Search for TeV-Scale Neutrino Dimuon Events with 10.7 Years of IceCube Data

by

Sourav Sarkar

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Physics
University of Alberta

Abstract

The IceCube neutrino observatory is a gigaton Cherenkov detector located at the geographic south pole. The experiment is designed to detect neutrinos originating from the atmosphere and astrophysical objects over a huge energy range from 10 GeV to 10 PeV. Information about the interacting neutrinos is extracted from the event reconstructions and identification of the event topologies, constituting probes to study high-energy particle physics and astrophysics. Typically, muon neutrinos produce a single muon from their interactions in ice via deep inelastic scattering (DIS) processes. However, in some fraction of the cases, the Standard Model predicts the production of two closely spaced high energy muons from the same neutrino interaction, referred to as dimuon. Detection of these events in IceCube can be a powerful tool to study rare standard model neutrino interactions and search for physics beyond the Standard Model at an energy scale beyond the reach of current accelerator neutrino experiments. However, identifying dimuon events in IceCube is very challenging due to the limited detector resolution, and no such events have been observed prior to this work.

The primary contribution of the dimuon events within the Standard Model framework arises from the charm quark production in neutrino DIS processes. An additional subdominant standard model channel, called neutrino trident production, can also produce dimuon events. This thesis performs an inclusive search for such dimuon events in IceCube. The search analysis uses an event classification method developed based on advanced machine learning algorithms called graph neural networks to identify the dimuon events. The outcome of the event classification is then used to define two overlapping signal regions for the search. The analysis of 10.67 years of IceCube

data used in this work reports the first non-zero observation of dimuon event candidates. A single event is observed in one signal region with a prediction of 0.21 background contamination, resulting in a dimuon signal significance of 1.31σ. The measurement in the second, higher statistics signal region reports the observation of 4 events against a background prediction of 2.13 events. The pure single muon background-only hypothesis is rejected at 1.15σ for this region.

Preface

The work presented in this thesis is my own and is original except where noted below. The use of others' work is cited in the text where relevant.

The IceCube neutrino observatory's operation results from an international research collaboration comprising approximately 300 scientists. The technical details of the detector and neutrino detection principle in Chapter 3 are the works of many collaboration members over the past 20 years. The development of the neutrino dimuon simulation software and the production of the Monte Carlo simulation in Chapter 4 are my original work which depends on several external and collaboration software packages cited in the text. The Monte Carlo production of the background events and the preliminary event selection process in Chapter 4 are primarily developed by Spencer Axani and Christopher Weaver with the help of other members in the collaboration. The classification method in Chapter 5 is independently developed by myself using the software framework for neural network algorithms cited in the text. The literature review in Chapter 2 and the final analysis method and results in Chapter 6 are my original work.

Acknowledgements

Pursuing a Ph.D. is a long journey and is impossible without the guidance, inspiration, and invigoration from many individuals along the way. I want to express my sincere gratitude to my thesis advisor, Prof. Roger Moore, whose astute insights, expert guidance, and unwavering support were invaluable to successfully completing this work. His commitment to excellence was a constant source of inspiration and motivation. The intellectual and personal growth I experienced under his mentorship was truly transformative.

I express my heartfelt appreciation to Prof. Juan Pablo Yanez. The success of this work would not be possible without his erudite ideas and rigorous engagement in overcoming the research challenges. My appreciation also extends to the IceCube and P-ONE local research group at the University of Alberta. Their constructive feedback and incisive critique enriched my thinking and deepened my understanding of the subject matter. The people in the IceCube collaboration are some of the brightest minds I have ever met. I am genuinely thankful for their continuous support and monumental contribution to the field, which shaped the outcome of this research.

Furthermore, I take great pleasure in extending my gratitude to my dear friends, Nick, Catherine, Sebastian, and many others (I apologize in advance for not including everyone individually). Their warm-hearted company made my transition to graduate school extremely smooth. Indulging in intellectual and stimulating conversations and exciting outdoor and leisure activities with them made the arduous journey of completing this thesis much more enjoyable. I shall cherish their support, encouragement, and kindness for a lifetime.

Lastly, I must take a moment to express my profound gratitude and appreciation to one of the most important persons in my life, Michelle Sullivan. Her steadfast belief in my abilities, firm confidence in my potential, and constant source of encouragement have sustained me through the most challenging days of my academic life. She has also been instrumental in shaping the outcome of this thesis through her keen intellect, persistent commitment to my academic success, and unrelenting attention to detail. Her love, support, and patience have been integral to my academic journey, and I am forever grateful for her presence in my life.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The study of neutrino interactions with matter is pivotal for advancing critical concepts in particle physics, astrophysics and cosmology [1–3]. Theoretical models developed in these areas undergo rigorous validation and testing based on the experimental observation of neutrino interactions in the detectors. Conversely, precise measurements performed in neutrino detectors can also lead to advancements in our understanding of fundamental physics [4–6]. Therefore, neutrino experiments offer a rich ground for new discoveries in both fundamental physics and understanding the nature and evolution of the universe.

The IceCube neutrino observatory, located at the geographic south pole, is the world's largest neutrino detector and observes neutrinos originating from atmospheric and astrophysical processes. Detection of neutrinos with a massive energy range ($10\ \text{GeV} - 10\ \text{PeV}$) and a large international collaboration analyzing the experimental data allow IceCube to run numerous scientific programs to study multiple aspects of neutrino physics. The analyses involving the observation of low energy neutrinos ($E_\nu \in [10 - 500\ \text{GeV}]$) primarily focus on neutrino oscillation studies [7–14]. The extremely high energy neutrinos ($E_\nu > 100\ \text{TeV}$) detected in IceCube originate from sources outside the solar system [15, 16]. The analyses of these very high energy neutrinos in IceCube currently lead the field of neutrino astronomy and have made several ground-breaking discoveries, such as the first detection of the Glashow resonance [17] and evidence of neutrino emissions from active galaxies [18, 19]. In the high energy regime ($E_\nu \in [1 - 100\ \text{TeV}]$), IceCube has detected a large number of neutrinos primarily from the interaction of cosmic rays in the earth's atmosphere, with

an increasing contribution from astrophysical neutrinos in the higher energy regime ($E_\nu \gtrsim 10$ TeV). Most of the IceCube analyses based on these high energy neutrino events involve the study of weak interactions and quantum chromodynamics (QCD) within the Standard Model framework [20–22] and the search for new physics [23–25]. Historically, these studies have been typically performed in accelerator neutrino and collider experiments. However, the high and very high energy neutrino interactions in IceCube often exceed the energy limit of the laboratory-based experiments, making it a prime area for further testing of the Standard Model and searching for new physics [26–30].

The work in this thesis focuses on identifying a new class of events in IceCube called dimuons, where an interacting neutrino produces two outgoing, high energy muons. The dimuon events in the Standard Model neutrino interactions mainly come from two processes, namely charm and trident production. They have been observed in accelerator neutrino experiments with an approximate energy range of $10 - 100$ GeV, providing important implications for studying weak interactions, QCD physics, and new physics searches [31–38]. The observation of TeV-scale dimuons in IceCube presents an opportunity to perform similar studies in a kinematical region beyond the energy scale of current accelerator neutrino experiments [39–43]. In Chapter 2, we briefly introduce the neutrino interactions in the Standard Model and provide a detailed discussion of the theory describing the dimuon production processes. The use of neutrino trident dimuon events as a probe to search for new physics is also reviewed in the chapter. The details of the neutrino detection mechanism in IceCube must be well understood for developing the dimuon search analysis presented in this work. Therefore, we provide a concise overview of the neutrino detection principle, the identification of different event topologies, and the sources of detected neutrinos in IceCube in Chapter 3.

The first step towards the analysis development requires investigating the dimuon event properties, such as the event kinematics and the expected event rates in the detector. This is achieved by developing the charm and trident Monte Carlo (MC) event generation framework and simulating the detector response for the generated events, presented in Chapter 4. The chapter also discusses the simulation of background processes and the preliminary event selection procedure relevant to the anal-

ysis. Identifying two closely spaced muons in dimuon events is extremely challenging in IceCube due to the limited detector resolution. A novel event classification method based on advanced machine learning (ML) algorithms is developed to achieve the best possible identification for these events and is described in Chapter 5. The output of the event classification then defines the final analysis regions, which are used in the statistical dimuon search analysis. The steps involving the selection of the analysis regions and the application of the statistical tests are discussed in Chapter 6. The results of this work are the first measurements of the dimuon search in IceCube and are reported in the later section of the chapter. Following the results, concluding remarks and the future implications of this work are discussed in Chapter 7.

Chapter 2

Theory of Neutrino Interactions

The discovery of the neutrino by the Reines-Cowan experiment [44] in 1956 marked the beginning of the study of experimental neutrino physics. Although Wolfgang Pauli first postulated neutrino's existence in 1930 [45] and later formally hypothesized by Enrico Fermi [46] to explain the missing energy in beta decays, it took 20 years to detect the particle due to its rare interaction with matter. Since its discovery almost 70 years ago, neutrinos have significantly contributed to the development of particle physics, notably the discovery of W and Z bosons and the unification of electroweak interactions [47]. The field of neutrino physics is still rapidly evolving to study some of the most fundamental concepts in understanding the nature of matter and the origins of the universe. This chapter will discuss the details of the neutrino interactions within the Standard Model framework, and the critical roles neutrinos can play in understanding physics beyond the Standard Model (BSM).

2.1 Particles in the Standard Model

The Standard Model of elementary particles has been developed by combining the theories of special relativity and quantum mechanics. In this framework, all particles are interpreted as the excitations of relativistic quantum fields, and the interactions between the particles dictated by the fields' behaviour define the Standard Model Lagrangian. The Lagrangian describes the particles that constitute matter (called *fermions*) and three of the fundamental forces of nature (strong, electromagnetic and weak force) using a unified gauge symmetry. A summary diagram of the particles is shown in Figure 2.1. The forces arise from the spin-1 vector fields, and the quantiza-

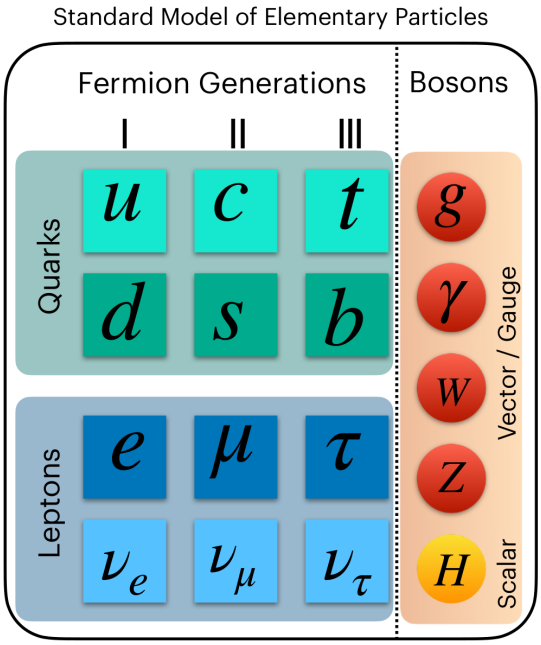


Figure 2.1: The illustration showing all the matter and force particles in the Standard Model framework.

tion of those fields creates the force carriers called *gauge bosons*. The electromagnetic interaction is mediated by the *photon*(γ), the weak interaction is mediated by three bosons - $W^{\pm}$ and Z, and the strong interaction is mediated by *gluons*(g). In addition, *higgs bosons*(H) in the Standard Model are spin-0 particles from the quantum excitation of the Higgs field, which provides a mechanism for the elementary particles to acquire their masses. Table 2.1 summarizes the properties of the gauge and higgs bosons.

As discussed above, fermions are matter particles with spin-$\frac{1}{2}$ and are subdivided into two groups - *quarks* and *leptons*. Quarks interact via all three forces and are bound by the strong interaction to form colourless composite particles called *hadrons*. The hadrons containing three quarks are called *baryons* (such as proton and neutron) and have half-integer spins (e.g. $\frac{1}{2}, \frac{3}{2}, ...$). The hadrons with the bound states of quark and anti-quark pairs are called *mesons* (such as pions and kaons) and have integer spins (e.g. $0, 1, 2, ...$). Based on the observations in collider experiments, the Standard Model provides an inclusive list of six quarks in three quark-doublet configurations. Each doublet (pair of quarks) belongs to one of the three *generations* of fermions. Quark doublets across the generations have identical properties in strong

Particle	Mass ($\mathrm{GeV/c^2}$)	Charge (e)	Spin ($\hbar$)	Decay Width (GeV)
γ	0	0	1	stable
g	0	0	1	stable
$W^{\pm}$	80.385 ± 0.015	± 1	1	2.085 ± 0.042
Z	91.1876 ± 0.0021	0	1	2.4952 ± 0.0023
H	125.09 ± 0.24	0	0	< 1.7 (95% CL)

Table 2.1: Properties of the Standard Model bosons. The measurements are from Ref. [48]. The decay width (Γ) is defined as the inverse of the mean lifetime (τ) of the particles, $\Gamma = 1/\tau$.

and electromagnetic interactions and only differ in mass and flavour quantum numbers. The doublets denoted by (*Up(u)*, *Down(d)*), (*Charm (c)*, *Strange (s)*), and (*Top(t)*, *Bottom (b)*) quarks form the first, second, and third generation of quark pairs respectively. A summary of all the quarks is shown in Table 2.2.

Generation	Particle	Mass	Charge (e)
1	u	$2.2^{+0.6}_{-0.4}$ MeV/c^2	$+\frac{2}{3}$
	d	$4.7^{+0.5}_{-0.4}$ MeV/c^2	$-\frac{1}{3}$
2	c	1.27 ± 0.03 GeV/c^2	$+\frac{2}{3}$
	s	96^{+8}_{-4} MeV/c^2	$-\frac{1}{3}$
3	t	173.21 ± 0.71 GeV/c^2	$+\frac{2}{3}$
	b	$4.66^{+0.04}_{-0.03}$ GeV/c^2	$-\frac{1}{3}$

Table 2.2: Summary of the Standard Model quarks. The measurements are from Ref. [48].

Similarly, there are three lepton doublets that belong to each generation of fermions. Each pair has a charged lepton (*electron(e), muon*(μ), *tau*(τ) for the three generations, respectively) and a neutral lepton (collectively called *neutrinos*(ν)). Unlike quarks, leptons do not participate in the strong interaction. The charged leptons

Generation	Particle	Mass	Charge (e)	L_e	L_μ	L_τ
1	e^-	0.511 MeV/c^2	-1	$+1$	0	0
	ν_e	< 2 eV/c^2	0	$+1$	0	0
2	μ^-	105.66 MeV/c^2	-1	0	$+1$	0
	ν_μ	< 2 eV/c^2	0	0	$+1$	0
3	τ^-	1.777 GeV/c^2	-1	0	0	$+1$
	ν_τ	< 2 eV/c^2	0	0	0	$+1$

Table 2.3: Summary of the Standard Model leptons. The measurements are from Ref. [48].

interact via electromagnetic and weak forces, and the neutrinos (being neutral leptons) participate only in the weak interaction. The summary of the leptons is shown in Table 2.3. Three lepton flavour quantum numbers (L_e, L_μ, L_τ) related to the three generations are assigned to each lepton and are postulated to be conserved in all the leptonic interactions within the Standard Model framework. On the contrary, the framework allows for violating quark flavour quantum number conservation in weak interactions such as heavier meson ($K^\pm, D^\pm, B^\pm$) decays. The mechanism of the quark flavour mixing is formulated by Cabibbo, Kobayashi and Maskawa using a mixing matrix called CKM matrix [49, 50]. The matrix elements dictate the strength of the weak interactions resulting in the transitions between two corresponding generation of quark flavours.

All the fermions discussed above are accompanied by the corresponding antiparticles that carry the opposite sign of the charges and flavour quantum numbers and participate in the conjugate interactions compared to their counterparts. Precise experimental measurements of the particle properties over several decades have cemented the tremendous success of the Standard Model in describing the quark and lepton interactions [48]. However, the discrepancy in the solar neutrino flux measurement by the Homestake experiment [51] (known as the 'solar neutrino problem') was the first profound experimental evidence that the description of the three forces in the Standard Model is incomplete. The problem was later resolved by the discovery of

the phenomenon of *neutrino oscillations* [4, 5], similar to the flavour mixing observed in the quark sector.

2.2 Neutrino Oscillation

The neutrinos are treated as massless particles in the Standard Model. However, neutrino oscillation requires that the neutrinos are massive and that their mass eigenstates (denoted as $\nu_i \in \{\nu_1, \nu_2, \nu_3\}$) are not the same as their flavour eigenstates (denoted as $\nu_\alpha \in \{\nu_e, \nu_\mu, \nu_\tau\}$). A neutrino of certain flavour ν_α at production (time $t = 0$) can be expressed in terms of mixing of the mass eigenstates as

$$|\nu_\alpha\rangle = \sum_i U_{\alpha i} \, |\nu_i\rangle \quad , \tag{2.1}$$

where $U_{\alpha i}$ are the elements of the 3×3 ($\alpha = e, \mu, \tau$ and $i = 1, 2, 3$) unitary PMNS matrix $\boldsymbol{U}$, named after Pontecorvo, Maki, Nakagawa, and Sakata [52, 53]. Using the properties of the unitary matrix, the elements of $\boldsymbol{U}$ can be parameterized in terms of the three rotational angles (also known as *mixing angles*) - $\theta_{12}, \theta_{23}, \theta_{13}$ and a CP-violating phase δ_{CP}. The subscript of the mixing angles denotes the angle between corresponding mass eigenstates. The elements of the PMNS matrix can then be expressed as the transformation factors from three rotation matrices and a global phase (δ_{CP}),

$$\boldsymbol{U} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \cdot \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \cdot \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \tag{2.2}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$. When the neutrino of flavour ν_α travels in the vacuum after its production, the state after time t can be expressed using the time-dependent propagation of its mass eigenstates with energy $E_i (i = 1, 2, 3)$,

$$|\nu(t)\rangle = \sum_i U_{\alpha i} \; e^{-iE_i t} \, |\nu_i\rangle \quad . \tag{2.3}$$

In the relativistic limit,

$$E_i \approx p + \frac{m_i^2}{2p} \quad , \tag{2.4}$$

where p is the momentum of the neutrino and m_i is the mass of the corresponding mass eigenstate. During the propagation, the phases of the mass eigenstates vary due to different masses, which leads to the mixing of the neutrino flavours for the state $\nu(t)$. The probability of detecting the neutrino as flavour ν_β after time t can then be expressed as,

$$P(\nu_\alpha \to \nu_\beta) = |\langle \nu_\beta | \nu(t) \rangle|^2 = \left| \sum_i U_{\alpha i} U^*_{\beta i} e^{-iE_i t} \right|^2 \quad . \tag{2.5}$$

Assuming the difference between the energy eigenvalues in terms of the masses as $\Delta m^2_{ij} = m^2_i - m^2_j$ and L as the distance travelled by the neutrino in time t, we can expand the above equation as,

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 4 \sum_{i>j} \mathbf{Re} \left[U^*_{\alpha i}\, U_{\alpha j}\, U_{\beta i}\, U^*_{\beta j} \right] \sin^2 \left(\frac{\Delta m^2_{ij} L}{4E} \right) + 2 \sum_{i>j} \mathbf{Im} \left[U^*_{\alpha i}\, U_{\alpha j}\, U_{\beta i}\, U^*_{\beta j} \right] \sin \left(\frac{\Delta m^2_{ij} L}{2E} \right) \quad . \tag{2.6}$$

Here E is the energy of the neutrino, and **Re**, **Im** are the real and imaginary parts of the matrix element multiplications, respectively. In the presence of CP invariance (i.e. $\delta_{CP} = 0$), the imaginary term vanishes, and the flavour transition probability can be expressed using only the first two terms in the above equation.

From Equation (2.6), we see that the experiments studying the neutrino oscillation depend on the relative differences in neutrino masses (i.e. Δm^2_{ij}), and the absolute mass measurement requires different experimental design and effort [6]. In addition, the L/E term in Equation (2.6) dictates the phase of the oscillation and varies for the neutrino sources with different scales of the energy (E) and propagation distance (L). Therefore, the experiments detecting neutrinos from different sources are sensitive to measuring different oscillation parameters. Solar neutrino experiments like SNO [4], Borexino [54], and Super-Kamiokande (SK) [55] detect MeV neutrinos from the sun and are sensitive to the measurement of the mixing angle θ_{12} and Δm^2_{21}. The experiments detecting GeV-scale neutrinos from the atmosphere (SK [56], IceCube [13]) and the accelerators (MINOS [57], T2K [58], NOvA [59]) are designed to measure Δm^2_{31} and θ_{23}. Several reactor neutrino experiments like Daya Bay [60] and Double

Oscillation Parameter	Best Fit $\pm$ 1σ
$\sin^2\theta_{12}$	$0.307^{+0.013}_{-0.012}$
$\sin^2\theta_{23}$	$0.565^{+0.025}_{-0.120}$
$\sin^2\theta_{13}$	$0.02195^{+0.00075}_{-0.00074}$
δ_{CP} $(^0)$	228^{+51}_{-33}
$\frac{\Delta m^2_{21}}{10^{-5}eV^2}$	$7.40^{+0.21}_{-0.20}$
$\frac{\Delta m^2_{31}}{10^{-3}eV^2}$	$+2.515^{+0.035}_{-0.035}$

Table 2.4: Global fit results of neutrino oscillation parameters from Ref. [64], assuming Normal Ordering (NO).

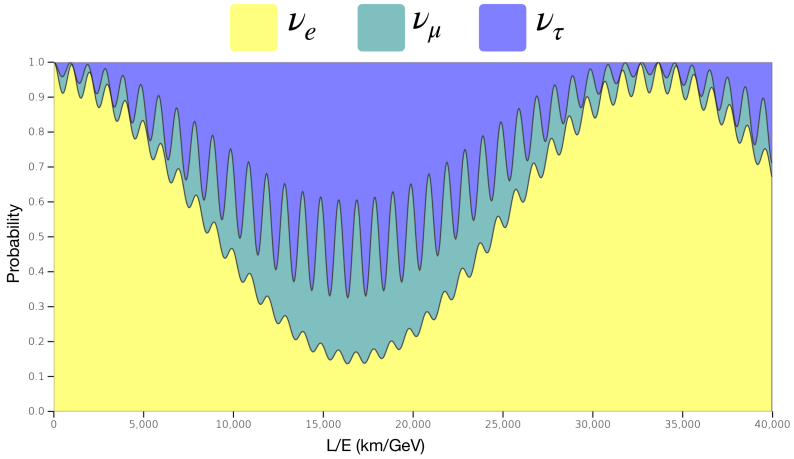


Figure 2.2: Survival probability of the three flavours due to neutrino oscillation in vacuum. The plot is generated with an initial ν_e flux using Ref. [65] for the NO scenario.

Chooz [61] have the sensitivity to measure the mixing angle θ_{13}. In addition, the CP-violating phase δ_{CP} can be measured from long-baseline (LBL) neutrino oscillation experiments [62]. Table 2.4 summarizes the NuFit [63, 64] global fit values of the six oscillation parameters used in this work. An example of three-flavour neutrino vacuum oscillation probability for an initial ν_e flux is shown in Figure 2.2 as the function of L/E.

Precise knowledge of the leptonic flavour mixing parameters is essential to further developing our understanding of the true nature of the particles [66, 67]. The values of

the parameters (mass-squared splittings and the mixing matrix) presently come from observation and have several unknowns related to the phenomenon. The absolute value of Δm^2_{31} (also known as *atmospheric mass splitting*) is extracted with high precision from the atmospheric neutrino oscillation measurements. However, the sign of the atmospheric mass splitting is still unknown and is referred to as the neutrino *mass ordering* problem. This leads to two possible scenarios, $\Delta m^2_{31} > 0$ (known as *Normal Ordering (NO)*) and $\Delta m^2_{31} < 0$ (known as *Inverted Ordering (IO)*). In addition, the current measurements of the CP-violating phase δ_{CP} have significant errors and cannot infer if there is CP-violation ($\delta_{CP} \neq 0$) in neutrino oscillations. Therefore, the experimental study of neutrino oscillation remains a highly active field of research to investigate the properties of neutrinos which can lead to the search for BSM physics signatures [68, 69]. Apart from the neutrino oscillation, the interaction of neutrinos with matter can also be used as a probe to study several BSM scenarios, as discussed later in Section 2.7.

