI: 10.1007/BF00058655.

[Bur20] A. Burgman. “Bright needles in a haystack: A search for magnetic monopoles using the IceCube Neutrino Observatory”. PhD thesis. Uppsala University, High Energy Physics, 2020, p. 166.

[BV04] S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004, pp. 298–300.

[Čer34] P. A. Čerenkov. “Vidimoe svechenie chistyh zhidkostej pod dejstviem γ-radiacii”. In: *Proc. USSR Acad. Sci.* 2 (1934), pp. 451–454.

[Čer37] P. A. Čerenkov. “Visible radiation produced by electrons moving in a medium with velocities exceeding that of light”. In: *Phys. Rev.* 52.4 (1937), pp. 378–379. DOI: 10.1103/PhysRev.52.378.

[CG16] T. Chen and C. Guestrin. “XGBoost: A scalable tree boosting system”. In: *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data. Min.* (2016), pp. 785–794. DOI: 10.1145/2939672.2939785.

[Chi17] D. Chirkin. *Flasher data-derived ice models*. Tech. rep. 2017.

[Chr11] B. J. Christy. “A search for relativistic magnetic monopoles with the IceCube 22-string detector”. PhD thesis. University of Maryland, USA, 2011.

[CS06] M. Cirelli and A. Strumia. “Cosmology of neutrinos and extra-light particles after WMAP3”. In: *J. Cosmol. Astropart. Phys.* 2006.12 (2006), pp. 013-1–013-18. DOI: 10.1088/1475-7516/2006/12/013.

[Cur94] P. Curie. “Sur la symétrie dans les phénoménes physiques, symétrie d’un champ électrique et d’un champ magnétique”. In: *J. phys. (Paris)* 3.1 (1894), pp. 393–415.

[Der+98] J. Derkaoui et al. “Energy losses of magnetic monopoles and of dyons in the earth”. In: *Astropart. Phys.* 9.2 (1998), pp. 173–183. DOI: 10.1016/S0927-6505(98)00016-4.

[Dir31] P. A. M. Dirac. “Quantised singularities in the electromagnetic field”. In: *Proc. R. Soc. Lond. A* 133 (1931), pp. 60–72. DOI: 10.1098/rspa.1931.0130.

[FC98] G. J. Feldman and R. D. Cousins. “Unified approach to the classical statistical analysis of small signals”. In: *Phys. Rev. D* 57.7 (1998), pp. 3873–3889. DOI: 10.1103/PhysRevD.57.3873.

[Fed+19] A. Fedynitch et al. “Hadronic interaction model SIBYLL 2.3c and inclusive lepton fluxes”. In: *Phys. Rev. D* 100.10 (2019), pp. 103018-1–103018-31. DOI: 10.1103/PhysRevD.100.103018.

[Fey70] R. Feynman. *The distinction of past and future – Chapter 5 of the character of physical law*. Cornell University. 1970.

[Fle+94] R. S. Fletcher, T. K. Gaisser, P. Lipari, and T. Stanev. “SIBYLL: An event generator for simulation of high energy cosmic ray cascades”. In: *Phys. Rev. D* 50.9 (1994), pp. 5710–5731. DOI: 10.1103/PhysRevD.50.5710.

[Flo99] T. Flournoy. *Des Indes à la planète Mars: Étude sur un cas de somnambulisme avec glossolalie*. Slatkine, 1899, pp. 344–345.

[Fre+84] C. G. Freeman, T. I. Quickenden, R. A. J. Litjens, and D. F. Sangster. “Visible and ultraviolet emission from pulse irradiated amorphous and polycrystalline H_2O ice”. In: *J. Chem. Phys.* 81.12 (1984), pp. 5252–5254. DOI: 10.1063/1.447691.

[FZ12] J. A. Formaggio and G. P. Zeller. “From eV to EeV: Neutrino cross sections across energy scales”. In: *Rev. Mod. Phys.* 84.3 (2012), pp. 1307–1341. DOI: 10.1103/RevModPhys.84.1307.

[Gai12] T. K. Gaisser. “Spectrum of cosmic-ray nucleons, kaon production, and the atmospheric muon charge ratio”. In: *Astropart. Phys.* 35.12 (2012), pp. 801–806. DOI: 10.1016/j.astropartphys.2012.02.010.