2.3 Neutrino Deep Inelastic Scattering

The success of the neutrino oscillation experiments relies on the flavour tagging of the interacting neutrinos in the detector medium. When an (anti-)neutrino interacts via the exchange of a W boson (referred to as charged current (CC) interaction), it produces an outgoing charged lepton of the same flavour ($e^{\pm}, \mu^{\pm}, \tau^{\pm}$). Identification of the charged lepton is then used to infer the flavour of the interacting neutrino. Neutrinos of all flavours also participate in the neutral current (NC) interaction via the exchange of the Z boson and produce an outgoing neutrino of the same type. Detection of these neutrino interactions in IceCube occurs at energies $E_\nu \gtrsim 10\ GeV$, where the exchanged weak bosons couple to the individual quarks inside the target nucleons (proton, neutron), as shown in Figure 2.3. The outgoing quark from these interactions can carry a significant fraction of the incoming neutrino energy to leave the bound state of the target nucleon. Due to the colour confinement, the individual quark then recombines with the rest of the nucleon through the *hadronization* process to produce multiple higher-energy outgoing hadrons. As these interactions result in the shattering of the target nucleons and probe into the hadronic structure, they are

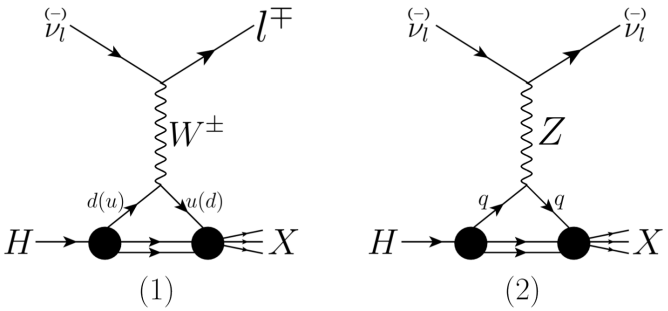


Figure 2.3: The example feynman diagrams showing the neutrino CC (1) (coupling to u or d quark) and NC (2) DIS processes (coupling to any quark q). Here $l = e, \mu, \tau$ denotes the different lepton flavours of the same interaction process, and H is the target hadron.

known as neutrino deep inelastic scattering (DIS). Both CC and NC DIS reactions are shown below,

$$\begin{aligned} \textbf{CCDIS}: &\quad \nu_l(\bar{\nu}_l) + H \longrightarrow l^-(l^+) + X \\ \textbf{NCDIS}: &\quad \nu_l(\bar{\nu}_l) + H \longrightarrow \nu_l(\bar{\nu}_l) + X \quad , \end{aligned} \tag{2.7}$$

where H is the target nucleon (proton or neutron) and X is the cumulative set of outgoing hadrons. The processes occur for any flavour of $l = e, \mu, \tau$. In the CC DIS interaction, the incoming neutrino deposits most of its energy in the detector medium by producing a charged lepton and jets of hadrons. However, in the NC DIS interactions, the outgoing neutrino escapes the detector without further interaction, and only a fraction of the incoming neutrino energy is deposited through the outgoing hadrons (same for all flavours) in the medium. Therefore, the detection of CC DIS processes provides more information about the interaction (such as neutrino flavour and energy) compared to the NC DIS processes.

The constituents of the hadrons are modelled using the quantum chromodynamics (QCD) framework and consist of real quarks (known as *valence quarks*) and virtual particle-antiparticle pairs (known as *sea quarks* and *gluons*). The interacting quark in the neutrino DIS interactions can come from either valence or sea quarks of the target nucleon. Several experiments [70–75] have performed a detailed observation of the neutrino DIS processes to study the quark and gluon content (together referred to as *partons*) of the target nucleons. The probability distributions for the fractions of the

hadron's momentum carried by the partons are called *parton distribution functions* (PDFs) and define the parton content of the hadron. Therefore, the PDFs are one of the key components in calculating the total and differential cross sections of the neutrino DIS processes.

For the general form of the DIS process shown in Equation (2.7), let us denote the 4-momenta of the incoming neutrino and the target as k and p, respectively. Let us also define the 4-momenta of the outgoing lepton and the hadronic system as k' and p', respectively. The 4-momentum of the virtual gauge boson can then be computed as $q = k - k'$. We can now introduce a few useful Lorentz invariant variables, namely the Bjorken scaling variable x, inelasticity y, and the momentum transfer squared Q^2. They are defined as,

$$x = -\frac{q^2}{2p \cdot q}, \quad y = \frac{p \cdot q}{p \cdot k}, \quad Q^2 = -q^2 \quad . \tag{2.8}$$

The inelasticity can be expressed in a simpler form as $y = \frac{E_{had}}{E_\nu}$, where E_ν and E_{had} are the energy of the incoming neutrino and the outgoing hadrons, respectively and represents the fraction of incoming neutrino energy transferred to the hadronic system. The general form of the neutrino DIS differential cross section on an isoscalar (average of proton and neutron) nucleon target in terms of these measurable quantities is expressed as [76],

$$\frac{\mathrm{d}^2\sigma^{\nu N}_{CC,NC}}{\mathrm{d}x\ \mathrm{d}Q^2} = \frac{G_F^2 M_{W,Z}^4}{4\pi(Q^2 + M_{W,Z}^2)^2\ x} \left[Y_+ F_2(x,Q^2) - y^2 F_L(x,Q^2) \pm Y_- x F_3(x,Q^2)\right] \ , \tag{2.9}$$

where G_F is the Fermi coupling constant, $M_{W,Z}$ is the mass of the corresponding interacting weak boson in the CC and NC processes, and $Y_\pm = 1 \pm (1-y)^2$. The terms F_2, F_3, F_L are known as *structure functions* and, as the name suggests, contain information about the structure of the target nucleon. These functions are directly related to the PDFs and are calculated using the perturbative QCD methods discussed in Ref. [77–79]. The level of approximation in the calculation is dictated by the power of the QCD coupling constant $\alpha_s(Q^2)$, resulting in the use of *Leading-Order* (LO), *Next-to-Leading-Order* (NLO), and *Next-to-Next-Leading-Order* (NNLO) approximations of the PDFs. Several collaborations like CTEQ [80], HERA [81], NNPDF [82],

and MSWT [83] have performed extensive analyses to produce the PDFs using data from collider and accelerator neutrino experiments. The differential cross section is then derived using the PDF measurements, and the total cross section is calculated by integrating the differential cross section. The total cross section of neutrino CC and NC DIS processes as a function of neutrino energy (E_ν) is calculated by CSMS [84] using HERA1.5 PDFsets [81] and is shown in Figure 2.4.

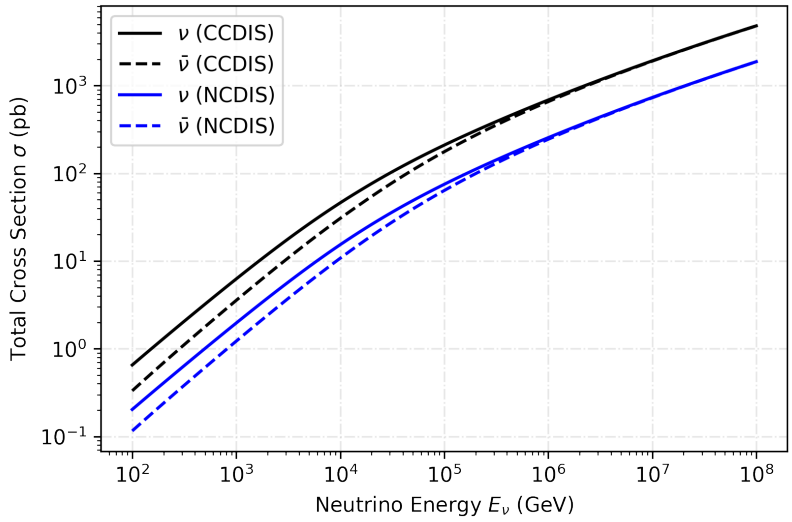


Figure 2.4: Total cross section of neutrino (solid line) and antineutrino (dashed line) CC and NC DIS interactions calculated by CSMS [84].

2.4 Charm Production

Identification of the heavy quark production (such as charm, top) in the neutrino DIS processes is useful for further testing of the PDFs and the development of QCD physics [85, 86]. Moreover, a study of the heavy quark production in neutrino DIS processes can directly probe into the CKM mixing matrix by observing the weak decay of the quarks. The production of heavy quarks requires a higher neutrino energy threshold due to heavier masses of the outgoing partons (m_{charm} = 1.3 GeV/c^2, m_{top} = 173 GeV/c^2). In particular, the CCFR [32], CHARM-II [31], NuTev [33], CHORUS [34], and NOMAD [35] experiments have measured the charm production in neutrino DIS events with mean incoming neutrino energy $\langle E_\nu \rangle$ in $\mathcal{O}(10 - 100$ GeV). The most prominent channel to observe charm production is where an incoming muon neutrino

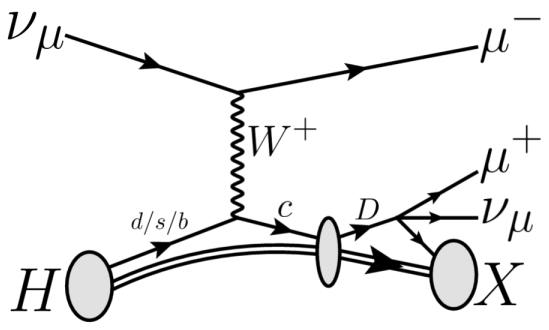


Figure 2.5: Feynman diagram showing the neutrino CC DIS interaction producing outgoing charm quark and decaying into a secondary muon from the reaction.

(ν_μ) interacts with an initial-state $d/s/b$ quark (either from the valence (d) or sea (d, s, b) quark of the nucleon) via CC DIS and produces an outgoing muon and a charm quark. The contribution from the interacting b quark is vanishingly small due to its large mass and strong CKM suppression in the quark mixing. After the hadronization process, the charm hadron carries a large fraction of the invariant mass of the hadronic system, and $\sim$10% of the time decays semileptonically into a muon. The outgoing muon from the primary neutrino interaction and the secondary muon from the charm hadron decay produce an observable dimuon event signature in the detector. This particular process is called ν-induced charm dimuon production and is shown in Figure 2.5. The fraction of the neutrino CC DIS events producing the charm quark as a function of the incoming neutrino energy is also shown in Figure 2.6. According to the CKM matrix, the strange quark dominates the contribution to the charm production at low Bjorken x [35]. So the observation of dimuons offers a direct probe to measure the strange quark content of the nucleons.

IceCube's ability to detect very high-energy neutrinos creates a unique opportunity to study the DIS processes at an energy scale that is not achievable by any accelerator-based neutrino experiments [21, 88, 89]. In recent years, the possibility of observing ν-induced charm dimuon events in IceCube has drawn more attention [42, 43]. The detection of such TeV-scale dimuon events in IceCube can potentially be used to measure the strange quark PDF at a much larger factorization scale Q than any accelerator neutrino experiments [90]. In addition, as we will see in later sections,

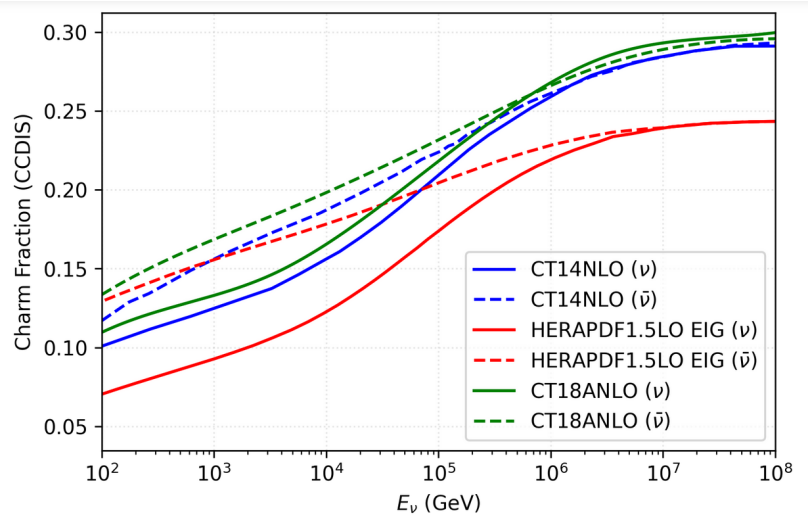


Figure 2.6: Comparison of the charm fraction in CC DIS interaction, computed using different PDFsets (CT18ANLO [80], CT14NLO [87], and HERAPDF1.5LO EIG [81]) and the CSMS [84] total cross section.

the charm dimuons become a background in the search for other Standard Model and BSM dimuon processes. So the accurate measurement of the charm dimuon cross section is necessary to be able to observe these additional processes.

2.5 Glashow Resonance

The scattering cross section of neutrinos with the electrons in the detector medium is usually very negligible due to the small electron mass (m_e). However, when an electron antineutrino interacts with an electron via CC, it can have an s-channel contribution, as shown in Figure 2.7. This process allows the production of on-mass shell W-boson, and the cross section is enhanced due to the resonance (known as *Glashow Resonance*) [91]. The requirement of centre-of-mass (CM) energy of the system to be higher than the mass of W-boson ($M_W = 80.4$ GeV) is achieved at the neutrino energy $E_\nu = M_W^2/2m_e = 6.3$ PeV. The real W-boson can then decay into either hadrons ($W \to q\bar{q}$ with branching ratio (BR) 67.41%) or leptons ($W \to \bar{\nu}_l l^-$ with BR 32.59%). In 2016, IceCube detected an event that can be attributed to the Glashow resonance [17]. The quarks from the hadronic decay of W-boson in this process can produce two high-energy muons creating a dimuon event signature. However, the requirement of leptonic decays (producing muons) for both the quarks

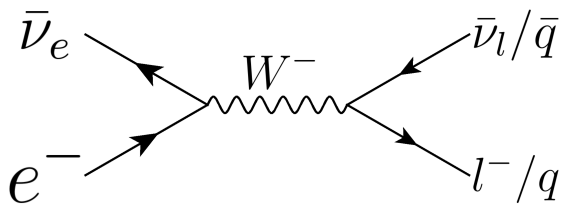


Figure 2.7: Feynman diagram showing the Glashow resonance process. Both the leptonic and hadronic decay of W boson are shown as the outgoing particles.

and the small neutrino flux at this energy scale, i.e. observation of one Glashow resonance event in ~ 10 years lead to a vanishingly small rate of such dimuon events in IceCube. Therefore, the production of dimuon events from Glashow resonance is not considered in this analysis.

2.6 Neutrino Trident Production

The ν-induced charm production and the Glashow resonance are not the only neutrino interaction processes within the Standard Model framework that can produce dimuon event signatures in the detector. Neutrino *trident* production is a sub-dominant Standard Model interaction that can also produce such an event topology. In trident interactions, an incoming neutrino interacts in the Coulomb field of a nucleus and produces two outgoing charged leptons, an outgoing neutrino, and a recoiled nucleus,

$$\nu_\alpha + H \longrightarrow \nu_{\alpha(\beta)} + l^+_{\alpha(\beta)} + l^-_{\alpha(\beta)} + X \quad , \tag{2.10}$$

where H, X are the initial and final state nuclei respectively, and $l^+_{\alpha(\beta)}, l^-_{\alpha(\beta)}$ are the charged leptons with flavours $\alpha, \beta \in \{e, \mu, \tau\}$ obeying the conservation of lepton flavour numbers of the process. In the context of dimuon production, we focus on the trident processes where both the outgoing charged leptons are muons. These neutrino interactions have been historically important in developing the weak interaction theory and studying the $W - Z$ interference [92–96]. The Feynman diagrams of one such example trident interaction are shown in Figure 2.8, where an incoming muon neutrino (ν_μ) interacts via both CC and NC channels and produces the outgoing leptons of the same flavour (referred to as *CC+NC* channel). Two additional channels, called *CC only* and *NC only* involve in the trident interactions via the exchange of

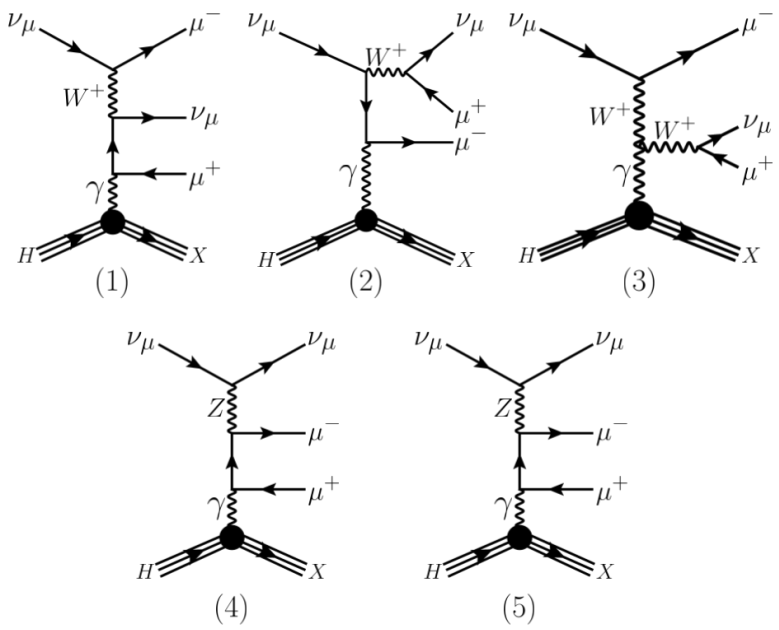


Figure 2.8: The Feynman diagrams for ν_μ CC+NC channels. The diagrams (2) and (3) have s-channel W-boson production and dominate above ∼6 TeV of neutrino energy.

only $W^\pm$ and Z bosons, respectively and produce different non-overlapping flavour combinations of the outgoing leptons, $\{\nu_{\alpha(\beta)}, l^-_{\alpha(\beta)}, l^+_{\alpha(\beta)}\}$. The outgoing particles from the three channels of ν_μ trident interactions are shown below,

$$\begin{aligned}
&\textbf{CC+NC}: \nu_\mu + H \longrightarrow \nu_\mu + \mu^- + \mu^+ + X \\
&\textbf{CC Only}: \nu_\mu + H \longrightarrow \nu_{e(\tau)} + \mu^- + e^+(\tau^+) + X \\
&\textbf{NC Only}: \nu_\mu + H \longrightarrow \nu_\mu + e^-(\tau^-) + e^+(\tau^+) + X \quad .
\end{aligned} \tag{2.11}$$

For this work and the rest of the thesis, we will refer to the $\nu_\mu(\bar{\nu}_\mu)$ CC+NC channel as the trident interaction, as this is the most dominating trident process producing dimuon events. In diagrams (2) and (3) of Figure 2.8, W bosons are produced via s-channel. If the center-of-mass (CM) energy of the neutrino-nucleus system is above the W-boson mass limit (occurring at $E_\nu \gtrsim 6$ TeV), the trident interaction is significantly enhanced due to on-shell W-boson production. Therefore, the trident processes can probe into the production of real W boson at much lower neutrino energy than the Glashow resonance [97, 98].

Based on the photon momentum transfer squared (Q^2), the trident interaction can

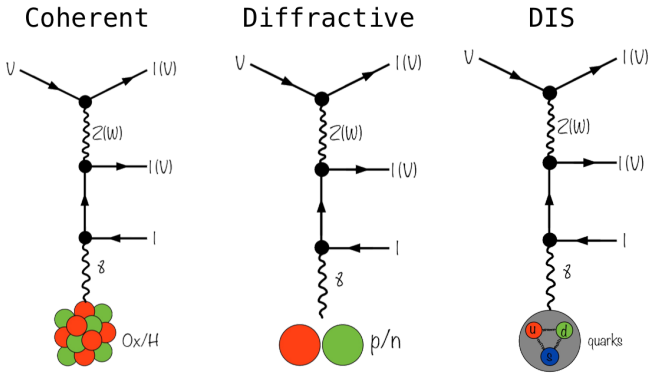


Figure 2.9: Diagrams showing the different interaction regimes in neutrino trident interaction.

be divided into three different nuclear regimes: coherent, diffractive, and deep inelastic (DIS) as illustrated in Figure 2.9. Calculation of the total and differential cross sections for all three regions is needed to produce a Monte Carlo (MC) simulation of these events. In the past, the trident cross section is calculated using approximation methods. For the low energy (MeV to sub-GeV) neutrinos, Fermi four-point approximation is applied in calculating the cross section [99]. This method simplifies the calculation by ignoring the presence of the explicit weak propagator terms and assuming a single interaction vertex for the four leptons (incoming neutrino and three outgoing leptons). The approximation works well for the reactor and accelerator neutrino experiments with incoming neutrino energy much below the W-boson production threshold. However, this work explores the high-energy neutrino interactions ($E_\nu \geq 100\ GeV$) where the Fermi four-point approximation breaks down due to the lack of weak interaction propagators in the calculation. Another approximation method called equivalent photon approximation (EPA) simplifies the calculation for the coherent and diffractive regimes by decoupling the interaction into two parts: the neutrino-photon interaction and the nuclear effects (from the nucleus (coherent) and individual nucleons (diffractive) as shown in Figure 2.9) producing that interacting photon [39, 100, 101]. For the neutrino-photon interaction $\nu + \gamma \rightarrow \nu + \mu^- + \mu^+$, the total cross-section is denoted as $\sigma_{\nu\gamma}(s)$, and the CM energy s is expressed in terms of the neutrino energy E_ν and photon momentum transfer Q as $s = 2E_\nu Q$. The approximation of the total trident cross section ($\sigma_{\nu H}$) can then be calculated as the

neutrino-photon cross section weighted by the probability $P(s, Q^2)$ of the nucleus/nucleon producing a virtual photon with virtual-mass squared Q^2,

$$\sigma_{\nu H} = \int ds\ \sigma_{\nu\gamma}(E_\nu, Q) \int dQ^2\ P(s, Q^2) \quad . \tag{2.12}$$

The probability distribution $P(s, Q^2)$ in the above equation is evaluated from the nuclear form factor for each elastic regime (coherent and diffractive). For the DIS regime, the cross section is calculated using the convolution of the parton cross section with the quark PDFs (d, u, c, s). To avoid double-counting, artificial cuts on the photon momentum transfer are imposed to restrict the overlap of the three regimes. However, with the EPA method, the total cross section is reported to be overestimated by as much as 200% [102].

Recent work in [40, 103] reports an extensive calculation of the trident total cross sections free from any approximation method and indicates a promising number of trident dimuon events in IceCube. The total cross sections of all incoming neutrino flavours and interaction channels (CC+NC, CC only, NC only) are shown in Figure 2.10, along with a comparison to the standard neutrino CC DIS cross section. However, the work still lacks an open tool to calculate total and differential cross sections that are important to build an MC generator for simulating trident events at high energies. In the first part of this work, we have developed a general-purpose MC trident event generator, which will be discussed in detail in Chapter 4.

Although the total cross section is different, the 4-momenta phase space of the outgoing particles in the EPA method is the same as in the full exact calculation (discussed in Ref. [103]). Therefore, we have implemented the EPA method described in Ref. [101] for calculating the differential cross sections (discussed below) and used the total cross sections from the exact calculation in generating trident MC events.

2.6.1 Trident Elastic Interactions

In the **coherent regime**, for a given neutrino energy E_ν, the total cross section in Equation (2.12) is expressed as,

$$\sigma_{\nu H}(E_\nu) = \frac{Z^2\alpha}{\pi} \int_{Q_{min}}^{Q_{max}} dQ\ \ \frac{\sigma_{\nu\gamma}\left(E_\nu, Q\right)}{Q} \int_{Q^2}^{Q^2_{max}} dQ'^2\ \ \frac{F_c^2\left(Q'^2\right)}{Q^2} \quad , \tag{2.13}$$

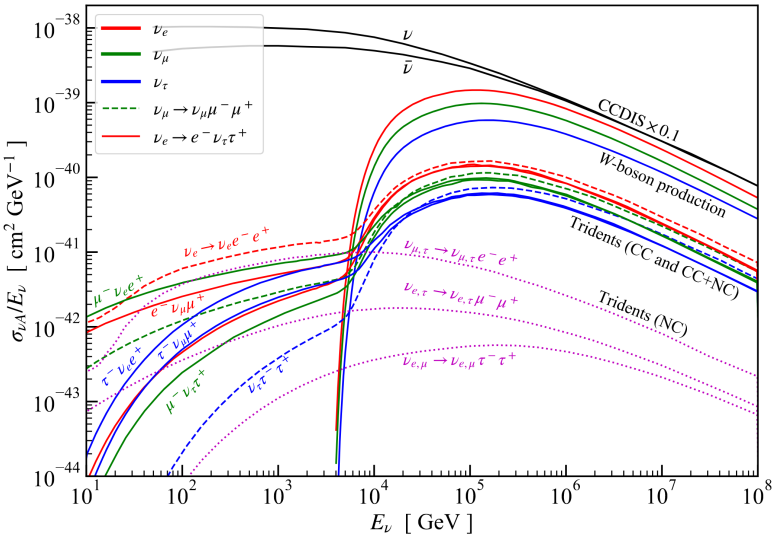


Figure 2.10: Total cross section of the trident interactions with O^{16} as the target nucleus, from Ref. [103]. The solid lines (red, green, blue) are the CC only channels, the dashed lines are the CC+NC channels, and the magenta dotted lines are the NC channels. The corresponding antineutrino cross sections are the same as the neutrino couples in the coulomb field of the nucleus. The plot also shows the cross sections of the standard ν CC DIS interaction and the W-boson production (i.e. the real W as the final outgoing particle from the interaction) for comparison.

where Z is the atomic number of the target nucleus, and α is the fine-structure constant. $\sigma_{\nu\gamma}$ is the total cross section of the neutrino-photon system, and F_c is the nuclear form-factor in the coherent interaction regime. The first integral runs on the photon momentum transfer variable, Q with a lower limit of $Q_{min} = \frac{(2m_\mu)^2}{2E_\nu}$, which is computed from the minimum CM energy required to produce dimuon events ($m_\mu = 105.7$ MeV is the muon mass). The upper limit cut-off for the coherent regime is defined as $Q_{max} = \frac{\Lambda_{QCD}}{A^{1/3}}$ [96], where Λ_{QCD} is the QCD scale parameter, and A is the total number of nucleons in the nucleus. Above this limit, the photon momentum transfer starts probing into the individual neutrons and protons, i.e. the interaction in the diffractive regime. For each value of Q, the second integral in the above equation runs on the variable Q'^2 with a range $[Q^2, Q^2_{max}]$. From Equation (2.13), we can derive the differential cross section of the photon momentum transfer as,

$$\frac{d\sigma}{dQ}(E_\nu, Q) = \frac{Z^2\alpha}{\pi} \cdot \frac{\sigma_{\nu\gamma}(E_\nu, Q)}{Q} \int_{Q^2}^{Q^2_{max}} dQ'^2 \quad \frac{F_c^2(Q'^2)}{Q^2} \tag{2.14}$$

with an imposed cut on the momentum transfer,

$$\frac{(2m_\mu)^2}{2E_\nu} < Q \leq \frac{\Lambda_{QCD}}{A^{1/3}} \quad . \tag{2.15}$$

The Wood-Saxon (WS) form-factor [104, 105] is used for $F_c(Q^2)$ in our calculation and is expressed as the Fourier transform of the nuclear charge distribution,

$$F_{WS}(Q^2) = \frac{1}{\int \rho(r)\ d^3r} \int \rho(r)\ \exp(-i\overrightarrow{q}\cdot\overrightarrow{r})d^3r \quad , \tag{2.16}$$

where the charge density is expressed as,

$$\rho(r) = \frac{\rho_0}{1+\exp\left(\frac{r-r_0}{a}\right)} \quad . \tag{2.17}$$

The values, $\rho_0 = 3/\left(4\pi A r_0^3\right)$, $r_0 = 1.126A^{1/3}$, and $a = 0.523$ fm come from the fits to experimental data as discussed in Ref. [105]. An alternative expression from Ref. [102] avoids the computationally expensive numerical integration in the above implementation and is found to be in good agreement with the full calculation. Therefore, the same approach is followed in this work, where the alternative form of F_{WS} is expressed as,

$$F_{WS}(Q^2) = \frac{3\pi a}{r_0^2+\pi^2a^2}\frac{\pi a\ \coth(\pi Qa)\ \sin(Qr_0) - r_0\ \cos(Qr_0)}{Qr_0\ \sinh(\pi Qa)} \quad . \tag{2.18}$$

In the **diffractive regime**, the differential cross section has the similar form with few changes,

$$\frac{d\sigma}{dQ}(E_\nu, Q) = \frac{Z\alpha}{\pi}\cdot\frac{\sigma_{\nu\gamma}\left(E_\nu, Q\right)}{Q}\int_{Q^2}^{Q^2_{max}} dQ'^2\ \ \frac{F_d^2\left(Q'^2\right)}{Q^2}\ \ ;\ \ \frac{\Lambda_{QCD}}{A^{1/3}} < Q \leq 1.3\ \text{GeV}\ \ , \tag{2.19}$$

where the lower limit Q_{min} is the cut-off for the coherent regime upper bound and the upper limit Q_{max} is set at 1.3 GeV, the energy scale where the PDFs for DIS interaction start probing into the quark structure of the target nucleon. The form-factor in this regime is computed from the hadronic tensor of the nucleons (proton and neutron) in the squared matrix element, and the final expression is derived in Ref. [95] as,

$$F_d(Q^2) = \frac{G_{dip}(Q^2) + \tau\ \xi\ G_{dip}(Q^2)}{1+\tau} \quad , \tag{2.20}$$

where $\tau = Q^2/4M^2$ with the average of proton and neutron mass, $M = (m_p + m_n)/2$. ξ is the difference between the magnetic moment of proton (μ_p) and neutron (μ_n)

and is expressed in units of the nuclear magneton (μ_N) as, $\xi = (\mu_p - \mu_n)/\mu_N \approx 4.7$. $G_{dip}(Q^2)$ is the electric dipole form factor of proton and is expressed as,

$$G_{dip}(Q^2) = \left(1 + \frac{Q^2}{0.71 \;\; GeV^2}\right)^{-2} \quad . \tag{2.21}$$

2.6.2 Trident Inelastic Interactions

As explained in Ref. [103], the inelastic trident interaction has two contributions, the photon-initiated and quark-initiated subprocess. In the photon-initiated subprocess, the hadronic coupling occurs through the virtual photon content of the nucleons. It requires PDFsets that consider only the inelastic component of the photon content, as the elastic contribution is already accounted for in our calculation of the coherent and diffractive regimes. In the quark-initiated subprocess, the quark content of the nucleon is directly involved in the scattering as the initial-state particle. The use of PDFsets to sample exchanged virtual photons for simulating inelastic interactions is discussed in more detail in Section 4.1.2.

2.7 Probing New Physics with Neutrino Interactions

In recent years, the interest in studying the trident processes has been revived in the context of new physics searches [39, 100, 106]. In particular, a proposed extension to the Standard Model can be studied using the trident dimuon processes [107, 108]. This model introduces a new $U(1)$ gauge group associated with the difference between the muon and tau lepton numbers ($L_\mu - L_\tau$) and proposes the existence of a new gauge boson, Z'. The proposition gained popularity as it explains the discrepancy in several anomalous measurements (such as muon $g-2$ anomaly [109] and B meson decay [110]) and provides a mechanism for neutrino mass modelling and non-standard interactions. The model allows the neutrinos to couple to Z' similarly to the NC channels shown in Figure 2.8 and produce identical outgoing particles. Therefore, any excess in the observation of trident dimuon events compared to the Standard Model prediction can indicate the existence of such a new particle. The accelerator neutrino experiments CHARM-II [36], CCFR [37], and NuTeV [38] have detected trident dimuon events

with the energy of the incoming neutrinos in $\mathcal{O}(10 - 100$ GeV). The results from these observations are then used in Ref. [100] to probe the mass-coupling parameter space of Z'. The detection of trident dimuons in IceCube is sensitive at an energy scale $E_\nu > 1$ TeV, which is beyond the reach of any current accelerator neutrino experiments, and such detection is possible due to the enhancement of the cross section from the W-boson production. Therefore, IceCube can also contribute to this new physics search by observing dimuon events.

In addition to the Standard Model extension discussed above, possible observation of 'dimuon-like' events in IceCube from many other BSM scenarios has been proposed [29, 41, 111–113]. In most cases, supersymmetric (SUSY) models predict the existence of long-lived charged particles such as *staus*, the super-partner of the tau lepton. These particles can be produced in pairs and travel a significant distance due to their long lifetime, mimicking a dimuon-like event signature in the detector. IceCube performed past searches for these BSM particles by looking for two widely separated ($\mathcal{O}(100$ m)) parallel tracks and observed no events [114, 115]. The search for dimuons in this work allows for a much smaller separation ($\mathcal{O}(10$ m)) of two diverging tracks originating from the same interaction vertex. The analysis uses a new and improved strategy based on advanced machine learning algorithms to address the challenging task of identifying such events and, thus, is complementary to the previous searches.

In summary, dimuon events in IceCube serve as an important testbed for studying QCD physics and new physics. For many BSM physics scenarios, dimuon events from the Standard Model processes are considered as background and require accurate characterization in order to extract signal information from data. In this work, we have performed an inclusive search for dimuon events in IceCube and compared it to the Standard Model expectations, offering important implications for several aspects of neutrino physics.

Chapter 3

Detection of Neutrinos in IceCube

The IceCube neutrino observatory is designed to detect some of the highest energy neutrinos (TeV-PeV) passing through the earth. It began operation in 2011 and has been taking data with high detector uptime ($\gtrsim$ 99%). The experiment is located at the geographic south pole and uses a cubic kilometre of instrumented Antarctic glacial ice as the detector volume. This chapter will introduce the working principle of neutrino detection in such a transparent medium. In addition, we will discuss the details of the detector configuration, the origins of the detected neutrinos, and their physics implications.

3.1 Neutrino Interactions in Ice

3.1.1 Cherenkov Radiation

The existing techniques and tools probe the neutrino weak interactions by detecting the appearance of outgoing particles from the interaction vertex. The detector observes the outgoing charged particles via their electromagnetic interactions in the medium. Information on the weak interaction and the interacting neutrino is then extracted from the properties (e.g. energy, direction) of the electromagnetically visible outgoing particles. When relativistic charged particles travel in a medium (of refractive index n) with a speed (v_x) larger than the phase velocity of light in that medium ($v_p = c/n$), they emit electromagnetic radiation called *Cherenkov radiation*. As a result, a conical wavefront of photons (also referred to as Cherenkov photons) is created with the vertex located at the moving charged particle, as shown in Figure 3.1. The Cherenkov photons are emitted at an angle (θ_c) with respect to the

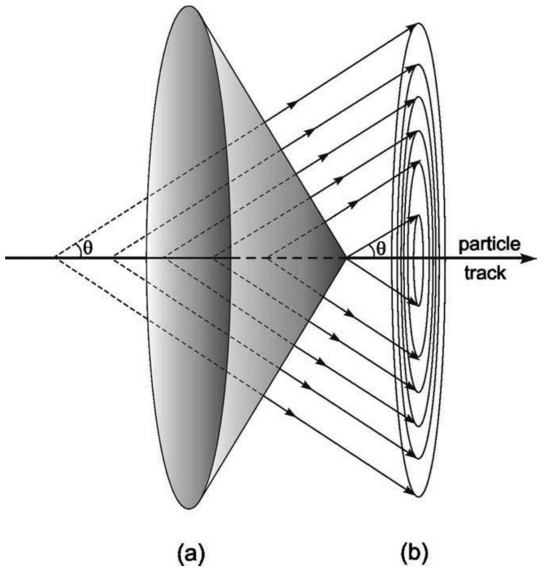


Figure 3.1: Cherenkov conical wavefront is shown in (a) and the photon propagation direction is indicated by the arrows (b). The angle θ in the diagram refers to the Cherenkov angle θ_c. The diagram is taken from Ref. [116].

particle travel path and depend on the particle's speed and the refractive index of the medium,

$$\theta_c = \cos^{-1}(1/n\beta) \quad , \tag{3.1}$$

where $\beta = v_x/c$. For a charged particle travelling with $\beta \sim 1$ in ice ($n = 1.31$), the Cherenkov angle is $\theta_c \approx 41^0$. The produced Cherenkov photons in a transparent medium, such as ice in IceCube, can travel a long distance. Many detectors, including IceCube, utilize this mechanism to collect the Cherenkov photons efficiently using light sensors and are known as Cherenkov detectors. These sensors typically operate in the optical photon wavelength range of 300-600 nm. The Cherenkov photon emission spectra of a particle with charge Ze can be estimated per unit particle path length and per photon wavelength as [117],

$$\frac{\mathrm{d}^2N}{\mathrm{d}x\ \mathrm{d}\lambda} = \frac{2\pi\alpha Z^2}{\lambda^2}\left(1 - \frac{1}{\beta^2 n^2(\lambda)}\right) \quad , \tag{3.2}$$

where λ is the photon wavelength, α is the fine-structure constant and the refractive index, $n(\lambda)$, is a function of the wavelength.

3.1.2 Hadronic and Electromagnetic Showers

Charged particles lose a tiny fraction of their energy via Cherenkov radiation and, instead, deposit most of their energy through other interactions with the detector material. In neutrino CC and NC DIS, the outgoing hadrons undergo further nuclear interactions with the surrounding matter to produce additional secondary particles. Starting from the primary outgoing hadrons, multiple iterations of such process form a cascade of high energy particles, collectively known as a *hadronic shower*. The majority of the particles produced in a hadronic shower are light hadrons such as pions and kaons. The mean path of the hadrons between two consecutive interactions is defined as the *nuclear interaction length* (λ_I) and dictates the size of the longitudinal shower development. The nuclear interaction length for light hadrons in ice is $\lambda_I^{ice} =$ 90.8 cm [118].

In addition to the hadronic shower in the CC DIS processes, the outgoing charged lepton also interacts electromagnetically in the detector. The charged leptons lose energy via ionization and excitation processes at low energy, and the loss rate has a logarithmic dependence on the lepton energy. However, at high energy, the energy loss is dominated by Bremsstrahlung and pair production processes for electrons and by an additional photonuclear process for muons and taus. These processes are referred to as radiative loss, and the energy loss rate for all participating leptons ($e^{\pm}, \mu^{\pm}, \tau^{\pm}$) varies linearly with the lepton energy. The critical energy (E_c) of the charged leptons is defined as the energy at which loss from the ionization process becomes equal to the radiative loss. In IceCube, the energy of the outgoing electron from an electron neutrino interaction is much higher than its critical energy in ice ($E_c = 78.6$ MeV). Therefore, the electron triggers an iterative production of lower energy photons and electrons through Bremsstrahlung and pair production processes. This cascade of electrons ($e^{\pm}$) and photons (γ) is called an *electromagnetic (EM) shower*. Similar to the nuclear interaction length in the hadronic showers, a characteristic length called *radiation length* (X_0), describes the longitudinal size of the EM showers. The radiation length in ice, $X_0^{ice} = 39.31$ cm $< \lambda_I^{ice}$, indicates that the EM showers dissipate energy in a shorter distance than hadronic showers.

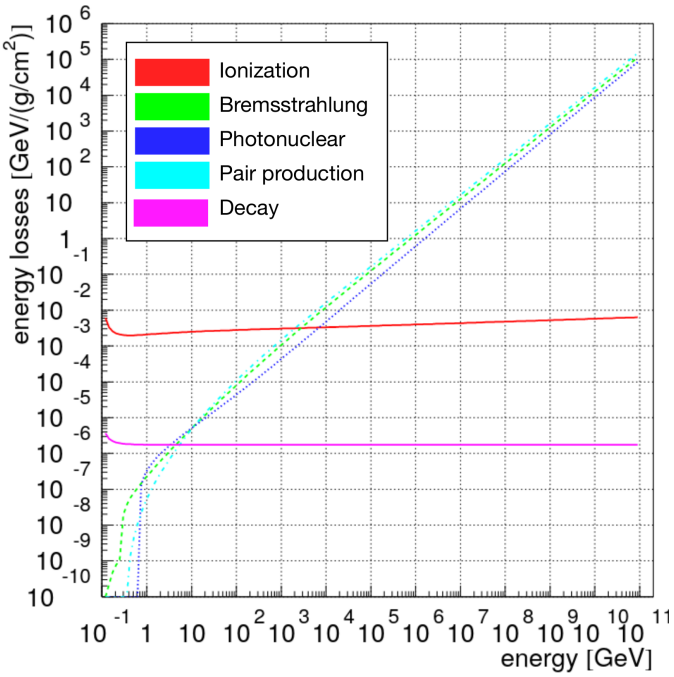


Figure 3.2: Muon energy loss via different EM interaction processes in ice, calculated in Ref. [120]. The critical energy E_c is defined as the intersecting point of the ionization and Bremsstrahlung energy loss.

3.1.3 Muon Energy Loss

The high energy muons produced in ν_μ CC DIS process have the same interactions with ice as the electrons. The average stopping power of muons is expressed as [119],

$$-\left\langle \frac{\mathrm{d}E_\mu}{\mathrm{d}x} \right\rangle = a(E_\mu) + b(E_\mu) \cdot E_\mu \quad , \tag{3.3}$$

where $a(E_\mu)$ is the electronic stopping power, and $b(E_\mu)$ is the radiative stopping power. The values of a and b in ice are calculated in Ref. [120] to be approximately constant for the muons in the energy range $20-10^{11}$ GeV (the energy range used in the fit for the calculation) and are 0.246 GeV.m^{-1} and 4.31×10^{-3} m^{-1}, respectively. The critical energy for muons in ice is $E_c = 1.031$ TeV, below which they have small energy losses due to the ionization process. Radiative processes dominate energy loss of multi-TeV muons above the critical energy. These processes are stochastic in nature and are responsible for significant muon energy loss, as shown in Figure 3.2. However, muons ($m_\mu = 105.66$ MeV/c^2) are more massive than electrons ($m_e = 0.51$ MeV/c^2), so they lose a much smaller fraction of their energy in the interactions. As a result, muons are much more penetrating than electrons and can travel a long distance through the ice

before stopping. Figure 3.2 shows the individual muon stopping power for ionization, Bremsstrahlung, pair production, photonuclear, and decay processes.

Tau neutrino CC interaction produces an outgoing tau with a very short lifetime. 64.8% of the time, the tau decays into hadrons, producing a hadronic shower. The tau decays into an electron and neutrinos with a branching ratio of 17.8% and generates an EM shower. The other 17.4% of the time, the tau decays into a muon and neutrinos. Due to the short travel distance of the tau and IceCube's inability to differentiate between EM and hadronic shower, the hadronic shower from the primary interaction vertex and the EM/hadronic shower from the tau are often inseparable. Therefore, the event signature of most ν_τ CC DIS processes is similar to the ν_e or ν_μ CC DIS interactions in IceCube, depending on the tau decay channel.

All the charged particles produced in the hadronic shower, EM shower, and the muon's interaction in ice emit Cherenkov light if their energy is above the threshold for the Cherenkov radiation, i.e. $n\beta > 1$ in Equation (3.1). Cumulative light yield from these particles is detected by Cherenkov detectors like IceCube and forms the primary observable to study neutrino interactions. Precise modelling of the showers, muon propagation, and optical properties of the ice is crucial to establish the correlation between the detected Cherenkov light and event properties like energy, direction, and event topology. In Chapter 4, we will discuss the reconstruction of these events in more detail.

3.2 Detector Description

The IceCube experiment was designed based on the knowledge and experience acquired from the success of its predecessor experiment, AMANDA [121]. The design is optimized for detecting high energy ($\gtrsim$ 1 TeV) neutrinos from astrophysical objects outside the solar system, along with many additional scientific goals involving atmospheric neutrinos. In order to detect Cherenkov light from neutrino interactions, the IceCube detector uses a total of 5160 light sensor units, known as the digital optical modules (DOMs). These modules are permanently buried deep in the Antarctic glacial ice at a depth between 1450 m and 2450 m below the surface. Placement of the DOMs in ice was achieved by boring holes using hot water drills [122] and deploying

a vertical string system, each with 60 DOMs attached, before the melted ice in the holes refroze. The attached string ensures the regular spacing between the DOMs during deployment and provides power supply and communication to the modules. The primary in-ice configuration has 78 such strings (called IceCube strings) installed in a triangular grid with a horizontal spacing of approximately 125 m. The cross-sectional area of the strings forms a hexagonal shape, as shown in Figure 3.3. The vertical spacing of the DOMs in these strings is 17 m. The total instrumented volume from the array configuration of the DOMs is approximately 1 km^3.

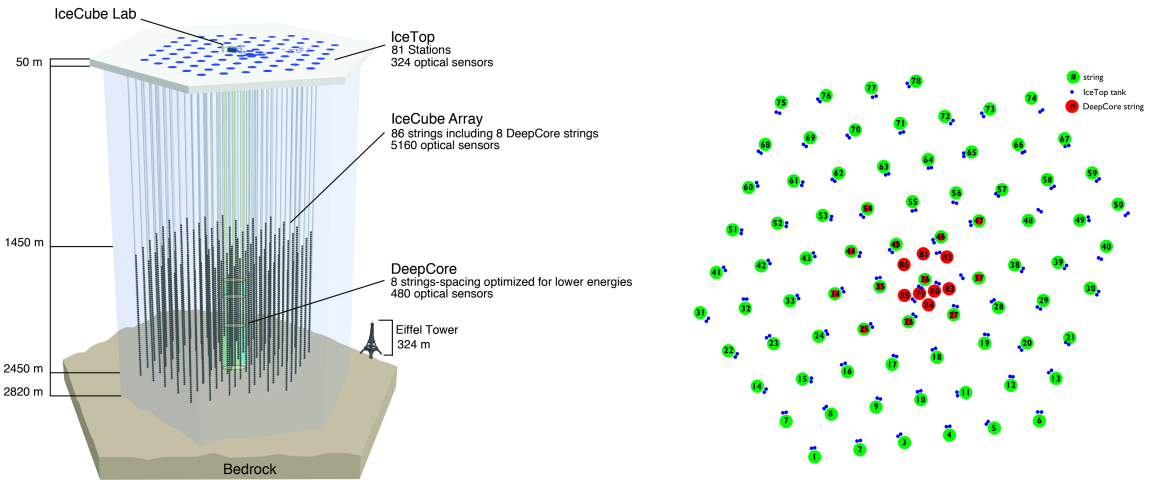


Figure 3.3: The IceCube detector (left) showing the IceTop, IceCube and DeepCore components. For reference, the Eiffel tower is shown on the side, to the scale of the detector. (Right) shows the top view of the detector, where the green points are the IceCube strings, the red points are the DeepCore strings, and the blue points are the IceTop tanks.

An additional eight strings with a smaller vertical DOM spacing are inserted in the deepest and clearest ice of the detector volume. The DOMs from these eight strings (known as DeepCore strings) and the seven IceCube strings around the centre of the detector form a subset of the total instrumented volume known as DeepCore. The bottom 50 DOMs in DeepCore strings have a vertical spacing of 7 m at a depth between 2100 m and 2450 m. The remaining 10 DOMs in each string are located at a depth above 2000 m with 10 m vertical spacing and serve as the veto cap to reject the atmospheric muon background for detecting neutrino interactions inside DeepCore. The horizontal spacing of the strings in DeepCore ranges from 41 m

to 105 m, with an average spacing of 72 m. The denser configuration of sensors in DeepCore allows for a higher Cherenkov light yield and, thus, lowers the energy threshold of the detected neutrinos in DeepCore ($\sim$ 10 GeV) compared to the total detector volume ($\sim$ 100 GeV) [123].