[GG74] H. Georgi and S. L. Glashow. “Unity of all elementary-particle forces”. In: *Phys. Rev. Lett.* 32.8 (1974), pp. 438–441. DOI: 10.1103/PhysRevLett.32.438.

[Gin40] V. L. Ginzburg. “Quantum theory of light radiation of an electron uniformly moving through a medium”. In: *J. Exp. Theor. Phys.* 10 (1940), pp. 589–600.

[Goo56] I. J. Good. “Some terminology and notation in information theory”. In: *Proc. Inst. Electr. Eng. C* 103.3 (1956), pp. 200–204.

[Ham92] E. Hammer. “Über die Planisphere von Aitow und verwandte Entwürfe, insbesondere neue flächentreue ähnlicher Art”. In: *Petermanns geogr. Commun.* 38 (1892), pp. 85–87.

[Hec+98] D. Heck et al. *CORSIKA: A Monte Carlo code to simulate extensive air showers.* 1998.

[Hin+12] G. E. Hinton et al. *Improving neural networks by preventing co-adaptation of feature detectors.* 2012. arXiv: 1207.0580 [cs.NE].

[Hon+07] M. Honda et al. “Calculation of atmospheric neutrino flux using the interaction model calibrated with atmospheric muon data”. In: *Phys. Rev. D* 75.4 (2007), pp. 043006-1–043006-50. DOI: 10.1103/PhysRevD.75.043006.

[HR03] G. C. Hill and K. Rawlins. “Unbiased cut selection for optimal upper limits in neutrino detectors: the model rejection potential technique”. In: *Astropart. Phys.* 19.3 (2003), pp. 393–402. DOI: 10.1016/S0927-6505(02)00240-2.

[ISO15] ISO Central Secretary. *Information technology — Automatic identification and data capture techniques — QR Code bar code symbology specification.* Standard ISO/IEC 18004:2015. Geneva, CH: International Organization for Standardization, 2015.

[KI14] J. L. Kelley and IceCube Colloboration. “Event triggering in the IceCube data acquisition system”. In: *AIP Conf. Proc.* 1630.1 (2014), pp. 154–157. DOI: 10.1063/1.4902795.

[KL51] S. Kullback and R. A. Leibler. “On information and sufficiency”. In: *Ann. Math. Stat.* 22.1 (1951), pp. 79–86. DOI: 10.1214/aoms/1177729694.

[Koe+13] J.-H. Koehne et al. “PROPOSAL: A tool for propagation of charged leptons”. In: *Comput. Phys. Commun.* 184.9 (2013), pp. 2070–2090. DOI: 10.1016/j.cpc.2013.04.001.

[KYG77] Y. Kazama, C. N. Yang, and A. S. Goldhaber. “Scattering of a Dirac particle with charge *Ze* by a fixed magnetic monopole”. In: *Phys. Rev. D* 15.8 (1977), pp. 2287–2299. DOI: 10.1103/PhysRevD.15.2287.

[Lee+09] C. Lee et al. “Study of the luminescence of H_2O and D_2O ices induced by charged-particle bombardment”. In: *Appl. Surf. Sci.* 255.9 (2009), pp. 4716–4719. DOI: 10.1016/j.apsusc.2008.05.344.

[LM01] D. Lelgemann and A. Marchenko. “The Earth’s mass and the radial density profile derived from seismic data: an application of generalised Darwin’s functions”. In: *DGK Acad. Sci, Theo. Geod. A* 117 (2001), pp. 31–51.

[Ma+11] K. J. Ma et al. “Time and amplitude of afterpulse measured with a large size photomultiplier tube”. In: *Nucl. Instrum. Methods Phys. Res. A* 629.1 (2011), pp. 93–100. DOI: 10.1016/j.nima.2010.11.095.

[Max61] J. C. Maxwell. “XLIV. On physical lines of force”. In: *Philos. Mag.* 21.140 (1861), pp. 281–291.

[Ng04] A. Y. Ng. “Feature selection, L_1 vs. L_2 regularization, and rotational invariance”. In: *Proc. 21st Int. Conf. Mach. Learn.* (2004), pp. 78-1–78-8. DOI: 10.1145/1015330.1015435.

[Par70] E. N. Parker. “The origin of magnetic fields”. In: *Astrophys. J.* 160 (1970), pp. 383–405. DOI: 10.1086/150442.