In addition, a surface array of DOMs is installed on the ground near the approximate in-ice string positions (also shown in Figure 3.3). There are 81 stations in the surface-array system (known as IceTop), with each station having two ice-filled tanks and each tank containing a pair of DOMs. The IceTop is constructed primarily to detect the Cherenkov radiation from the cosmic ray air showers and to be used as the surface veto for in-ice neutrino event detection. The role of the IceTop array is not directly relevant to this work. Thus, any further discussion of the detector refers only to the in-ice array, i.e. IceCube and DeepCore.

3.2.1 The Digital Optical Module

Photomultiplier tubes (PMTs) are one of the basic photon detection technologies used in the Cherenkov detectors and are capable of detecting Cherenkov light with single photon precision and high timing resolution ($\mathcal{O}(ns)$). They consist of a photocathode, a series of dynodes, and an anode. The elements are contained inside a vacuum tube, and a high voltage is applied to the electrodes. When a photon hits the photocathode, it can emit an electron (known as a photoelectron (PE)). In reality, every incident photon does not produce the corresponding photoelectron, and the efficiency (known as *quantum efficiency*) with which photons are converted into photoelectrons is a characteristic of the specific PMTs and incident photon wavelengths. Following its production, the photoelectron accelerates due to the high electric potential and collides on the first dynode to release more electrons. The successive collisions along the dynode series produce an avalanche of electrons. The anode then collects the amplified electrical signal and indicates the detection of an incident light pulse. The general working principle of a PMT is shown in Figure 3.4. The primary component of the DOM is a downward-facing PMT with a 25.4 cm (10") diameter [125]. In addition, the module contains circuit boards, which are responsible for the PMT power supply, control, calibration, and data acquisition. All the components are housed in a spherical glass vessel that can withstand high pressure, as shown in Figure 3.5. The

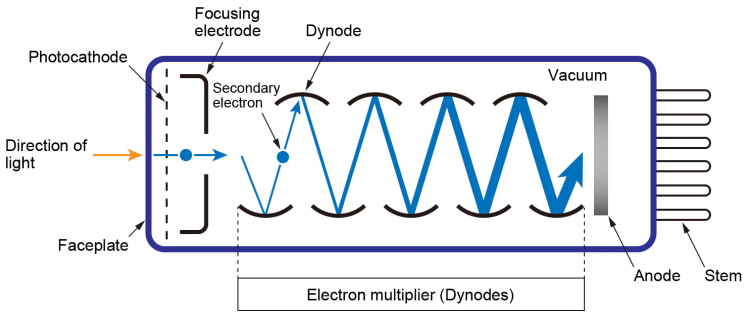


Figure 3.4: A diagram showing the working principle of a PMT, from Ref. [124]. The blue lines in the figure refer to the creation of the photoelectron from the incident photon and the subsequent production of electrons in the dynode series. The stem in the diagram refers to the electrical connector that transfers the electrical signal from the anode to the later part of the PMT electronics for further processing.

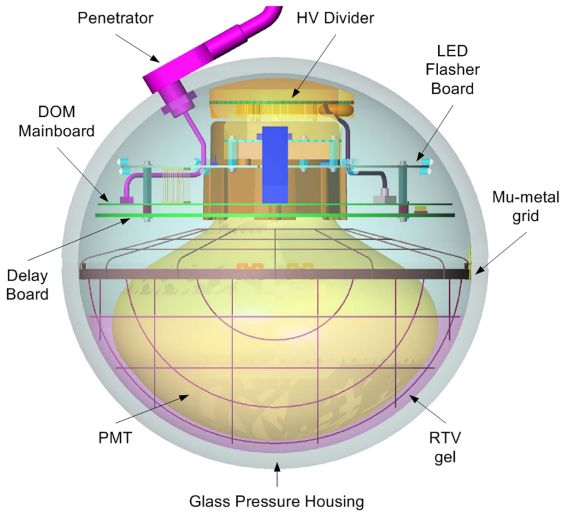


Figure 3.5: A diagram of the DOM showing all the components inside the glass housing, from Ref. [126].

PMT has the sensitivity to detect photons in the wavelength range of 300 nm-650 nm. The IceCube strings use the R7081-02 PMT from Hamamatsu photonics with a peak quantum efficiency of 25% at 390 nm [127]. Most of the DOMs in the DeepCore strings have R7081-02MOD PMTs, which operate with a higher peak quantum efficiency of 34%.

When the DOM detects a photon, the PMT signal of the single photoelectron (SPE) is digitized as a waveform by the on-board electronics and is transmitted with a time-stamp. The waveform contains the charge (area under the electric signal) and time information of the detected pulse. An average SPE waveform in IceCube is

shown in Figure 3.6. Since the spatial and temporal distribution of the light and the overall light yield are used for event reconstruction and particle identification, precise characterization of the SPE waveform is an important step. The characterization and study of various PMT properties were performed in the laboratory before deployment and are described in detail in Ref. [125]. The PMTs operate at a gain of 1×10^7, and the timing resolution of the SPE pulse is measured to be 2.7 ns.

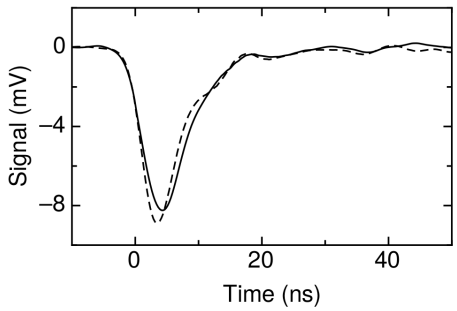


Figure 3.6: An average SPE signal measured from the detection of 10,000 individual photons [125]. The dashed and solid lines show the measurements for the old and new configuration of the PMT electronics, respectively (discussed in Ref. [125]).

In addition to the signal pulse, the PMTs are susceptible to various noise pulses that can affect many physics analyses in IceCube. Without any incident photon on the photocathode, PMTs can still generate a current called dark noise. In IceCube, the dark noise is primarily caused by the radioactivity and scintillation of the glass material and the thermionic emission of electrons from the photocathode. The dark noise rate of the PMTs used in IceCube operating at -40^0 C is measured to be $\sim$ 300 Hz [125]. There are also noises which are correlated with the signal pulses. Ionization of a residual gas atom (traces of gas elements in the vacuum tube) and luminous reactions in the electrodes, i.e. light emission due to the electron bombardment, can cause a second pulse in the PMT. The late occurrence of the luminous reaction and longer drift time of the heavier ionized gas atom result in a late arrival of the second pulse after the first signal pulse from the photoelectron [128]. The time difference between the two pulses is significantly high (in the range of 300 ns - 11 μs), thereby mimicking an additional, later photoelectron detection in the PMT. This noise is known as an afterpulse. The afterpulse measurement in the IceCube PMTs reported the prominent noise peaks appearing at 600 ns, 2 μs, and 8 μs af-

ter the main response peak. On average, the integral of pulses from 300 ns to 11 μs has an afterpulse contribution of 0.06 SPE per primary photoelectron (measured with primary pulses going up to 10^6 PE). The early arrival of a pulse, known as a prepulse, occurs when a photon skips the photocathode and hits one of the dynodes directly. Prepulses usually have less charge than normal pulses and take place with small probability. For IceCube PMTs, prepulse is only observed during the detection of a large amount of light ($\gtrsim$ 5000 photons) within a short time period (30 ns).

3.2.2 Data Acquisition and Processing

The IceCube data acquisition (DAQ) system can be divided into two main parts. The first part is a decentralized DAQ subsystem where each DOM autonomously digitizes the PMT signal and outputs the detected waveform with a timestamp. In the second part, the DOM outputs are collected in a central location, the IceCube laboratory (ICL), near the centre of the detector footprint on the surface. The DOM initiates the waveform capture and digitization process when the PMT signal exceeds the discriminator threshold of 0.25 PE [126]. The Analogue Transient Waveform Digitizer (ATWD) starts sampling the PMT signal into three channels with different gains ($\times$16, $\times$2, and $\times$0.25). A total of 128 samples are stored as a digitized waveform with a sampling rate of 300 MSPS. To reduce dead time in the digitization process, each DOM contains two ATWDs so that one unit is available while the other ATWD is engaged. In addition, a fast analog-to-digital converter (FADC) continuously samples the PMT signal in parallel with a lower sampling rate of 40 MSPS. The time window to record the FADC output is chosen to be 6.4 μs which ensures the digitization of the waveform for a longer physics signal. The DAQ system in the ICL controls the selection and flow of the data from the individual DOMs. The digitized waveform called a hit is stored in the on-board buffer in each DOM until a trigger decision is made. When a DOM detects a PMT signal, the trigger system opens a $\sim$ 1 μs time window to look for a local coincidence (LC) by sending and receiving trigger messages to and from the nearest DOMs in both directions along the string. The DOMs that satisfy the LC trigger condition are identified with LC tags. IceCube has two basic operating modes based on the LC trigger conditions on each DOM. In a soft local coincidence (SLC) mode, the DOM hits with LC tags keep both the ATWD

and FADC data for transmission to the surface pending fulfillment of higher-level trigger conditions. The isolated DOM hits with no LC tag keep only the FADC data in SLC mode. In hard local coincidence (HLC) mode, the hits with no LC tags are entirely discarded. These triggering conditions reduce the data transmission rate in each DOM from $\sim$ 700 Hz of raw PMT SPE rate to $\sim$ 10 Hz of LC tagging rate by primarily discarding the noise hits and sacrificing a negligible fraction of physics hits. Figure 3.7 summarizes the information flow of the DOM-level DAQ subsystem.

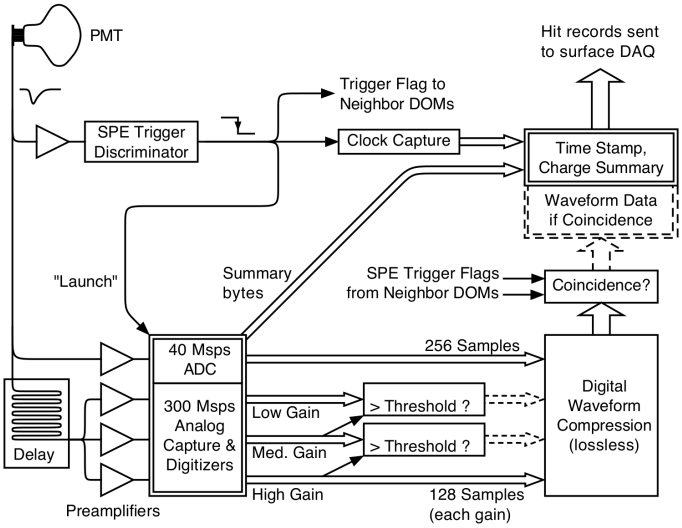


Figure 3.7: Flow diagram of the DAQ subsystem within each individual DOM [126].

After sending the HLC and SLC hits from all the DOMs to the surface, additional software triggering algorithms are applied for further data reduction and physics event selection. The central DAQ system at the ICL performs a time sequencing of the data and merges all the hits within a specific time window (set by the trigger condition) around the trigger into an 'event'. The simplest triggering algorithm to build an event is the Simple Multiplicity Trigger-8 (SMT8). This trigger looks for eight or more HLC hits within a sliding time window of 5 μs [123]. The SMT8 trigger window for an event starts when an initial trigger condition ($>$ 8 HLC hits in 5 μs) is met and continues to extend until there is no HLC hit in the last 5 μs sliding window. Two additional readout windows are added at the start (4 μs window) and the end (6 μs window) of the trigger to construct the overall event time window, as shown in Figure 3.8. The approximate event rate from the SMT8 trigger is 2.7 kHz and follows

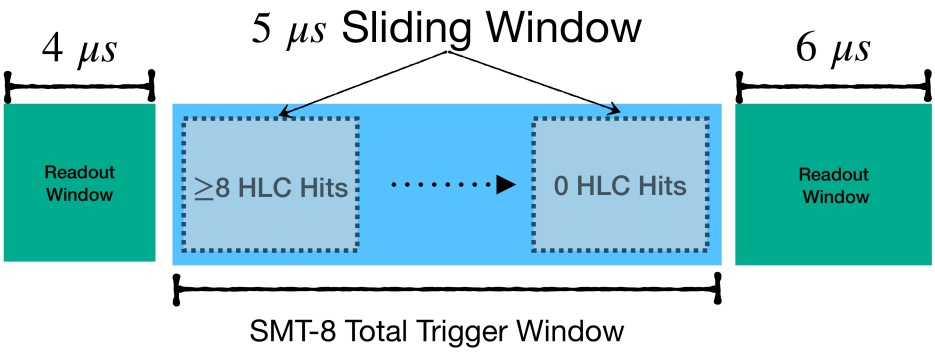


Figure 3.8: Total event time window for SMT8 triggers of in-ice DOMs. All the hits within the readout window and the SMT8 trigger window are merged to construct an event.

a seasonal variation due to the change in the production rate of atmospheric muons and neutrinos. The event selection for this work starts with the SMT8 trigger, and further filtering processes are discussed in Chapter 4 in detail. The raw detector data rate from the DAQ at this stage is $\sim$ 1 TB/day, which is approximately ten times higher than the satellite bandwidth allocation for IceCube from the south pole (100 GB/day). The online processing and filtering system handles this issue by performing an aggressive data compression procedure described in detail in Ref. [123]. A massive compression (9%) of the raw data for an average event is achieved by applying a deconvolution algorithm to the digitized waveform and storing it in a compact data format. This process converts the waveform into an estimated number of PE (with 30% charge resolution) per coarser time bins with a typical accuracy of 1 ns for ATWD data and 8 ns for FADC data [129], generating a series of discrete pulses for each DOM. All the pulses from an event are stored in a *pulse series* object which maps the pulses to their respective DOMs and forms the final fundamental basis for most high-level event reconstruction and identification in IceCube.

3.2.3 Detector Calibration

Various calibration and monitoring systems are in place to maintain a good quality of the acquired data for further analysis. For time calibration, all the DOMs are synchronized with a single GPS-run master clock at the ICL [126]. The charge associated with the detected PMT pulse is extracted from the width and amplitude of the

waveforms. The calibration of the SPE charge distribution for each DOM was initially determined in the laboratory [125] and later improved with an updated in-situ measurement [130]. The DOM characteristics measured in the laboratory can change after deployment in ice due to different conditions like change of environment, aging, and transport. The calibration of the overall light detection efficiency of the detector is, thus, a crucial step for every physics analysis. This is achieved by measuring an average scale factor attributed to all the DOMs, known as *DOM efficiency*. It uses atmospheric muon events in the detector for the measurement and represents the relative change to the light detection efficiency measured in the laboratory [131].

Each DOM also contains 12 light-emitting diodes (LEDs) on-board that support additional calibration procedures. Controlled emission and measurement of the light from the LEDs in neighbouring DOMs are used to verify time calibration, determine the DOM positions, and study the ice's optical properties.

3.2.4 South Pole Ice Properties

The light propagation in the detector undergoes scattering and absorption in the ice. Therefore, the ice properties must be well understood to extract meaningful event information from the pulses. The use of the LEDs as the controlled light source, followed by the detection of the light in the DOMs form the primary mechanism to study the scattering and absorption properties of the ice [132]. The measurements from the study are then used to develop South Pole Ice (SPICE) models that give a complete description of the optical properties. Most of the ice in the detector volume (known as 'bulk ice') was naturally formed in layers over 165,000 years and was susceptible to a broad range of weather conditions over a long period [133]. This results in a variation of the impurity content at different depths of the ice sheet and leads to the depth-dependent modelling of the scattering and absorption. The average distance between two consecutive scattering of light is denoted as $1/b$, where b is the scattering coefficient. In IceCube, a more useful form called effective scattering coefficient (b_e) is expressed as $b_e = b(1 - \langle \cos\theta \rangle)$, where θ is the deflection angle at each scattering of light. The average distance travelled by the light before it is absorbed is expressed in terms of the absorption coefficient (a) as $1/a$. The SPICE model uses a parametric form of the scattering and absorption coefficients (b_e, a) as

a function of photon wavelength and depth. The initial version of the ice model developed in Ref. [132] is known as SPICE MIE. Since then, improved calibration and modelling methods have contributed to developing a better description of the ice. This work uses one of such improved ice models, SPICE3.2, for event simulation and reconstruction. Figure 3.9 shows the effective scattering (b_e) and absorption (a) coefficients as a function of depth for different wavelengths. The plots in the figure show a significant increase in the scattering and absorption of light at a depth between 2000 m to 2200 m. This is caused by a higher amount of dust in the ice, so this ice layer is referred to as the 'dust layer'.

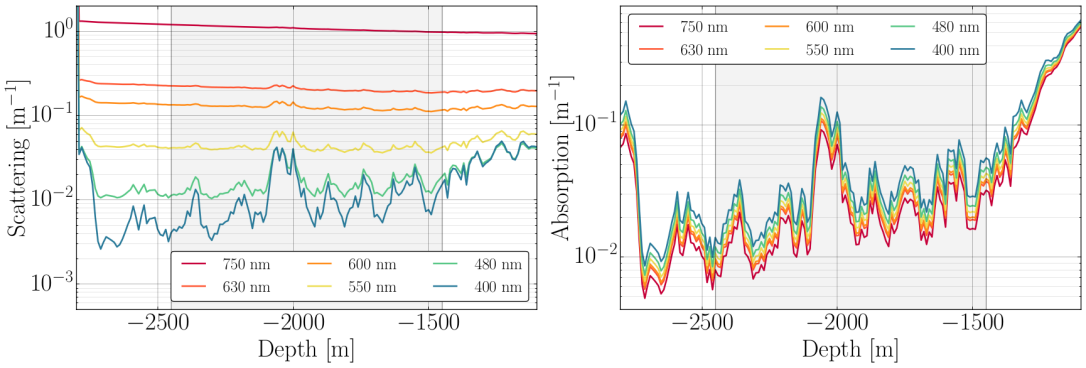


Figure 3.9: Effective scattering (left) and absorption (right) coefficients of the south pole ice (at string 63) at different photon wavelengths, from Ref. [134].

During the deployment of the DOMs, drilling $\sim$ 60 cm diameter holes into the ice created columns of melted ice for a short period. This process introduced air and other additional impurities into the water. Radially inward refreezing of the water resulted in the accumulation of air bubbles around the centre with a 5-10 cm diameter and caused a change in the ice properties of the hole columns compared to the bulk ice. These ice columns are referred to as 'hole ice' and show a higher scattering effect than bulk ice. The scattering of light in the hole ice affects the directionality of the incident photons and requires separate modelling to address the angular sensitivity of the DOMs. Further discussion on the hole ice modelling in this work's context is discussed in Section 6.2.1.

The antarctic glacial ice is not a stationary geographical structure but moves at a speed of $\sim$10 m/year [135]. In IceCube, the evidence of the ice flow is observed in two distinct effects. The measurements involving LEDs in Ref. [132] indicated

that the depth-dependent ice layers are not horizontal across the detector but have a 'tilt' correlated to the flow direction. In addition, the shearing of the ice layers due to the flow causes the crystal structure to form an azimuthal dependence for light scattering. This effect is referred to as ice 'anisotropy' [136]. Both of these effects are incorporated into the ice model used for this work.

3.3 Event Topologies

There are two primary event topologies observed in the detector from the neutrino interactions discussed in Section 3.1. The neutrino NC DIS interactions of all flavours produce a hadronic shower, typically confined within $\mathcal{O}(10$ m) and have multiple low energy shower particles produced with wide opening angle spread with respect to the incoming neutrino direction. Therefore, the shower produces a net Cherenkov radiation that propagates almost equally in all directions from the neutrino interaction vertex, creating a nearly spherical shape of light yield. The ν_e CC DIS and the majority of the (low energy) ν_τ CC DIS (where the outgoing τ decays into either hadrons or electrons) interaction produce two showers (EM+hadronic for ν_e and EM/hadronic+hadronic for ν_τ) very close to the interaction vertex. Due to the limited detector resolution and no detectable difference in the light emission pattern between EM and hadronic shower in IceCube, the cumulative light from the two showers generates an event signature similar to the NC interactions. This type of event topology is referred to as 'cascade'. An example cascade event detected in IceCube is shown in Figure 3.10. As the light emission from the cascade events is usually contained within the detector volume, the showers' deposited energy is reconstructed with a resolution of 30% for low energy (100 GeV) events and has improved reconstruction at higher energies, e.g. 8% resolution for 100 TeV events [129]. However, the directional reconstruction of the cascade events performs poorly with a typical resolution of 15^0 for $\gtrsim 100$ TeV cascades, as the directions of most of the detected photons are loosely coupled to the incoming neutrino direction.

The other common type of event topology comes from the detection of high energy muons and is referred to as 'track' events. The outgoing muon from ν_μ CC DIS and the tau decay channel, $\tau \to \mu\nu_\mu\nu_\tau$ in ν_τ CC DIS interactions carries a large portion

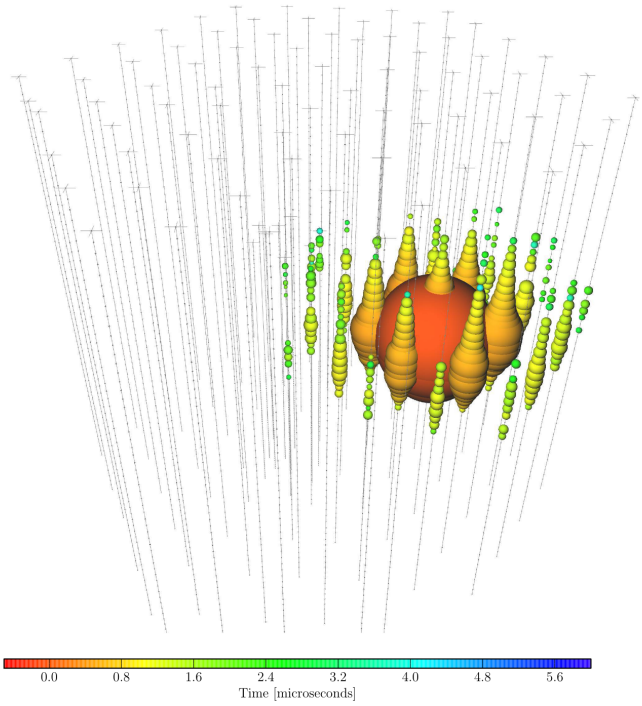


Figure 3.10: Event topology of a cascade event detected in IceCube. This event corresponds to an energy deposition of 1.14 PeV [137]. The size of the blobs refers to the amount of charge seen by the corresponding DOMs. The colour gradient refers to the hit times for the corresponding DOMs with the timescale showing at the bottom.

of the incoming neutrino energy and traverses a long distance ($\mathcal{O}$(km)) through the ice. The muon's interactions in ice result in the emission of Cherenkov light along the muon path, as shown in Figure 3.11. As the outgoing muon from the primary interaction has a small opening angle, the track events carry the initial neutrino directional information across the DOMs over a significant distance. It results in a high-quality direction reconstruction of the tracks ($\lesssim 1^0$ angular resolution for most events). However, a long propagation length ($\gtrsim$ 1 km for 200 GeV muons in ice) also means that the muon can escape the detector volume, limiting the energy reconstruction to only part of the track segment depositing light within the detector volume. If the interaction vertex for track events is inside the detector, the hadronic shower component of the interaction is visible, and the events are identified as starting tracks. Most high energy track events ($\geq$ 1 TeV) have the interaction vertex outside the detector, and the muon also escapes the detector volume, creating a through-going track signature.

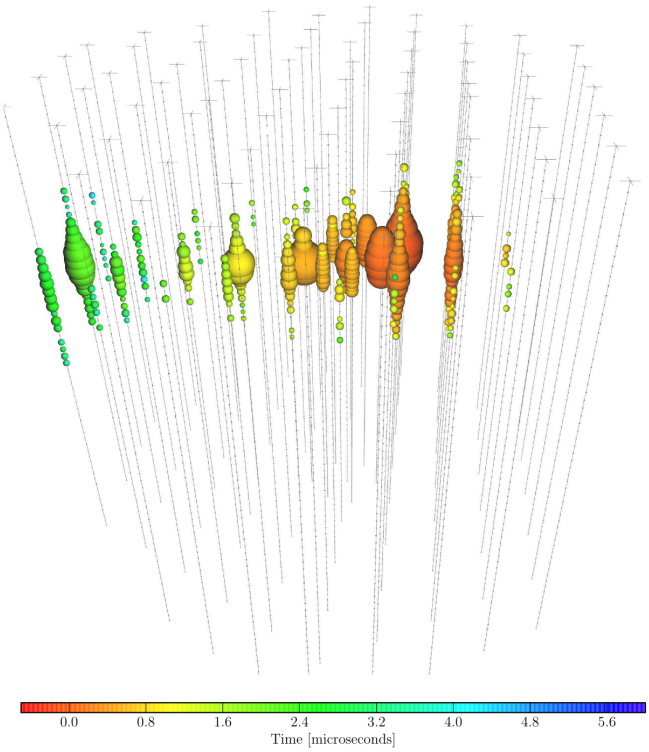


Figure 3.11: An example starting track event detected in IceCube with reconstructed energy of 71.4 TeV [137]. The size of the blobs refers to the amount of charge seen by the corresponding DOMs. The colour gradient refers to the hit times for the corresponding DOMs with the timescale showing at the bottom.

There are other subtle event topologies, which require rigorous analysis of the DOM hit patterns for their identification. If the energy of the outgoing τ from ν_τ CC interaction is high enough, it can travel a significant distance before producing a shower from its decay. A spatially separated pair of showers from this interaction creates a 'double-cascade' event topology in the detector. Several analyses to identify ν_τ double-cascade events have been conducted in IceCube [138, 139]. The recent result from the search has reported the detection of two ν_τ events with 2.8σ significance.

Another event topology, called 'double-track', occurs when two high energy muons propagate simultaneously with a short lateral separation (10-100 m) from one another. Prior to this work, no double-track event has been observed in IceCube. The neutrino interactions that produce such dimuon events are discussed in the previous chapter. The resolution of two distinct tracks in the event is limited due to the sparse geometry of the detector, making it a challenging task. The primary objective of this work is

to create an analysis framework for identifying the double-track events. Several novel methods are developed for the double-track event classification and are discussed in Chapter 5.

3.4 Sources of Neutrinos

3.4.1 Atmospheric Neutrinos

Most of the neutrinos detected by IceCube come from the collision of cosmic rays in the earth's atmosphere. Cosmic rays are high energy particles which can originate from the sun, galactic sources, and other astrophysical objects outside the galaxy. These particles interact mainly with the oxygen and nitrogen atoms in the atmosphere, creating hadronic showers. A dominant fraction of the secondaries in the shower consists of light mesons (i.e. pions and kaons). When a charged pion ($\pi^{\pm}$) decays, it produces an outgoing muon and a muon neutrino. The muon can further decay into an electron, an electron neutrino, and a muon neutrino. The same chain of reactions can also occur in charged kaons ($K^{\pm}$), as shown below,

$$\begin{aligned} \pi^{\pm}(99.99\%)/K^{\pm}(63.55\%) &\longrightarrow \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu}) \\ \mu^{\pm} &\longrightarrow e^{\pm} + \nu_e(\bar{\nu}_e) + \bar{\nu}_{\mu}(\nu_{\mu}) \quad . \end{aligned} \tag{3.4}$$

In addition, all the kaons ($K^{\pm}, K^0_S, K^0_L$) can also decay into charged pions triggering the identical decay channels (shown in Equation (3.4)) to produce neutrinos. These neutrinos form the primary *conventional* component of atmospheric neutrinos. Schematic diagram of an example cosmic ray air shower is shown in Figure 3.12. Equation (3.4) shows that the decay of each light meson produces approximately 2/3 ν_{μ} and 1/3 ν_e. However, this flavour composition changes when the neutrinos arrive at the detector due to neutrino oscillation. Mixing of the flavours depends on the neutrino energy and propagation length and can introduce a small but nonzero ν_{τ} contribution. Heavier charm mesons ($D^{\pm}, D^0, D_s$) are also created in the cosmic ray shower with several orders of magnitude smaller cross sections and carry a higher fraction of the shower energy than the light mesons. The heavier mesons' short lifetime ($\lesssim 10^{-12}$ s) triggers their fast decay without losing much energy in the prior hadronic interactions. Neutrinos produced from these decays are treated as an

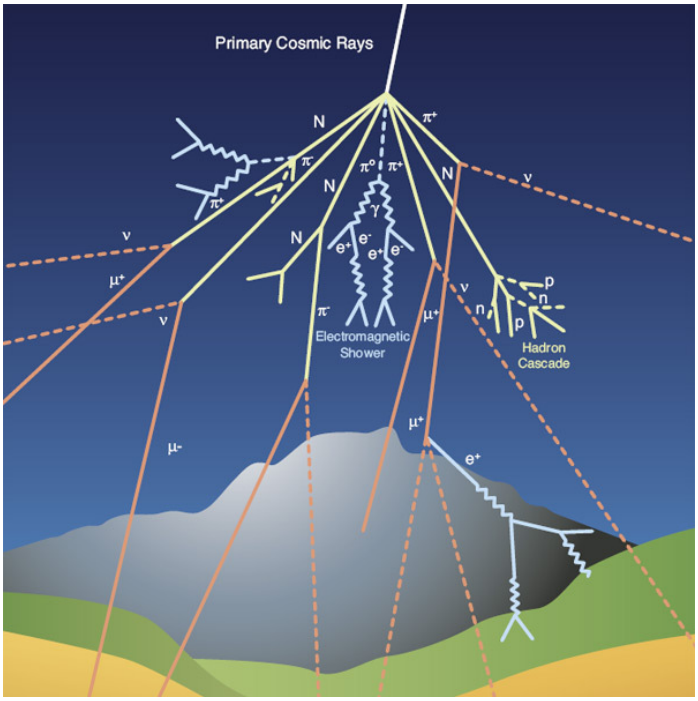


Figure 3.12: Diagram showing the interaction processes in a typical cosmic ray air shower event, from Ref. [140].

additional *prompt* component of the atmospheric neutrinos and have harder energy spectrum with a much smaller contribution compared to the conventional component.

The atmospheric neutrino fluxes are estimated using the initial cosmic ray spectrum and the hadronic interactions. Cosmic rays primarily consist of protons and a small fraction of helium and other heavier nuclei. This work uses the HillasGaisser2012 (H3a) cosmic ray model [141] and SIBYLL2.3c hadronic interaction model [142] to compute the atmospheric conventional neutrino flux. The cosmic ray model has been developed to calculate the cosmic ray flux, which follows a power law spectrum,

$$\frac{\mathrm{d}N(E)}{\mathrm{d}E} \propto E^{-\gamma} \quad , \tag{3.5}$$

where a global fit value of $\gamma = 2.7$ for energy below 10^6 GeV is reported in Ref. [48]. Above this energy, the spectrum has softer power law dependence with $\gamma = 3.1$ [143]. The model called BERSS [144] performed an extensive QCD calculation of heavier quark production in developing the prompt component of the atmospheric neutrino flux and is used in this work. The experimental measurements for the cosmic ray flux and the calculated atmospheric neutrino spectrum from different cosmic ray and hadronic interaction models are shown in Figure 3.13.

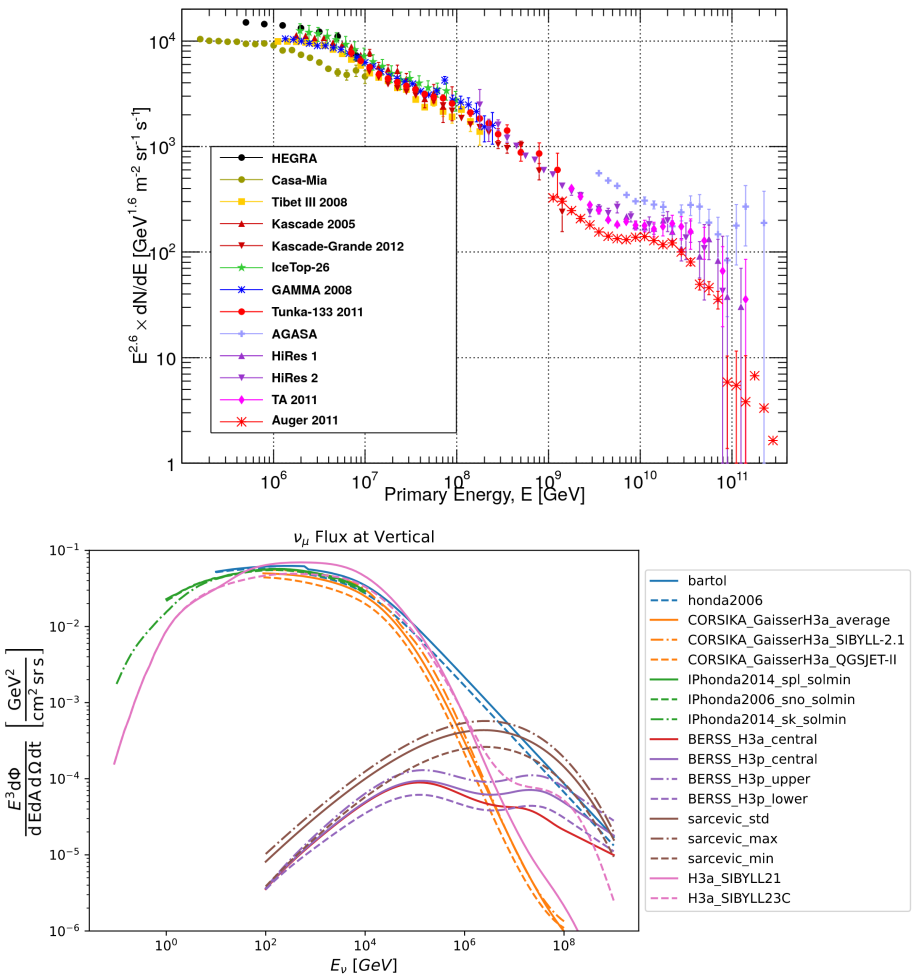


Figure 3.13: (Top) Cosmic ray flux measurements from various experiments [141]. (Bottom) The atmospheric neutrino flux models calculated using different cosmic ray and hadronic interaction models [145]. The fluxes labelled as 'H3a_SIBYLL23C' and 'BERSS_H3a_central' are used for the atmospheric conventional and prompt fluxes for this work, respectively.

3.4.2 Astrophysical Neutrinos

The highest energy neutrinos detected in IceCube come from astrophysical origins. The cosmic accelerators that are assumed to generate very high energy cosmic rays should also produce neutrinos since any accelerated protons or nuclei interacting with matter are predicted to generate neutrinos. The neutrinos from these distant objects can reach the detector unhindered by matter and EM fields, unlike photons and charged particles, carrying information about their origin. Extragalactic sources like active galactic nuclei (AGNs), blazars, and pulsars have a supermassive black hole

that feeds on the surrounding matter creating an accretion disk around it. The accelerated high energy particles (e.g. protons and other heavier nuclei) can then interact with the nearby EM radiation field, dust or molecular clouds to produce secondary charged mesons. For example, the photoproduction process of an accelerating proton with the surrounding radiation can create a resonant Δ baryon to produce pions,

$$\begin{aligned} p + \gamma \rightarrow \Delta^{+} &\rightarrow \pi^{0} + p \\ &\rightarrow \pi^{\pm} + n \quad . \end{aligned} \tag{3.6}$$

While the first output creates a pair of gamma rays from the decay of π^0, the second output produces neutrinos using the same mechanism shown in Equation (3.4). Similarly, the flavour composition of these neutrinos at the origin is assumed to be $\nu_e : \nu_\mu : \nu_\tau = 1 : 2 : 0$. Assuming this flavour composition at the source, the neutrino oscillation over a large astrophysical distance provides an expected flavour composition detected at the earth and should be roughly equal, $\nu_e : \nu_\mu : \nu_\tau \approx 1 : 1 : 1$. The region covering the expected neutrino flavour composition detected at the earth for any arbitrary flavour composition at the source can be constructed using the same method and is shown in Figure 3.14. Measuring the astrophysical neutrino fluxes for different event topologies in IceCube can thus probe into the interaction mechanism of some of the highest energy cosmic accelerators in the universe. IceCube first detected astrophysical neutrinos in 2013 [137]. Since then, several studies have found concrete evidence of neutrino emissions from extragalactic sources [18, 19].

The total neutrino flux from these distant objects is modelled to be isotropic in direction and follow a single power law spectrum. The diffuse flux spectrum can be expressed as,

$$\frac{\mathrm{d}\Phi}{\mathrm{d}E} = \Phi_N \cdot \left(\frac{E}{100 \text{ TeV}} \right)^{-\gamma} \quad , \tag{3.7}$$

where Φ_N is the normalization and γ is the spectral index. From the observation of astrophysical $\nu_\mu + \bar{\nu}_\mu$ events in IceCube, the diffuse flux for this event topology is measured as $\Phi_N^{\nu_\mu + \bar{\nu}_\mu} = 1.44^{+0.25}_{-0.26} \times 10^{-18}$ GeV^{-1}cm^{-2}s^{-1}sr^{-1} and $\gamma^{\nu_\mu + \bar{\nu}_\mu} = 2.37^{+0.09}_{-0.09}$ [146]. The measurement is used in this work as the flux model for astrophysical neutrinos and is further discussed in Section 6.2.2.

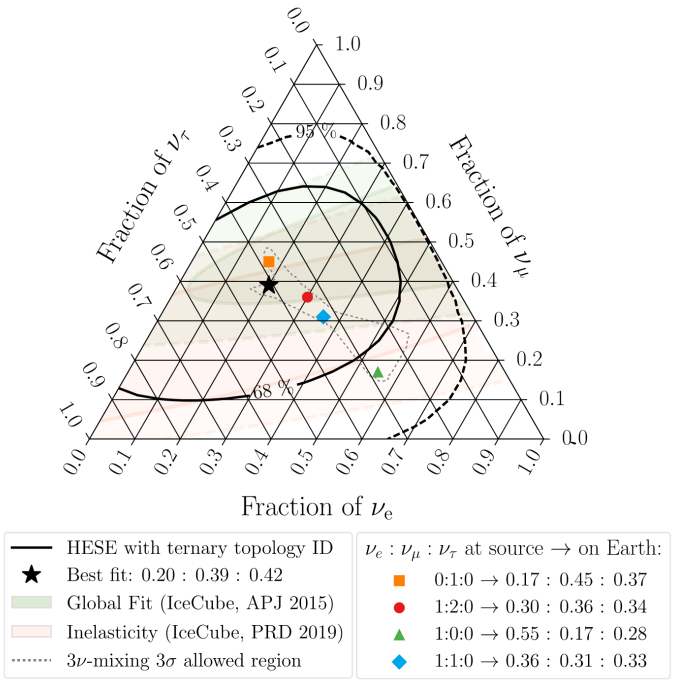


Figure 3.14: Measurement of the astrophysical neutrino flavour composition detected in IceCube [139]. Each side of the triangle refers to the fractional flavour contribution of the detected neutrinos. The star with the solid and dashed black line is the IceCube measurement. The enclosed region defined by the grey dotted line is the expected flavour composition at the earth for any arbitrary neutrino flavour ratio at the source.

3.5 Atmospheric Muons

Besides the atmospheric and astrophysical neutrinos, IceCube detects many atmospheric muons produced in the cosmic ray air showers. The atmospheric muon detection rate is approximately 2.15 kHz, six orders of magnitude higher than the neutrino detection rate of ~ 3.18 mHz [147]. As the atmospheric muons generate track events similar to the ν_μ CC DIS events, it is a background for most muon neutrino studies. Efficient rejection of this huge background is thus required for any precise measurement of neutrino-induced events. As the muons produced in the air showers enter the detector from the top, they are detected as down-going track events with zenith angles (defined as the angle between the incoming particle direction and the Z-axis of IceCube coordinate, as shown in Ref. [148]) $< 90^0$. Most of these events are vetoed by checking the detected tracks' reconstructed direction and starting position and discarding any down-going events starting at the detector edges. A small fraction

of the down-going atmospheric muons can be mis-reconstructed as up-going tracks, and the possibility of such backgrounds is discussed in Section 4.1.4. The down-going events that start inside the detector volume (excluding the DOMs that define the detector's outer surface) can be safely tagged as neutrino interactions. However, the effective flux for these events is reduced due to the muon veto, i.e. rejection of all through and down-going muon neutrino events. Contrary to the down-going tracks, the muons produced in the atmosphere at zenith angles $> 90^0$ travel through the earth and are absorbed before reaching the detector. As a result, most of the up-going (both through-going and starting) tracks in IceCube originate from neutrino interactions providing a higher purity ν-induced track events. Therefore, this work uses the up-going tracks to search for dimuon events.

Chapter 4

Simulation and Event Selection

The first step in developing a dimuon search analysis involves simulating the detector response of the relevant physics processes and the event reconstruction of those processes by the IceCube reconstruction software. The reconstruction of the simulated events is then used to determine the distributions of the event observables expected in the detector. The expected distributions are crucial in optimizing the event selection criteria and performing the search analysis (discussed in later chapters). In this chapter, we describe the simulation software framework developed for the dimuon events and the simulation of background processes relevant to this analysis. In addition, the later part of the chapter includes a discussion on the preliminary event selection procedure used in this work.

4.1 Monte Carlo Event Generation

4.1.1 Charm Dimuon Generator

The schematic diagram in Figure 4.1 shows the intermediate physics processes involved in producing the ν-induced charm dimuons. The software framework, `Charm Dimuon Generator`, implements these steps to inject the simulated interactions into the IceCube detector geometry. The physics processes can be divided into two primary steps in implementing the charm dimuon MC. In the first part, $\nu_\mu(\bar{\nu}_\mu)$ CC DIS interaction process produces the primary muon and the outgoing charm quark, followed by the hadronization process. The second part then simulates the charm hadron decay and interaction in ice to generate the secondary muon. The following steps in sequence form the complete simulation chain of charm dimuon event generation.

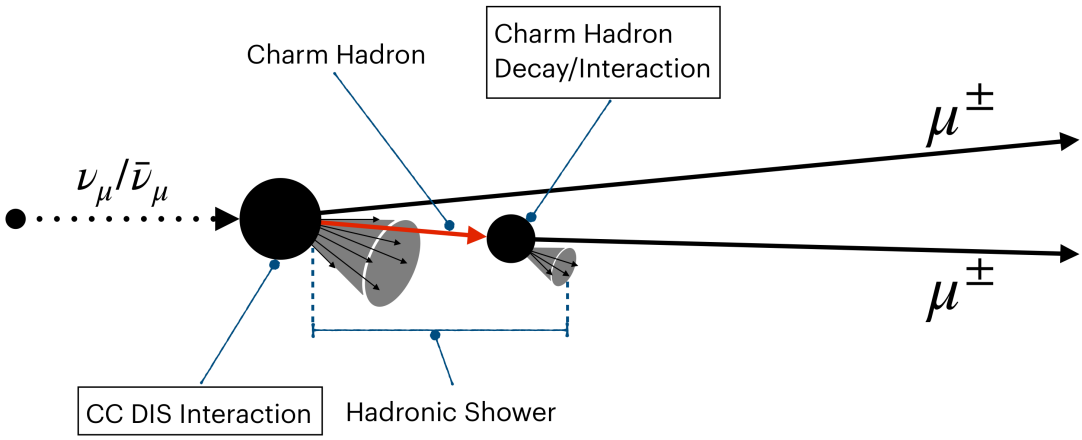


Figure 4.1: The steps involved in the charm dimuon production process. The hadronic shower component of the interaction is usually contained within $\sim$ 10 m of the interaction vertex. The interaction's high energy dimuons travel a significantly larger distance ($\mathcal{O}$(km)) than the shower size.

Energy and Geometry Sampling

The simulation starts with randomly sampling the incoming neutrino energy and the interaction geometry, i.e. the interaction vertex and direction of the incoming neutrino. The `LeptonInjector` software package [149] is used for this purpose. The step involves the independent sampling of the neutrino energy and direction, which are then used to construct an injection volume centred on the detector for sampling the interaction vertex. The atmospheric and astrophysical neutrino fluxes dictate the neutrino energy spectrum at the detector. However, the fluxes (discussed in Section 3.4) indicate low statistics for high energy events. In addition, they have associated uncertainties which need to be implemented in the simulation to account for the expected event rate fluctuation. These issues are addressed by employing the MC event weighting method, which allows the production of the simulated events with an arbitrary *generation spectrum* to achieve desired MC statistics. These events are later weighted to get the event rate expectation associated with the neutrino fluxes and their uncertainties. The neutrino energies for the charm dimuon events are sampled in the energy range $10^2 - 10^8$ GeV with a continuous power law spectrum $\mathrm{d}N/\mathrm{d}E \propto E^{-\gamma}$, where $\gamma = 1.5$. There are several configurable parameters (indicated in *italics*) related to the interaction geometry sampling. The incoming neutrino direc-

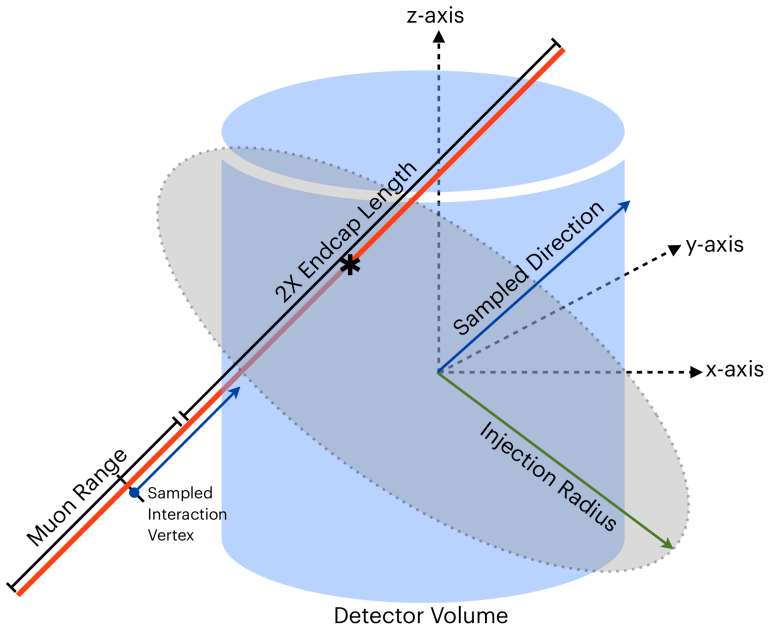


Figure 4.2: An illustration of the interaction geometry sampling using `LeptonInjector`.

tion is simulated by sampling the angles from a *zenith range*, $[90^0 - 180^0]$ (up-going neutrinos, sampled in cos(zenith)) and *azimuth range*, $[0^0 - 360^0]$. A circular disc with a radius (set by the *injection radius* parameter) of 700 m is then defined perpendicular to the neutrino direction and placed at the detector's centre. A point sampled evenly across the disc and the selected direction define the event's projected neutrino path as shown in Figure 4.2. The generation of an interaction vertex along the projected line requires calculating the maximum possible range for the outgoing muons. The maximum range is computed assuming one outgoing muon carrying the same energy as the incoming neutrino and using a parametric form discussed in Ref. [149]. In addition, an *endcap length* parameter with a value of 800 m is added along the projected line, on each side of the sampled point on the disc. The endcap lengths ensure the sampling of all possible interaction vertices which can lead to the generation of outgoing particles reaching the detector volume. The sum of the maximum muon range and the two endcap lengths define the line segment containing the possible interaction vertex, as shown in Figure 4.2. Finally, a uniformly sampled position on the line segment determines the interaction vertex of the simulated event.