[Pin+09] J. Pinfold et al. *Technical design report of the MoEDAL experiment*. Tech. rep. 2009.

[Pol04] J. Polchinski. “Monopoles, duality, and string theory”. In: *Int. J. Modern Phys. A* 19.supp01 (2004), pp. 145–154. DOI: 10 . 1142 / S0217751X0401866X.

[Pol18] A. M. Pollmann née Obertacke. “Search for mildly relativistic magnetic monopoles with IceCube”. PhD thesis. University of Wuppertal, Germany, 2018.

[Pol19] A. M. Pollmann née Obertacke. “Enabling a new detection channel for beyond standard model physics with in-situ measurements of ice luminescence”. In: *Proc. 36th Int. Cosmic Ray Conf.* 358 (2019), pp. 983-1–983-8. DOI: 10.22323/1.358.0983.

[Pol20] A. M. Pollmann née Obertacke. “Private Communication”. Uncertanty on measuremend of luminesence at IceCube. 2020.

[Pol74] A. M. Polyakov. “Particle spectrum in the quantum field theory”. In: *J. Exp. Theor. Phys.* 20.6 (1974), pp. 194–195.

[Pre84] J. Preskill. “Magnetic monopoles”. In: *Ann. Rev. Nucl. Part. Sci.* 34 (1984), pp. 461–530. DOI: 10.1146/annurev.ns.34.120184.002333.

[PWC00] P. B. Price, K. Woschnagg, and D. Chirkin. “Age vs depth of glacial ice at South Pole”. In: *Geophys. Res. Lett.* 27.14 (2000), pp. 2129–2132.

[QTS82] T. I. Quickenden, S. M. Trotman, and D. F. Sangster. “Pulse radiolytic studies of the ultraviolet and visible emissions from purified H_2O ice”. In: *J. Chem. Phys.* 77.8 (1982), pp. 3790–3802. DOI: 10.1063/1.444352.

[Qui+91] T. I. Quickenden et al. “The effect of temperature on the luminescence from electron-irradiated H_2O ice”. In: *J. Chem. Phys.* 95.12 (1991), pp. 8843–8852. DOI: 10.1063/1.461217.

[Rie+16] F. Riehn et al. “A new version of the event generator SIBYLL”. In: *Proc. 34th Int. Cosmic Ray Conf.* 236 (2016), pp. 558-1–558-8. DOI: 10.22323/1.236.0558.

[Ron19] M. Rongen. “Calibration of the IceCube Neutrino Observatory”. PhD thesis. RWTH Aachen University, Germany, 2019. DOI: 10.18154/RWTH-2019-09941.

[RS12] H. Rebel and O. Sima. “Information about high-energy hadronic interaction processes from extensive air shower observations”. In: *Rom. J. Phys.* 57.1-2 (2012), pp. 472–492.

[Rut11] E. Rutherford. “LXXIX. The scattering of α and β particles by matter and the structure of the atom”. In: *Philos. Mag.* 21.125 (1911), pp. 669–688. DOI: 10.1080/14786440508637080.

[SBS84] R. M. Sternheimer, M. J. Berger, and S. M. Seltzer. “Density effect for the ionization loss of charged particles in various substances”. In: *At. Data Nucl. Data Tables* 30.2 (1984), pp. 261–271. DOI: 10.1016/0092-640X(84)90002-0.

[SC05] B. Sherman and W. Cowper. “VlI. An extraordinary case of a costive person, in a letter from Mr B. Sherman to Dr Beeston of Ipswich; with a note on the same by Mr William Cowper, FR S”. In: *Phil. Trans. R. Soc.* 24.302 (1705), pp. 2110–2112. DOI: 10.1098/rstl.1704.0092.

[Sch90] R. E. Schapire. “The strength of weak learnability”. In: *Mach. Learn.* 5 (1990), pp. 197–227. DOI: 10.1007/BF00116037.

[Sco13] R. F. Scott. *Scott’s last expedition.* Oxford University Press, 1913, p. 177.

[Sco79] D. W. Scott. “On optimal and data-based histograms”. In: *Biometrika* 66.3 (1979), pp. 605–610. DOI: 10.1093/biomet/66.3.605.

[Smi66] S. Smiles. *Self-Help; with illustrations of character and conduct.* 1866, p. 339.