Primary Interaction and Hadronization

For the generated neutrino energies, the target nucleons (proton and neutron) are sampled based on their composition in ice (i.e. 10 protons and 8 neutrons for each H_2O molecule). The ν_μ CC DIS interaction with the target nucleon is then simulated using `PYTHIA8.2` [150] and `DIRE` [151] framework to produce the outgoing muon and the charm quark followed by the hadronization process. `PYTHIA8.2` was developed with a focus on simulating generic particle collisions (such as beams of protons, electrons, and heavier nuclei) in laboratory setups. As a result, the standalone version of the software lacks a complete implementation of neutrino DIS interactions [152]. The issue is resolved using the `DIRE` plugin to supplement `PYTHIA8.2` for accurately simulating the neutrino DIS interaction with an improved parton shower algorithm. In addition, the configuration of the PDFsets is required to provide a complete description of the partons in the nucleons. CT14 PDFsets from the CTEQ collaboration [87] are used to simulate the charm dimuon interactions. The PDFset with Next-to-Leading Order (NLO) approximation is used for the primary hard interaction, and the Leading-Order (LO) PDFset is used for the hadronization (as the use of the NLO PDFset for the hadronization is discouraged due to the possibility of NLO approximation probing into unphysical Bjorken x, as discussed in Ref. [153]).

`PYTHIA` simulates the interactions in a vacuum, and the approximation works well as the detector medium has a negligible effect on the primary interaction and the hadronization due to the short time scale of the processes. However, the longer lifetimes of the very high energy charm hadrons (produced in the hadronic shower) can lead to their interactions with the detector medium before decay. In addition, this simulation framework only focuses on producing muons from the decay and interaction of the charm hadrons. As a result, the generation of the secondary muons from the charm hadrons in ice needs to be simulated separately. Therefore, the current simulation step stops immediately after the hadronization process without simulating the decay of charm hadrons in a vacuum. From the simulated interactions in this step, the inelasticity parameter (fraction of the incoming neutrino energy transferred to the hadronic system) and the 4-momenta of the outgoing primary muon and the charm hadron(s) are extracted to be used in the next steps of the simulation chain.

Charm Hadron Decay and Interaction

The charm hadrons produced in the previous step are then simulated to either decay semileptonically into a muon or interact in ice before decay. A parameterization technique is developed for producing the outgoing MC muons from the decay and interaction of the charm hadrons. A particle gun in `PYTHIA8.2` simulates the standalone decay of seven charm hadrons ($D^{\pm}, D_0, D_s^{\pm}, \Xi_c^{\pm}, \Lambda_c^{\pm}, \Xi_c^0, \Omega_c^0$) at 350 discrete energies equally spaced in the log scale in the energy range $10^1 - 10^8$ GeV. 10^6 events are generated for the decay simulation of each charm hadron and energy point. A parametric table of the muon energy distributions as a fraction of the parent charm hadron energy is created from the simulated events. In the ultra-relativistic limit, the fractional energy distribution of the muons is approximately independent of the charm hadron energies, as shown in Figure 4.3 for two example charm hadrons ($D^{\pm}$). The cumulative distribution functions (CDFs) (also shown in Figure 4.3 for all charm hadrons) are then evaluated from the parametric table to sample the muon energies based on the simulated charm hadron decays. This simulation framework imposes a cut on the individual muon energy threshold at 10 GeV (a conservative limit well below the threshold for IceCube to distinguish a muon from the hadronic shower) for efficient MC production of dimuons.

For the charm hadron interactions in ice, a software package called `chromo` [154] is

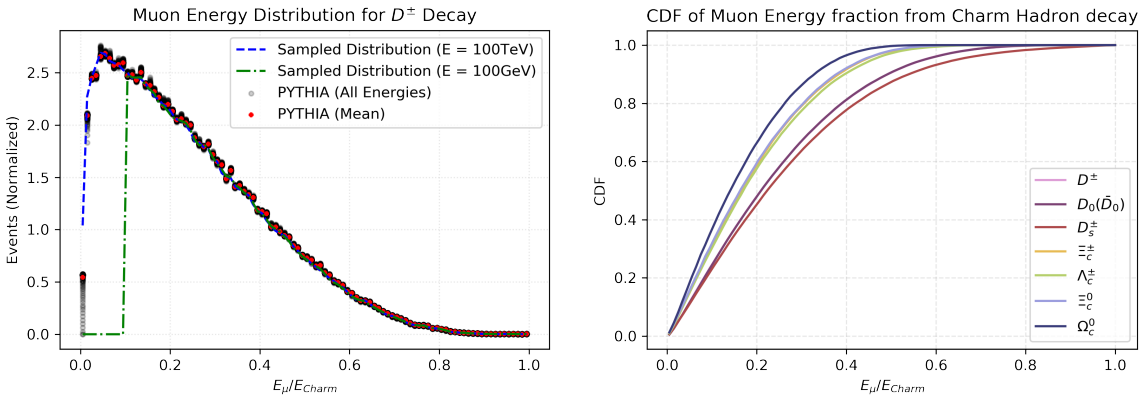


Figure 4.3: (Left) The fractional energy distributions of the outgoing muons sampled from the CDF in comparison with the simulated events for a charm hadron decay ($D^{\pm}$). The vertical dashed lines show the sampling cut-off where the muon energy falls below the 10 GeV threshold. (Right) The muon energy CDFs generated for all charm hadrons are shown as a fraction of parent charm hadron energy.

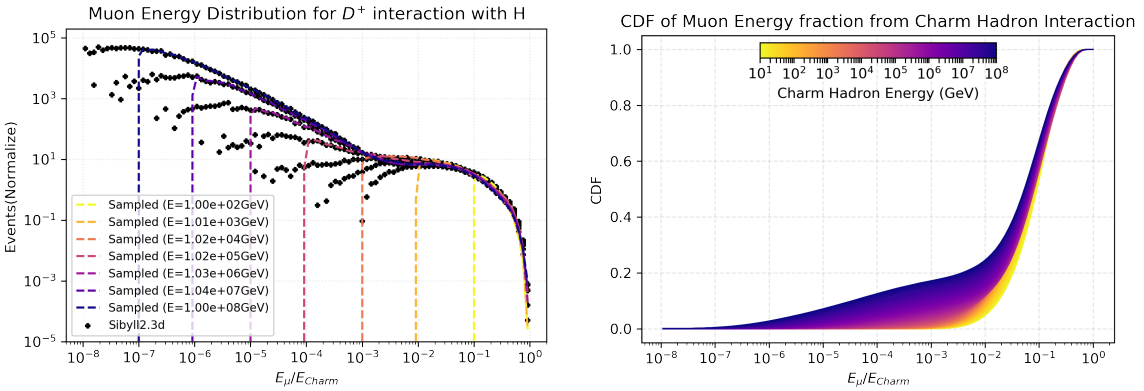


Figure 4.4: (Left) The sampled fractional energy distributions of the muons in comparison with the simulated events at different D^+ energies. The vertical lines of the sampled distributions show the muon energy threshold cut. (Right) The muon energy 2D CDF generated for D^+-H interaction at all energies.

used as it has been developed for MC simulation of hadronic interactions (primarily for cosmic ray models). A parametrization framework similar to the one developed for the charm hadron decay is implemented from the interactions of the charm hadrons with Hydrogen (H) and Oxygen (O) targets. The hadronic interaction model, SIBYLL2.3d [155] is used to generate 10^6 events for each combination of the seven charm hadrons, two target atoms (H and O), and 300 discrete energy values in the range $10^2 - 10^8$ GeV (equally spaced in log-scale). The parametric tables of the fractional muon energy from the interactions are then generated to sample the muon energies. However, unlike decay, the fractional energy distribution of the outgoing muon varies as a function of the charm hadron energy. Therefore, 2-D interpolated CDFs are created for the interactions (compared to 1-D CDFs for the decay) and are shown in Figure 4.4. The 10 GeV muon energy threshold cut is also applied here.

Since the charm hadron decay and interaction processes are developed separately, a method for selecting one of the two processes based on their relative probabilities is needed to decide on the origin of the second outgoing muon for a simulated event. The decay and interaction probability are the functions of decay and interaction lengths (denoted as l_d and l_i, respectively) which are expressed as,

$$l_d = \frac{E\tau}{mc}, \quad l_i = \frac{m_{ice}}{\rho N_A \sigma(E)} \quad , \tag{4.1}$$

where E, m, τ are the energy, mass, and lifetime of the charm hadron respectively,

m_{ice} = 18.02 g/mol is the molar mass of ice, ρ = 0.917 g/cm^3 is the density of ice, N_A is the Avogadro's number, and $\sigma(E)$ is the interaction cross section of the charm hadrons in ice. However, no concrete experimental data is available for the charm hadron interaction cross sections. After exploring several cross section models (shown in Figure 4.5), the already established work in Ref. [156] is found to be in good agreement with the other methods and, thus, is adopted to implement a similar framework in this work. A parametric form of the total cross section from `CORSIKA`

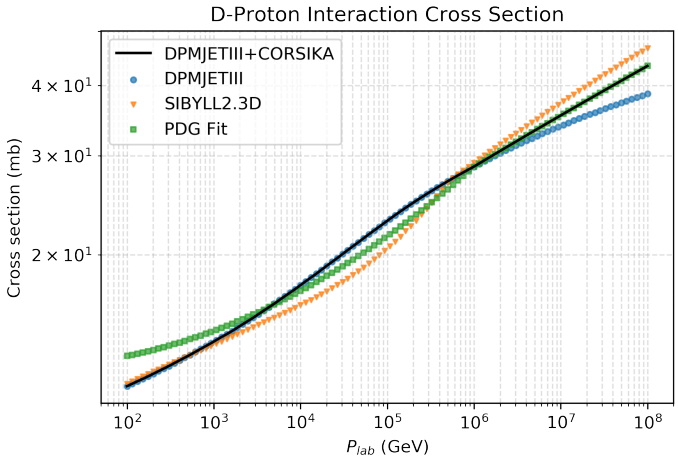


Figure 4.5: Different cross section models for proton-D meson interaction. The PDG fit is performed using the proton-pion cross section data on a second-order polynomial equation. 'DPMJETIII+CORSIKA' shown in the solid black line is used for this work.

[157] is used for the charm hadrons with energies above 1 PeV. Below 1PeV, the kaon-proton cross section $\sigma_{Kp}(E)$ from `DPMJETIII` [158] is used by scaling to match the `CORSIKA` cross section at 1PeV. The separate parametrization in `CORSIKA` for charm mesons and baryons results in a different scaling of the `DPMJETIII` cross section for the corresponding hadrons. The final forms of the cross sections for the charm mesons (σ_{Mp}) and baryons (σ_{Bp}) are expressed as,

$$\sigma_{Mp}(E) = \begin{cases} 0.69799\ \sigma_{Kp}(E), & E < 1 \text{ PeV} \\ \exp[1.891 + 0.2095\ \log_{10}(E)] - 2.157 + 1.263 \log_{10}(E), & E \geq 1 \text{ PeV} \end{cases}, \tag{4.2}$$

and

$$\sigma_{Bp}(E) = \begin{cases} 0.97618\ \sigma_{Kp}(E), & E < 1 \text{ PeV} \\ \exp[2.269 + 0.207\ \log_{10}(E)] - 0.9907 + 1.277 \log_{10}(E), & E \geq 1 \text{ PeV} \end{cases}, \tag{4.3}$$

respectively. The total cross section of the charm hadrons interacting in ice is calculated using the two equations above as the weighted sum of the hydrogen and oxygen cross sections. The decay and interaction lengths for the simulated charm hadrons are finally calculated using Equation (4.1) and shown in Figure 4.6 as a function of the hadron energy. The figure also shows that the charm hadrons reach their critical

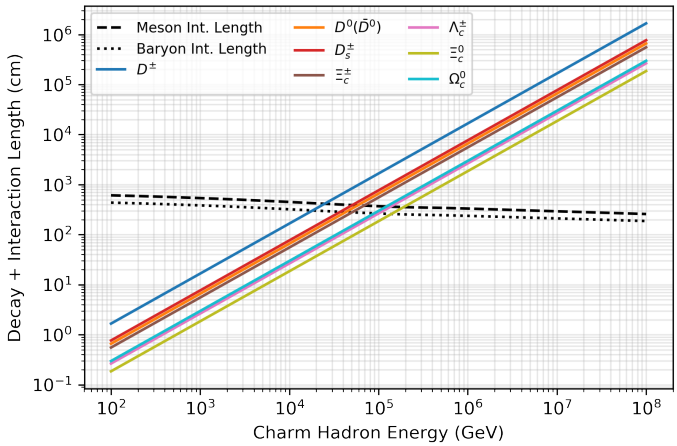


Figure 4.6: The decay and interaction lengths for the charm hadrons (mesons and baryons) in ice. The critical energies for $D^{\pm}, \Lambda^{\pm}$ are found to be 27 TeV, 129 TeV, respectively.

energies (at which the decay and interaction length is equal, different from lepton's critical energy introduced in Chapter 3) in the range of 20-200 TeV, above which the charm hadron interaction dominates over the charm hadron decay.

Finally, the following steps are repeated for each simulated charm hadron to produce the outgoing muon. Two distances s_d, s_i are sampled according to the decay and interaction probability distributions $e^{-s_d/l_d}, e^{-s_i/l_i}$, respectively. If $s_d \leq s_i$, the charm hadron decays, and muon energy is sampled using the parametric decay table. If $s_i < s_d$, the charm hadron interacts, and the outgoing muon energy is sampled from either H or O interaction table based on their composition in ice. The sampled muon is then injected into the final event output along the same direction as the parent charm hadron. Figure 4.7 shows the fractional distribution of the charm dimuon events generated from the decay and interaction processes.

In the hadronization step, a tiny fraction of the simulated events can produce more than one charm hadron. In this scenario, the same repeated process of the above steps is followed for each additional charm hadron in the event. After the production of

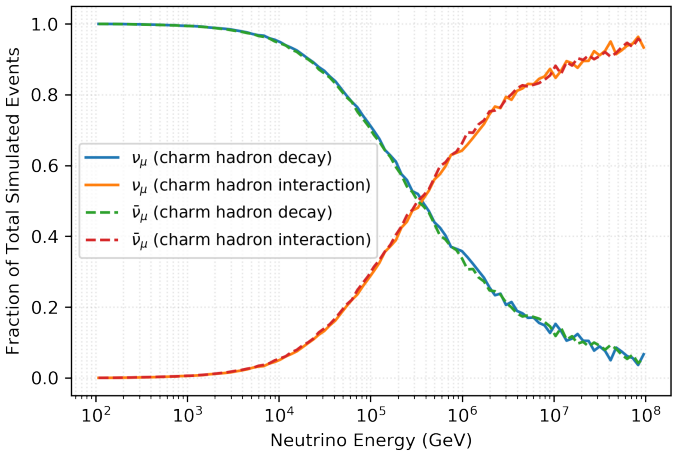


Figure 4.7: The fractional event distributions of the simulated charm dimuons from the charm hadron decay and interaction in ice. The muon production from the charm hadron interaction starts to dominate above ∼350 TeV of incoming neutrino energy.

the primary muon and the secondary muon(s) from the charm hadron(s), the excess energy of the system is calculated from the inelasticity parameter of the interaction and injected into a generic hadron along the outgoing charm quark direction to complete the list of outgoing particles for the charm dimuon event. As the cumulative Cherenkov light from a generic hadronic shower is produced within the detector simulation framework (discussed in Section 4.2), the detailed list of the involved shower particles from the primary interaction hadronization and the charm hadron decay or interaction is not relevant to this work.

Event Weight Calculation

Every MC event at the end of the simulation process carries an *event weight* (w_{MC}) which represents the probability of the event occurring given the true simulated quantities and is expressed in the unit of s^{-1}. In IceCube, a fluxless weight called *OneWeight* is constructed from the event generation details and is independent of any target neutrino flux model (i.e. atmospheric and astrophysical fluxes). The final weight w_{MC} can then be expressed as,

$$w_{MC} = OneWeight \times \Phi_{target} \quad , \tag{4.4}$$

where Φ_{target} is the target neutrino flux model. The calculation of *OneWeight* further depends on the various weight factors arising from different simulation steps and can

be expressed as,

$$OneWeight = \frac{w_{vol} \times w_{int} \times w_{prop}}{w_{gen} \times N_{events} \times N_{type}} \quad , \tag{4.5}$$

where w_{vol} is the weight associated with the sampling of the interaction position and direction within the injection volume. w_{int} is the interaction probability of the event. w_{prop} is the propagation probability of the neutrino to reach the detector from the earth's atmosphere. w_{gen} is the generation probability and depends on the neutrino *generation spectrum* used for energy sampling. N_{events} is the total number of simulated events which contain an equal number of generated neutrinos and anti-neutrinos. As a result, a type factor $N_{type} = 0.5$ is introduced to account for each half of neutrinos and anti-neutrinos in the simulation set. The expressions for each of the weight quantities are as follows:

- The volume weight is expressed as,

$$w_{vol} = \pi r_{inj}^2 \left(\cos\theta_{min} - \cos\theta_{max}\right)\left(\phi_{max} - \phi_{min}\right) \ \ \mathrm{cm}^2 \cdot \mathrm{sr} \quad , \tag{4.6}$$

  where r_{inj} is the *injection radius*, $[\theta_{min}, \theta_{max}]$ is the *zenith range*, and $[\phi_{min}, \phi_{max}]$ is the *azimuth range*. So the weight, w_{vol} is the total injection area multiplied by the solid angle over which the event positions and directions are sampled.

- The interaction weight is expressed as,

$$w_{int} = 1 - \exp\left[-\sigma_{int}(E_\nu) \cdot C_d \cdot N_A\right] \quad , \tag{4.7}$$

  where $\sigma_{int}(E_\nu)$ is the total cross section of the neutrino interaction producing the simulated final state particles as a function of the neutrino energy, C_d is the column depth which is computed by `LeptonInjector` using the range of the possible interaction positions (i.e. $2\times$*endcap length* + muon range) and considering the density of the material along the line, and N_A is the Avogadro's number.

- Since the events are injected close to the detector volume, a neutrino propagation weight w_{prop} is calculated separately using `nuSQuIDs` [159] software. This general-purpose software propagates neutrinos through different media accounting for neutrino oscillations and interactions with matter. For the current

simulation framework of charm dimuons, `nuSQuIDs` is used to calculate the transmission probability of neutrinos travelling from the atmosphere and reaching the detector volume. For the simulated zenith range, the neutrinos travel through different parts of the earth, and the PREM model [160] is used to provide its density profile. The computed propagation weight, w_{prop} for ν_μ as a function of the neutrino energy and zenith angle is shown in Figure 4.8. The plot shows that a significant fraction of the neutrinos above 100 TeV passing through the earth's core is absorbed before reaching the detector.

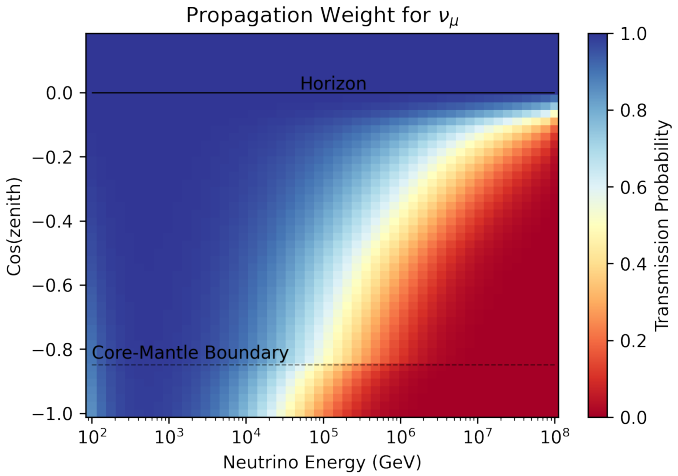


Figure 4.8: The propagation weight as a function of neutrino energy and cosine zenith for ν_μ. At zenith angles higher than the core-mantle boundary (horizontal dashed line), the neutrinos travel through part of the earth's core.

- The weight related to the generation spectrum is expressed as,

$$w_{gen} = \frac{E^{-\gamma}}{\int_{E_{min}}^{E_{max}} E^{-\gamma}\,\mathrm{d}E}\ \mathrm{GeV}^{-1} \quad , \tag{4.8}$$

  where E is the neutrino energy of the simulated event and $[E_{min}, E_{max}]$ is the neutrino energy range for the simulation. The spectral index, γ comes from the *generation spectrum*.

The total neutrino interaction cross section $\sigma_{int}(E_\nu)$ in Equation (4.7) for charm dimuon production is calculated by multiplying the ν CC DIS cross section σ^{tot}_{DIS} (from CSMS [84] calculation) with the fraction of events that produce charm hadrons

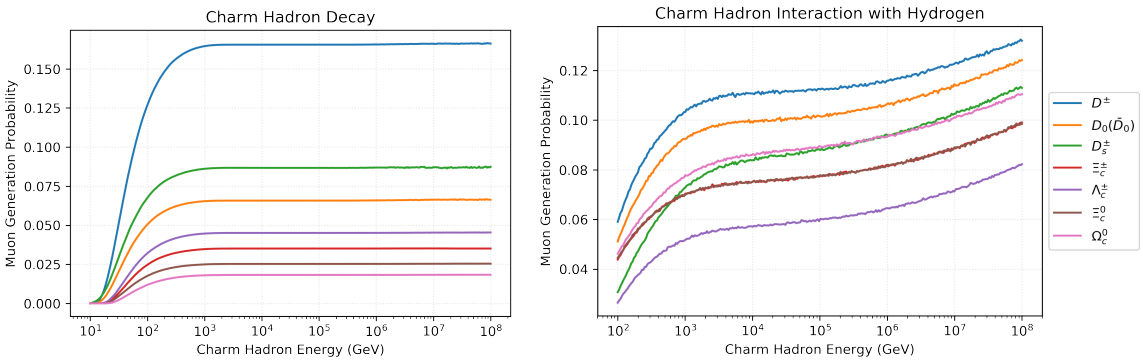


Figure 4.9: Probability of producing a muon with an energy greater than 10 GeV for all simulated charm hadron decays (left) and interactions (with H) (right) as a function of charm hadron energy. The drop in the probability below $\sim$ 1TeV is due to the imposed muon energy threshold cut.

(shown in Figure 2.6) and the probability of the charm hadron producing a muon either by decay or interaction (BR_μ). The final cross section is expressed as,

$$\sigma_{int}^{\mu\mu(charm)}(E_\nu) = \sigma_{DIS}^{tot}(E_\nu) \times \left[\frac{\sigma_{DIS}^{charm}(E_\nu)}{\sigma_{DIS}^{tot}(E_\nu)}\right] \times BR_\mu(E_{charm}^{decay/interaction}) \quad . \tag{4.9}$$

As discussed earlier, the primary interactions of the charm dimuons are simulated using the CT14NLO PDFset, which describes the interacting initial-state quarks (primarily d and s quarks) of the nucleon to produce the outgoing charm quarks. However, a recent measurement from the ATLAS experiment [161] reports more strange quark content in the nucleons compared to the previous measurements. A modern PDFset, CT18ANLO [80] shows better agreement with the recent result and, thus, is used to update the charm fractional cross section (the term within the square bracket) in Equation (4.9). The probability of producing a muon (BR_μ) is computed from the previously simulated charm hadron decays and interactions (used for generating the parametric tables) and is shown in Figure 4.9 as a function of the charm hadron energy (E_{charm}).

Calculation of the event weights completes the event generation process, and the flow diagram in Figure 4.10 shows the complete simulation chain of the `Charm Dimuon Generator` producing the final MC events. In the last step of the simulation chain, an `Event Converter` module stores the simulated events in an `I3File` format, the basic operating file structure within IceCube's software framework.