[SSH72] H. B. Steen, O. I. Sørensen, and J. A. Holteng. “Observation of x-ray- and U.V.-induced luminescence and absorption during and after the irradiation”. In: *Int. J. Radiat. Phys. Chem.* 4.1 (1972), pp. 75–86. DOI: 10.1016/0020-7055(72)90010-1.

[SSM06] U. Seljak, A. Slosar, and P. McDonald. “Cosmological parameters from combining the Lyman-α forest with CMB, galaxy clustering and SN constraints”. In: *J. Cosmol. Astropart. Phys.* 2006.10 (2006), p. 14. DOI: 10.1088/1475-7516/2006/10/014.

[Ste81] R. M. Sternheimer. “General expression for the density effect for the ionization loss of charged particles”. In: *Phys. Rev. B* 24.11 (1981), pp. 6288–6291. DOI: 10.1103/PhysRevB.24.6288.

[t H74] G. ’t Hooft. “Magnetic monopoles in unified gauge theories”. In: *Nucl. Phys. B Proc. Suppl.* 79.2 (1974), pp. 276–284. DOI: 10.1016/0550-3213(74)90486-6.

[Tan+18] M. Tanabashi et al. “Review of particle physics – Magnetic monopole searches”. In: *Phys. Rev. D* 98.3 (2018), pp. 030001-1823–030001-1825. DOI: 10.1103/PhysRevD.98.030001.

[Tan+20] M. Tanaka et al. “Search for proton decay into three charged leptons in 0.37 megaton-years exposure of the Super-Kamiokande”. In: *Phys. Rev. D* 101.5 (2020), pp. 052011-1–052011-12. DOI: 10.1103/PhysRevD.101.052011.

[TF37] I. E. Tamm and I. M. Frank. “Coherent radiation of fast electrons in a medium”. In: *Proc. USSR Acad. Sci.* 14 (1937), p. 107.

[Tro92] I. I. Trofimenko. “About one of the possible registration methods for magnetic monopoles in water Cherenkov detectors”. preprint. Inst. Nucl. Res. Russ. Acad. Sciences, 1992.

[VG15] R. K. Vinayak and R. Gilad-Bachrach. “DART: Dropouts meet multiple additive regression trees”. In: *Proc. Mach. Learn. Res.* (2015), pp. 489–497.

[WI14] D. Williams and IceCube Collaboration. “Light propagation in the South Pole ice”. In: *AIP Conf. Proc.* 1630.1 (2014), pp. 146–149. DOI: 10.1063/1.4902793.

[Wic+03] S. D. Wick, T. W. Kephart, T. J. Weiler, and P. L. Biermann. “Signatures for a cosmic flux of magnetic monopoles”. In: *Astropart. Phys.* 18.6 (2003), pp. 663–687. DOI: 10.1016/S0927-6505(02)00200-1.

Acknowledgements

I would like to acknowledge the people without whom writing this thesis would have been impossible.
Without my supervisor, **Prof. Dr. Klaus Helbing**, I would have never worked in the field of Beyond the Standard Model of particle physics nor joined the search for magnetic monopoles. Thanks to him, I got the chance to work at the IceCube Neutrino Observatory as a member of an amazing collaboration, which I am grateful for. Not only that, but I was also part of a stunning group here in Wuppertal, you all will be dearly missed.
Furthermore, I want to extend my gratitude to all those who took time out of their days to read my thesis and help me hunt down any typographical or logical error:

- **Yvonne Chim**,
- **Alexander Luke Kyriacou**,
- **Karol Debowski**,
- **Sonja Mayotte née Schröder**, and
- **Dr. Anna Maria Pollmann née Obertacke**.

I am indebted to all of you.
Wholeheartedly, I also want to thank my family who always stood by me and helped me in every way they could. Without you, I wouldn't be here.

Thank you all!

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Arbeit ohne fremde Hilfe und nur unter Verwendung der angegebenen Hilfsmittel und Quellen, selbstständig verfasst zu haben.
Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Arbeiten anderer entnommen sind, habe ich kenntlich gemacht.
Die Dissertation hat weder in der gegenwärtigen, noch in einer anderen Fassung, schon einem anderen Fachbereich einer wissenschaftlichen Hochschule vorgelegen.

Wuppertal, den	Frederik Lauber
Ort, Datum	Unterschrift