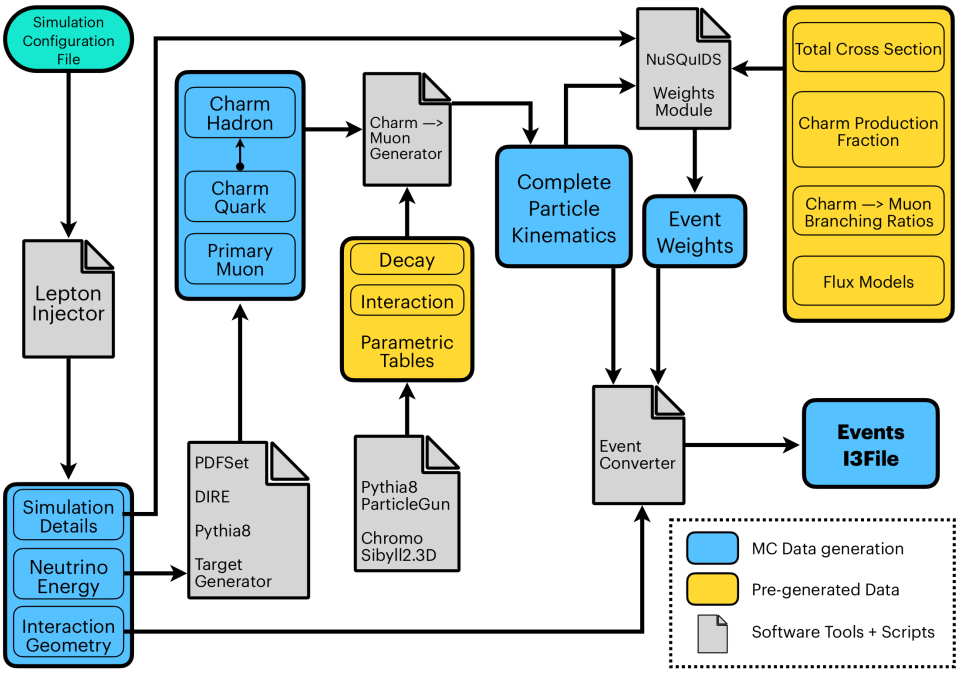


Figure 4.10: Simulation framework of generating charm dimuon MC events using `Charm Dimuon Generator`. The input simulation configuration file contains the configurable parameters of `Lepton Injector` discussed in the text.

4.1.2 Trident Dimuon Generator

Using the theory of the trident interaction described in Chapter 2, a general-purpose `Trident Dimuon Generator` is developed with the help of other external open-source software. The following sections describe the simulation chain to produce trident dimuon MC events.

Energy and Geometry Sampling

The energy and direction of the incoming neutrino and the interaction vertex are simulated using the same method described in the `Charm Dimuon Generator`. A relatively flat *generation spectrum* of E^{-1} is used to simulate high statistics trident dimuon MC at high energies where the cross section is much larger. The interaction geometry is simulated with the same injection volume configuration provided in the `Charm Dimuon Generator`.

Interaction Configuration Sampling and Event Generation

As discussed in Section 2.6, the trident interactions can occur under different nuclear interaction regimes leading to the use of different methods to simulate these events. For a given neutrino energy, a decision on the interaction regime is sampled based on the total cross sections from the elastic (coherent + diffractive) and inelastic contributions. In the case of interaction under the elastic regimes, the differential cross section of the photon momentum transfer (Q) is generated using Equation (2.14) and Equation (2.19). In these equations, the total cross section of the neutrino-photon system ($\sigma_{\nu\gamma}$) as a function of the CM energy is computed by simulating the process, $\nu_\mu + \gamma \rightarrow \nu_\mu + \mu^- + \mu^+$ in the `CalcHEP` event generator [162]. The neutrino-photon total cross section $\sigma_{\nu\gamma}$ and the differential distribution of Q for various neutrino energies are shown in Figure 4.11. A CDF is then constructed from the photon momentum transfer

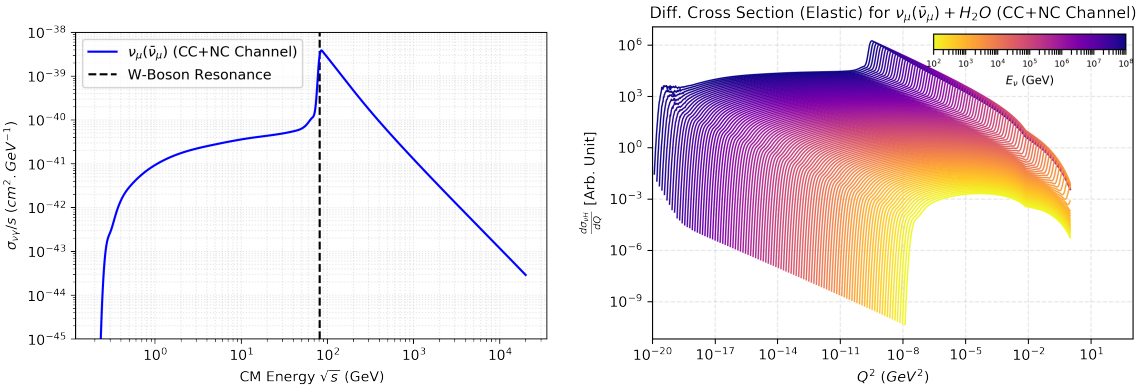


Figure 4.11: (left) Total cross section of the neutrino-photon system, which is used to calculate the differential cross section of the ν_μ trident process in water (H_2O) (right). In the differential cross section, a minor change in the shape of the photon momentum distribution can be seen at $Q^2 \sim 10^{-2}$ due to the change of interaction regime from coherent to diffractive.

distribution for a given neutrino energy to sample the interacting photon energy for the event. The generated neutrino and photon energy samples form the initial condition of the neutrino-photon interaction, and the configuration is transferred to the `CalcHEP` generator for simulating the outgoing particles.

The simulation of events under the inelastic regime follows a different treatment. If the interaction is sampled to occur in the inelastic (DIS) regime, the target nucleon is sampled from the number density of protons and neutrons in ice. The corresponding

PDF of the nucleon from the `LHAPDF` [163] interface is then used to describe the interacting quark and photon content. CT14qed PDFsets [164] are used to consider the interacting quarks (u, d, c, s) and only the inelastic component of the interacting photons. Finally, the neutrino interaction with the sampled nucleon's initial-state quark or photon is configured as the initial condition to simulate the outgoing particle kinematics in the `MadGraph` event generator [165].

Similar to the charm dimuon simulation, the detailed dynamics of the hadronic shower originating from the nuclear recoil (for the elastic regime) or the hadronization from the outgoing quark (for the inelastic regime) in the trident processes is not necessary for the context of the dimuon event topology. Therefore, the energy of the recoil nucleus or the quark is injected into a generic hadron which is later treated in the detector simulation for the Cherenkov light production.

Event Weights

The calculation of *OneWeight* for each simulated trident dimuon event also follows the same procedure discussed in the `Charm Dimuon Generator`. Apart from the *generation spectrum*, the only difference between the event weight calculation of the charm and trident dimuons comes from the total cross section of the neutrino interaction processes. Unlike charm dimuons, both the muons in the trident process are produced from the primary interaction. The total cross section for ν_μ trident interaction in ice from Ref. [103] is directly used as the dimuon interaction cross section $\sigma_{int}^{\mu\mu(trident)}$ to generate the interaction probability w_{int} in Equation (4.7).

Event Merging

During the event generation in `CalcHEP` and `Madgraph`, the interaction geometry sampled in the first step of the simulation chain (in energy and geometry sampling) is decoupled from the rest of the interaction configuration (energy of the interacting neutrino and target), and the process is simulated along a fixed z-axis (LHC-like geometry). The `event merger` module is implemented to perform a coordinate transformation of the event outputs (produced in z-axis) to recombine the interaction geometry, converting the events back to IceCube coordinates. The final outputs from all the simulated events are then stored in an `I3File` for later processing within the

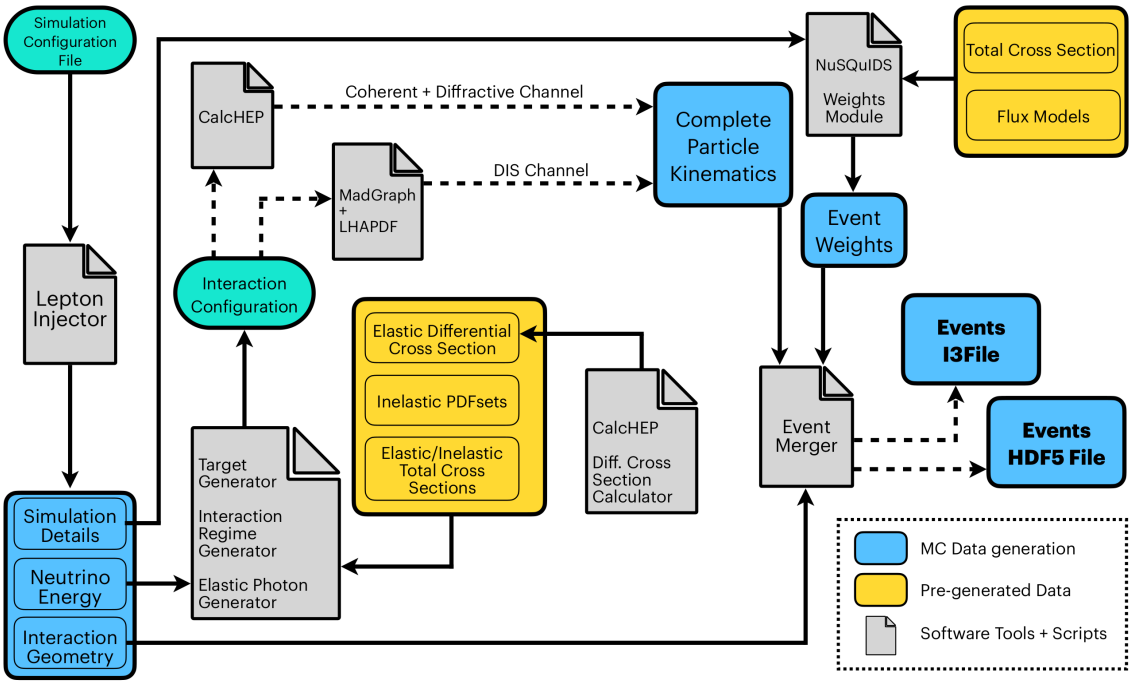


Figure 4.12: Schematic diagram of the process flow for the `Trident Dimuon Generator`. The dashed arrows denote the flow of the event generation process in one of the two parallel options. The final output of the generator can be stored either as `I3File` for use within IceCube or in `HDF5` file format for external use.

IceCube software framework. The `event merger` module also has the option to store the event outputs in an `HDF5` file format for external use with the publicly available version of the `Trident Dimuon Generator` [166].

The flow diagram of the steps discussed above for the simulation chain is shown in Figure 4.12. To validate the `Trident Dimuon Generator`, its simulation of W-boson production process ($\nu_\mu + H \rightarrow \mu^- + W^+ + X$) is compared with the results reported in Ref. [40]. Two MC datasets of the ν_μ W-boson production process (where the on-shell W boson is treated as the final state outgoing particle along with the outgoing muon) are generated using the `Trident Dimuon Generator` framework at discrete neutrino energies of 100 TeV and 1 PeV. The simulated events are then used to compute the energy distributions of the outgoing muon and the W boson. A comparison of the distributions with the reported result is shown in Figure 4.13 and found to be in good agreement with each other.

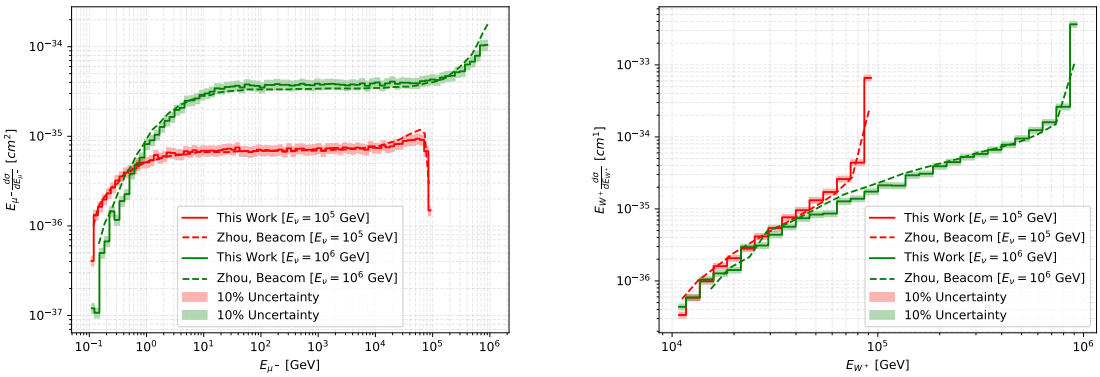


Figure 4.13: Comparison of the differential cross sections of the W-boson production from muon neutrinos. The outgoing muon energy distribution is shown on the left, and the W boson energy distribution is shown on the right.

4.1.3 Summary of Dimuon Processes

The dominating contribution to the dimuon events in both trident and charm production comes from the muon neutrino ($\nu_\mu + \bar{\nu}_\mu$) interaction channels discussed above. There are other channels which can also produce dimuon events with smaller contribution. Figure 4.14 shows the total cross sections from all possible trident and charm production channels which can produce dimuon events. The plot shows that most dimuon events are from charm production in ν_μ CC DIS interactions. The second largest dimuon cross section comes from the charm production in the ν_τ CC DIS interaction. However, tau neutrinos make up $\lesssim 10^{-4}\%$ of atmospheric neutrinos. Instead, they primarily come from the small astrophysical flux ($\sim$ 80 events/year, all flavour, all sky) above 100 TeV of neutrino energy. This makes the all sky (both up-going and down-going) dimuon contribution from ν_τ CC DIS interactions to be $\lesssim 0.1$ events/year, which is too small to be included in this work. The trident production from ν_μ (CC+NC channel) is the next leading channel (estimated at $\sim 10\%$ of the total dimuons above 10 TeV) and is included in the analysis. The ν NC DIS interaction (of all flavours) can produce a pair of outgoing charms, both of which can potentially decay into muons creating the dimuon events. The simulation of these events showed that the charm quark from the primary interaction vertex carries away a large energy fraction of the hadronic system, resulting in a small available energy for the other charm quark. A limit calculated from the simulation indicates that

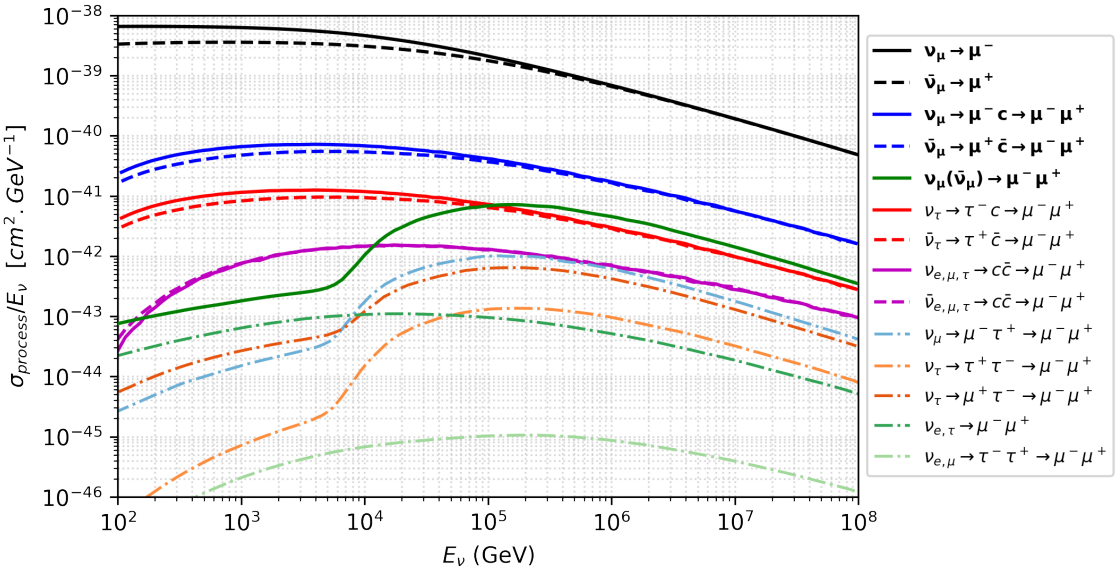


Figure 4.14: Total cross sections for all dimuon processes from trident [103] and charm production [84]. The processes in **bold** labels are included in this work. For comparison, the total cross section of ν_μ CC DIS interaction producing single muon background events is also included in the plot, showing in black solid and dashed lines.

$< 2.5 \times 10^{-5}\%$ of the NC interactions with charm production can have a second muon with energy $E_\mu > 10$ GeV. Therefore, the NC DIS channels have a vanishingly small contribution to the dimuon production and are not considered for the analysis. Figure 4.14 also shows the cross sections for other trident channels (CC only and NC only), which are significantly smaller than the dimuon channels discussed above and, thus, are not included in this work. In the hadronic shower component of the neutrino CC and NC DIS interactions, the production and decay of the pions and kaons can also contribute to additional secondary muons. However, these light mesons undergo a large number of nuclear interactions in ice due to their long lifetimes. As a result, their energy quickly falls below the required muon energy threshold of 10 GeV for the simulated dimuons.

4.1.4 Background Event Generation

The dominating background for the dimuon events comes from the ν_μ CC DIS processes, which produce a single $\mu^\pm$ and hadrons. The other background events considered for this work include ν_e and ν_τ CC DIS interactions and the atmospheric muons

from the cosmic ray showers. The MC datasets for the background processes used in this work are acquired from other physics analyses developed within the collaboration. A sterile neutrino analysis (involved the study of a new physics model and is not relevant to this work), called Matter-Enhanced Oscillations with Steriles (MEOWS), studied the energy and zenith distributions of the up-going track events in IceCube [167]. As these events predominantly come from the ν_μ CC DIS interaction, a high statistics MC dataset of the process was produced for developing the MEOWS analysis. In addition, ν_e and ν_τ CC DIS events were also simulated to consider the other background events for the MEOWS analysis. All the ν CC events in the MEOWS MC production were generated using `LeptonInjector`, and the corresponding event weights are computed using an accompanying event weighting software called `LeptonWeighter` [168]. The total cross section used in the event weight calculation of the ν_μ CC process considers the contributions from all the partonic interaction channels in the nucleon (regardless of the final states). A correction to the total cross section is, thus, applied to avoid the double counting of the charm dimuon production in these interactions. The updated cross section for the ν_μ CC process producing a single muon is then expressed as,

$$\sigma_{int}^{single\ \mu}(E_\nu) = \sigma_{DIS}^{tot}(E_\nu) - \sigma_{int}^{\mu\mu(charm)}(E_\nu) \quad , \tag{4.10}$$

where σ_{DIS}^{tot} is the total CC DIS cross section and $\sigma_{int}^{\mu\mu(charm)}$ is the cross section for charm dimuon production.

Any potential background events from the atmospheric muons due to their misreconstruction as up-going tracks are also considered in this work. The simulation of the down-going atmospheric muons used in the analysis is implemented with `CORSIKA` [157] software package and produced by an official simulation production (SimProd) group within IceCube. A summary of the simulation details for all the signal and background processes for this work is shown in Table 4.1.

4.2 Detector Simulation

In the next step, all the generated events discussed in the previous section are processed through the detector simulation. The IceCube software framework contains

Simulation Details	**Physics Processes**					
	ν_μ-Charm Dimuon	ν_μ-Trident Dimuon	ν_μCC	ν_eCC	ν_τCC	Atm. Muons
Energy Spectrum	$E^{-1.5}$	E^{-1}	E^{-2}	E^{-2}	E^{-2}	$E^{-2,2.6}$
Energy Range (GeV)	100-10^8	100-10^8	100-10^6	200-10^6	200-10^6	600-10^{11}
Zenith Range	90-180^0	90-180^0	80-180^0	80-180^0	80-180^0	0-90^0
Azimuth Range	0-360^0	0-360^0	0-360^0	0-360^0	0-360^0	0-360^0
Inj. Radius	700m	700m	800m	800m	800m	-
Endcap Length	800m	800m	1200m	1200m	1200m	-
No. of Events	2.10^7	6.10^6	5.10^8	5.10^8	5.10^8	$\sim 7.10^9$

Table 4.1: The parameter details for the simulation of all physics interactions considered in this work. Multiple `CORSIKA` MC datasets with two different energy spectrums are used for the atmospheric muon simulation. Although the signal MC processes are generated with higher maximum energy (10^8 GeV) compared to the background process, the final analysis cuts result in an event selection with energies much below 10^6 GeV as discussed in later chapters.

multiple components (referred to as 'modules') to simulate the full detector response of the events. Most of these modules require a detector description (e.g. the DOM properties and coordinates) provided via a Geometry Calibration Detector (GCD) file during the simulation process. The collaboration has released some of the modules to the public under an open-source license. The primary steps of the simulation process using these modules are discussed below.

4.2.1 Particle Propagation

The outgoing secondary particles from the event generations undergo the interactions discussed in Section 3.1 during their propagation through the ice. A software module called `PROPOSAL` [169, 170] simulates the muon propagation in the detector by modelling both the ionization and radiative loss processes. The muons reaching the

detector volume can travel through different media (ice, bedrock, air) depending on their energy and direction. At the bottom of the detector, the glacial ice has a boundary with the bedrock, which is part of the earth's crust. The ice overburden on the top of the detector volume ends at the surface, and the propagation medium changes to the earth's atmosphere. `PROPOSAL` coherently propagates the muons through these different media using a description of the IceCube detector volume and its surrounding environment. The software also simulates the propagation and decay of $\tau^{\pm}$ leptons.

The electromagnetic (EM) and hadronic shower development from the outgoing $e^{\pm}$ and hadrons are simulated using the cascade Monte Carlo (`cmc`) module [171]. As the EM shower produces only three types of particles (γ and $e^{\pm}$) through Bremsstrahlung and pair production processes, the primary particle's energy is dissipated via the cumulative Cherenkov radiation of the full shower with minimal variation. Therefore, the total Cherenkov light yield from a pure EM shower is proportional to the energy of the primary lepton. The hadronic shower development involves rather more complicated nuclear interactions producing a broad range of particles. The neutral hadrons and neutrinos produced in the shower do not contribute to the Cherenkov light emission resulting in a large variation of the cumulative light yield for a given hadron energy. The detailed simulation of the shower development for the energy scale in IceCube is computationally expensive, and instead, `cmc` uses various energy-dependent parametric models to produce the showers. This technique does not rely on the specific primary particle. Thus, any particle (e^-, e^+, or γ for the EM shower and any generic hadron for the hadronic shower) with the same energy follows the same corresponding parametrization method to simulate the EM and hadronic showers.

4.2.2 Cherenkov Photon Simulation

This step simulates the collective light emission from the propagated particles and the tracking of each photon through the ice to create the detected light in the DOMs for the event. A software package called `CLSim` [172] is used for this purpose. The simulation treats each final state charged particle in the EM and hadronic showers and each small segment along the propagated muon tracks as point-like light emitters. The Cherenkov photons from these emitters are created based on their energy and direction. `CLSim` then uses a ray tracing algorithm to propagate the photons through

ice and saves the ones that reach the DOMs. The saved photons are further processed to generate more realistic light yields in the respective DOMs, and the outputs are stored as MC photoelectrons (MCPEs). The conversion of the photons into MCPEs depends on several DOM characteristics, such as angular acceptance, quantum efficiency, and DOM efficiency. `CLSim` accesses these properties from either the GCD file or the configurable parameters during the simulation process.

Propagation of the photons through ice requires a complete description of the scattering and absorption properties of the ice. SPICE3.2 ice model describes the bulk ice in this simulation to produce the nominal MC dataset. However, the uncertainties related to the modelling of the ice properties lead to systematic effects in the photon propagation, referred to as *bulk ice systematics*. The uncertainty related to the hole ice modelling (called *hole ice systematics*) is considered a local phenomenon around the DOMs and is implemented using DOM angular acceptance of the simulated photons. The photon propagation and the MCPE generation must be re-simulated for each variation of the ice models around the nominal values to quantify the effects of these systematic uncertainties. In addition, the DOM efficiency factor directly affects the simulation of overall detected light for an event and also contributes to non-negligible systematic effects due to the related uncertainty. The treatment of these systematic effects is implemented in the analysis and discussed in Section 6.2.

4.2.3 DOM Simulation

In the next step, the MCPEs are processed by simulating the PMT response and the DAQ system to produce the final digitized event output, as seen in the detector. During the conversion of the MCPEs into PMT pulses, the simulation process also introduces noise hits arising from the dark noise, prepulse, and afterpulse associated with the PMTs. The modelling of the DOM on-board electronics then digitizes the PMT signals to determine the charge and timestamp of the simulated waveforms. The final outputs are stored as the simulated raw detector data of the MC events. This concludes the event simulation process, and all the subsequent processing of the MC simulation has identical treatment as the data.

4.3 Event Reconstruction and Preliminary Event Selection

The first stage of the event filtering and reconstruction process (known as 'Online Filtering') deploys several real-time triggering criteria on the data stream to discard non-physical noise events, as discussed in Section 3.2.2. High-level event reconstruction algorithms then use the pulse series (the dictionary containing all the hit charge and time information mapped to their respective DOMs) to reconstruct the properties of the secondary particles from the primary interaction. As this work focuses on identifying the up-going dimuon track events, the first step is to reject background events such as cosmic ray muons, $\nu_{e/\tau}$ CC events, and ν NC events. This is a common step for most IceCube analyses involving the study of ν induced track events. Several collaboration members have made extensive effort to develop a filter to select a high purity sample of up-going track events [134, 173]. The sample is referred to as MEOWS *platinum selection*, and a detailed description of the selection process can be found in Ref. [134]. The sample production method is adopted as a preliminary event selection in this work, and a summary of the selection steps is described below.

The platinum selection of the track events starts by triggering the 'Online Muon Filter', a part of the Online Filtering system. A track event's direction and position reconstruction require fitting five parameters (two angles for the direction and three coordinates for the position). The most basic reconstruction, known as the *LineFit*, performs a least square linear regression on the time of the first pulses in each hit DOMs. The results from the LineFit are used as the initial guess for better fitting algorithms based on maximum likelihood estimation. The single photoelectron fit (*SPEFit*) and multi-photoelectron fit (MPEFit) account for the Cherenkov light emission profile and the scattering and absorption of light in ice in their likelihood models to produce improved track reconstruction [129, 174]. This step of the reconstruction process is internally known as *Level 2 (L2)* reconstruction. A set of 'precut' criteria (e.g. cut on the reconstructed zenith angle) are applied to the reconstructed track to primarily select up-going events and reject a significant fraction of the atmospheric muons. An event-splitting algorithm called *Topological Splitter* [175] is designed to check for pulses from coincident atmospheric muons in an event

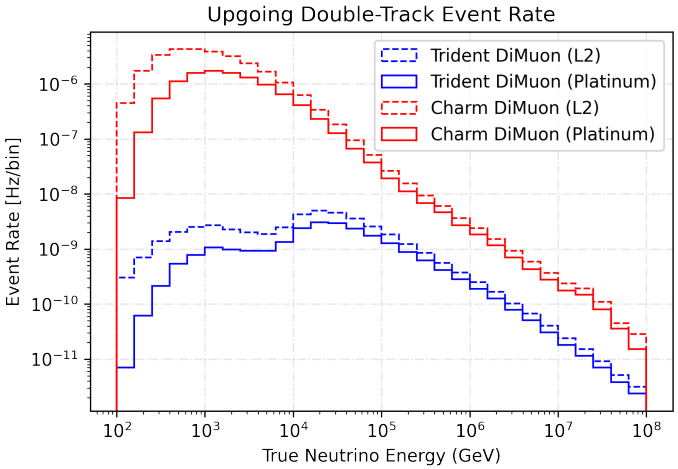


Figure 4.15: Comparison of expected signal events in the detector as a function of true incoming neutrino energy. Solid lines show the filtered track event rate at L2, and the dashed lines show the distribution of the events passing through MEOWS platinum event selection.

and splits the pulse series into multiple sub-events. The reconstruction algorithms are re-run on each sub-event, and the one with the best up-going track reconstruction is selected for further processing. The split pulse series of the selected sub-event is internally referred to as *TTPulses* and is used in the later part of the analysis (in Chapter 5). Another set of cut criteria are applied using various goodness-of-fit parameters and properties of the reconstructed track to reject the tracks with poor reconstructions. Finally, a sophisticated track energy reconstruction algorithm called *MuEx* [176] is applied to determine the energy of the track event. The output from MuEx only represents the measurement of the energy observed inside the detector and is denoted as E_{proxy}.

The MEOWS platinum selection sample contains $> 99.9\%$ ν_μ-induced up-going track events [134]. Therefore, this selection process reduces the analysis problem to separating dimuon neutrino events from the single muon ν_μ CC background. An event classification method is developed to address this problem and discussed in more detail in Chapter 5. A comparison of the expected dimuon event rates for the L2 and platinum selection in IceCube is shown in Figure 4.15. $\sim 60\%$ of the L2 level dimuon events pass the selection criteria for the platinum sample. Figure 4.16 shows the expected rate of the events passing the platinum selection as a function of the

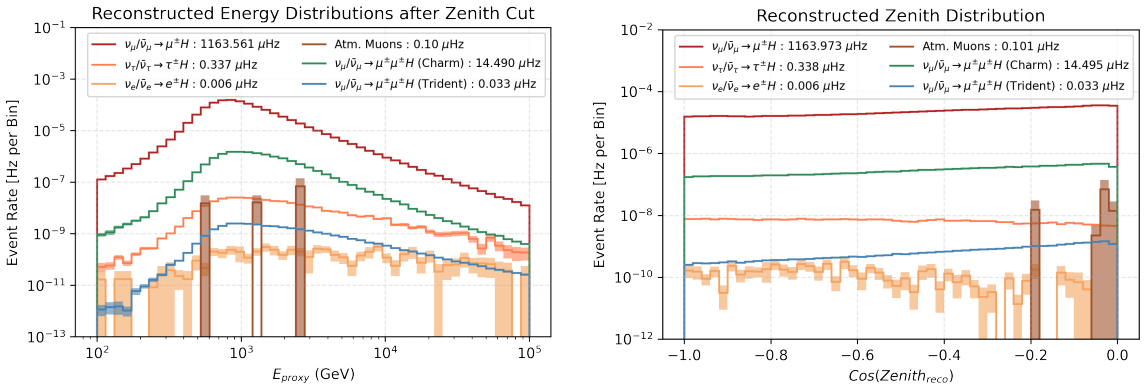


Figure 4.16: Reconstructed energy (left) and cosine zenith (right) distribution of all the signal+background processes after passing the platinum selection. The event distributions include a zenith cut of $[90^0 - 180^0]$ to consider only the up-going track events. The shaded region around the bins shows only the MC statistical uncertainty.

reconstructed energy and cosine zenith for all the physics processes considered in this work.

Chapter 5

Classification of Dimuon Events

The high-purity platinum sample of the up-going track events significantly reduces the background ($< 0.1\%$ contamination) from cascade and atmospheric muon events. However, Figure 4.16 shows that the majority of the events in the selected sample still come from ν_μ CC single muon production, which is two orders of magnitude higher than the dimuon event rate in the sample. Identifying the double-track signal events from a large number of single-track background events in the platinum sample remains challenging as the resolution of the two closely spaced muons is obscured due to their large overlap of the DOM hits in space-time. Therefore, the event identification task requires a powerful classification method to recognize the weak discriminating event signature patterns between single- and double-tracks. This is addressed by developing a novel machine learning (ML) based classification model, discussed in this chapter. Construction of the classification model starts with an initial characterization of the dimuon events based on their kinematical properties.

5.1 Dimuon Event Properties

The simulated dimuon events in this work have the two following underlying assumptions as implemented in `Charm Dimuon Generator` (Section 4.1.1) and `Trident Dimuon Generator` (Section 4.1.2),

- The energy of each muon has to be ≥ 10 GeV.
- Both the muons originate from the same primary neutrino interaction vertex.

Let us denote the leading and trailing energy muons in a dimuon event as μ_1 and μ_2, respectively, i.e. $E_{\mu_1} \geq E_{\mu_2}$. The true energy distributions of the two muons for the expected event rate in the detector from both charm and trident production are shown in Figure 5.1. The energy asymmetry between the two muons is defined as $(E_{\mu_1} - E_{\mu_2}) / (E_{\mu_1} + E_{\mu_2})$, and its expected distribution is also shown in the figure. The two plots indicate that the trailing muon energy can be much lower than the leading muon energy. Since the energy of the muons is directly related to the distance travelled, one of the tracks can stop well before the other track in a dimuon event. The

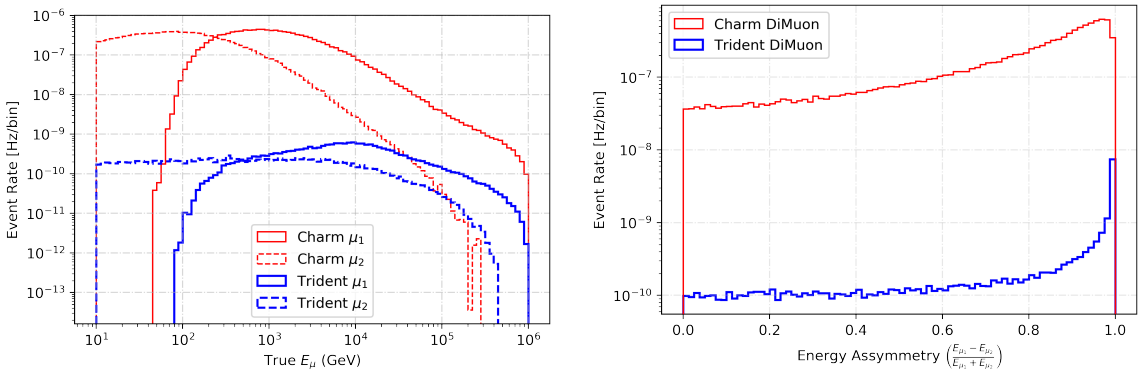
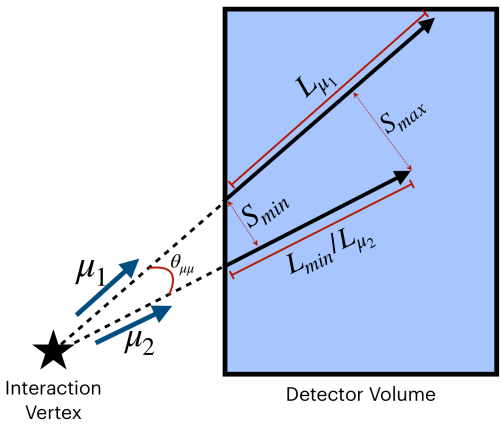


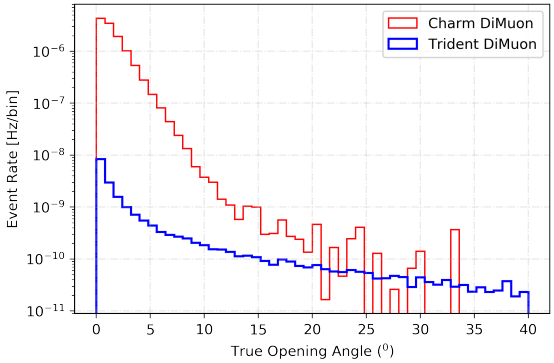
Figure 5.1: Expected event rate distributions as a function of the true muon energies (left) and dimuon energy asymmetry (right).

angle between the two muons (referred to as the *opening angle*, $\theta_{\mu\mu}$) dictates the lateral separation between the two tracks as their paths diverge from the interaction vertex. The lateral separations of the two muon track segments inside the detector volume can be computed assuming the detector is approximately a cylinder of radius 500 m and height 1000 m. The minimum track separation (S_{min}) is referred to as the lateral distance between the tracks when they enter or start in the detector volume, and the maximum track separation (S_{max}) is calculated at the point where either both the tracks exit the detector or the trailing muon stops inside the detector. An illustration of the double-track properties and the expected event rate distributions as a function of the true opening angle and the track separations are shown in Figure 5.2. The plots show that the two muons are almost collinear, i.e. $\theta_{\mu\mu} < 0.8^0$ and $S_{max} < 4$ m (based on the bin widths in the plots) for $\sim 95\%$ of the events and, thus, are almost indistinguishable from single-track background events.

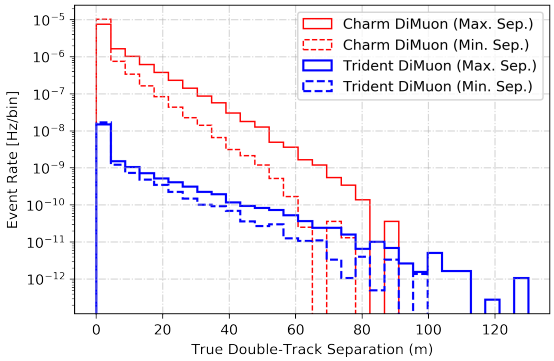
The lengths of the track segments inside the detector volume (denoted as L_{μ_1}, L_{μ_2}



(a)



(b)



(c)

Figure 5.2: (a) Illustration of the double-track properties computed from MC truth. (b) The expected event rate as a function of the opening angle ($\theta_{\mu\mu}$). (c) The minimum and maximum track separations inside the detector volume (S_{min} and S_{max}).

for μ_1, μ_2, respectively) together with S_{max} represent the characteristics of the simulated double-tracks. The minimum of the track segments (L_{μ_1} and L_{μ_2}), denoted as L_{min} and the maximum track separation S_{max} are used to divide all MC dimuon events into three subgroups, referred to as class A, B, and C events. The values of L_{min} and S_{max} are set to zero when the trailing muon of a simulated dimuon event stops before entering the detector. Only the leading muon from such an event contributes to the detected light creating a single-track event signature in the detector. The cut criteria to assign the events in each group are summarized in Table 5.1. The distribution of S_{max} vs. L_{min} in Figure 5.3 shows that the bin with highest event rate has $S_{max} < 5$ m, $L_{min} < 50$ m and belong to class C. Therefore, the event signature

Dimuon Subclass	MC Cut Criteria
Class A	$S_{max} \geq 25$ m, $L_{min} \geq 200$ m
Class B	$L_{min} \geq 200$ m
Class C	$L_{min} < 200$ m

Table 5.1: The cut criteria applied on each MC dimuon event to segregate the simulation dataset into three different quality double-track classes. The class A criteria contain one additional condition ($S_{max} \geq 25$ m) compared to class B and, thus, is a subset of class B dimuon events.

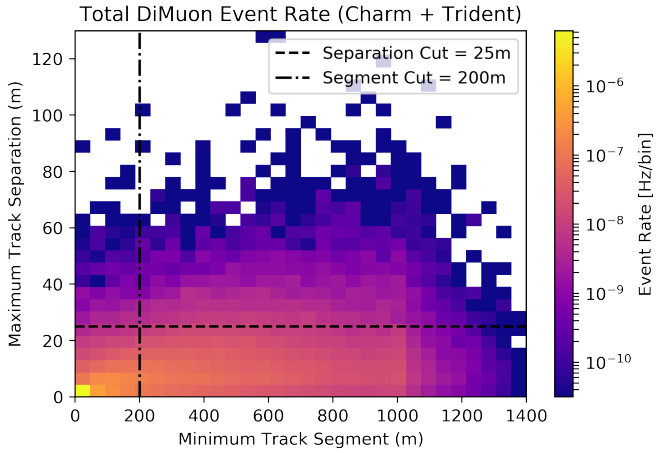


Figure 5.3: Expected event rate as functions of the minimum track segment (L_{min}) and maximum track separation (S_{max}). Two cuts (shown in dashed and dash-dotted lines) are used to divide all dimuon events into A, B, and C classes. The top right quadrant forms class A dimuons, the right two quadrants together form class B dimuons, and the left two quadrants form class C dimuons.

of the dimuons in class C contains either very little or no information from both the muons and is considered to be beyond the capabilities of the signal identification method developed in this work. All events with $L_{min} \geq 200$ m are defined as class B events and are further segregated into a class A subset of events with an additional criterion of $S_{max} \geq 25$ m. The high track separation and long track segments of class A dimuons in the detector make them the events most likely to have a clear double-track signature. Therefore, the search analysis developed in this work primarily focuses on identifying class A dimuons. An additional analysis step to select class B dimuon events is also implemented to help the ML algorithms achieve a better

performance in the event classification with an increased signal-to-background ratio. Figure 5.4 shows the reconstructed energy and cosine zenith distributions of the individual classes of dimuons from both charm and trident MC production. The plots indicate that approximately 3% of all simulated dimuons expected in the detector are class A events. A powerful event classification method is, thus, required to distinguish the highest quality (class A) dimuons which are three orders of magnitude smaller than the single-track background events in the platinum sample.

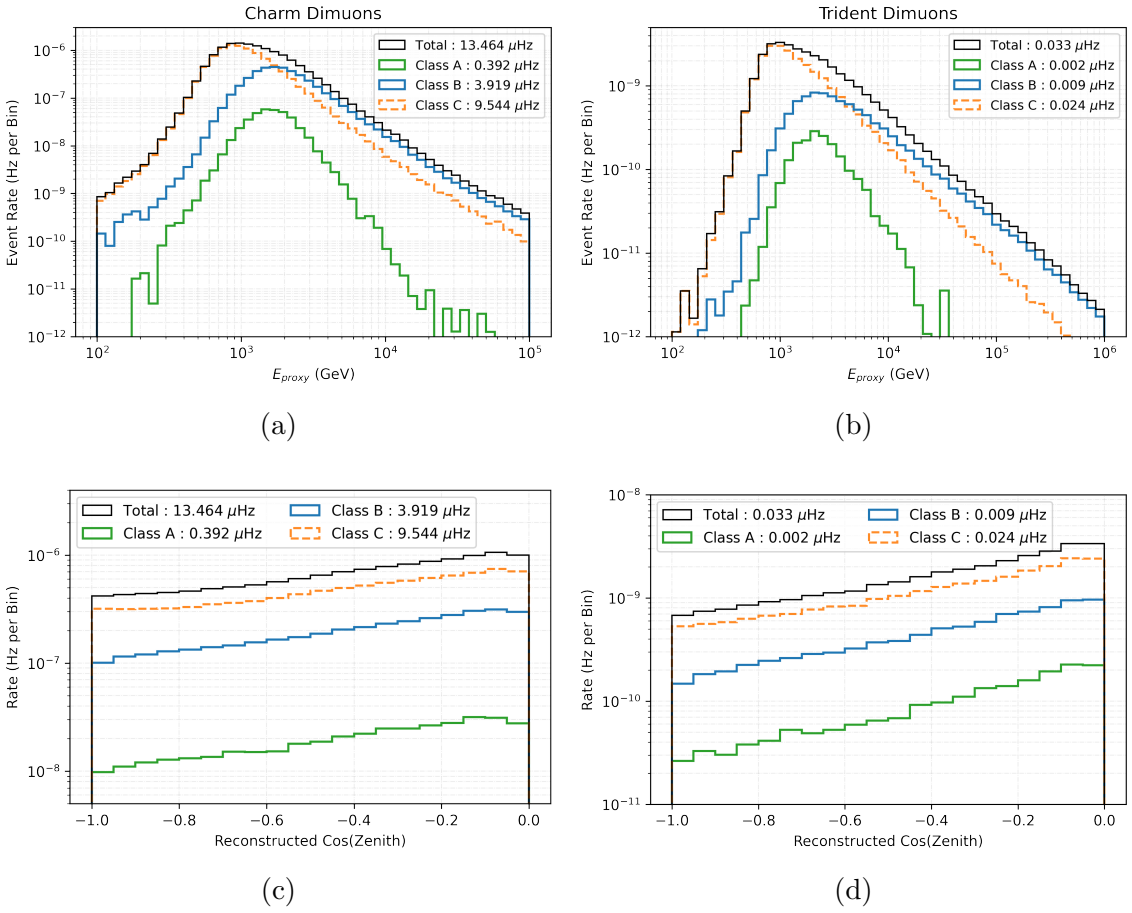


Figure 5.4: Expected event rate of the total and individual classes of dimuons. The two left plots show the reconstructed energy (a) and cosine zenith (c) for the charm dimuons. The two right plots show the reconstructed energy (b) and cosine zenith (d) for the trident dimuons.

5.2 Machine Learning Based Event Classification

A binary classification model is developed to identify each event in the platinum sample as either double-track or single-track. The model utilizes multiple high-level event properties constructed from the DOM hits and the reconstructed track of each event. These are referred to as *input features*. The details of the input features used in this work are discussed in Section 5.2.5. The model uses the input features for performing a multivariate analysis, which provides a mechanism to examine the relationships among the features and finds patterns and correlations in multiple variables to build the discriminant for the binary classifier.

Machine learning algorithms, developed based on the multivariate analysis framework are general-purpose tools that can perform various tasks like classification and regression. These algorithms achieve the task-solving process by using the input data to configure the model parameters. During the *training phase* (also called model training), the model automatically learns to optimize its parameters for the given task. The input dataset used for the purpose is called the *training dataset*. Finally, the trained model aims to accurately determine the outcome for a new input dataset. In recent years, the rapid growth of computational power and data generation has enabled the development of advanced machine learning algorithms which can perform human-level complex tasks like driving cars [177], image recognition [178], and natural language processing [179]. Deep learning is a family of such advanced algorithms developed based on a computational model called *artificial neural networks*. The complexity of the detectors and the vast amount of data collected in current high-energy physics experiments have paved the way for applying various deep learning algorithms in recent and ongoing physics analyses [180–182].

In IceCube, the performance of deep learning models in several analyses has shown improved results over conventional statistical analysis methods like maximum likelihood estimation [183–186]. This work also exploits a hybrid of neural networks to develop the dimuon binary classifier. The primary component of the model is a *graph convolutional neural network (GCNN)*, a type of deep learning algorithm that operates on graph-structured data. In addition, a basic neural network called a *multilayer perceptron (MLP)* is included as part of the hybrid neural network model. The de-

tailed mechanism of GCNN, MLP, and the complete dimuon classifier is discussed in the following sections.

5.2.1 Binary Classification Method

Let us denote each simulated event i as $(\boldsymbol{x_i}, y_i)$, where $\boldsymbol{x_i}$ is the set of input features for the event and y_i is the class label with $y_i \in \{0, 1\}$, where label 0 is for single-track, and 1 is for double-track events. The goal is to develop a classification model denoted as a function $f_{\boldsymbol{\theta}}$ with a set of model parameters denoted as $\boldsymbol{\theta}$. Given the input features, each event is mapped to its corresponding class label,

$$f_{\boldsymbol{\theta}}(\boldsymbol{x_i}) \mapsto y_i \quad . \tag{5.1}$$

Evaluation of the parameters $\boldsymbol{\theta}$ involves constructing a loss function $\mathcal{L}$ which measures the correctness of the model in mapping the events to their corresponding classes. An optimization process then finds the best parameter values by minimizing the loss measurement during model training. For binary classification, a loss function called binary cross-entropy is widely used and is expressed as,

$$\mathcal{L}(\mathbb{X}, \mathbb{Y}, \boldsymbol{\theta}) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log(f_{\boldsymbol{\theta}}(\boldsymbol{x_i})) + (1 - y_i) \cdot \log(1 - f_{\boldsymbol{\theta}}(\boldsymbol{x_i})) \quad , \tag{5.2}$$

where $\mathbb{X} = \{\boldsymbol{x_1}, \boldsymbol{x_2}, ..., \boldsymbol{x_N}\}$ is the set containing the sets of input features and $\mathbb{Y} = \{y_1, y_2, ..., y_N\}$ is the set of class labels for N events in the training set. In this framework, the output of the model $f_{\boldsymbol{\theta}}(\boldsymbol{x_i})$ can vary continuously in the range $[0, 1]$ and, thus, represents a 'double-trackness' measurement of the events (referred to as *classification score*). The binary cross-entropy loss as a function of the classification score is shown in Figure 5.5.

5.2.2 Graph Convolutional Neural Network

Let us denote a graph as the set $\{\mathcal{N}, \mathcal{E}\}$, where $\mathcal{N}$ is the set of nodes and $\mathcal{E}$ is the set of edges in the graph. The events in IceCube can be represented as a graph where each node is the input feature vector ($\overrightarrow{x}$) for a hit DOM, and each edge is the connection between two hit DOMs in an event. If a track event has M number of hit DOMs, and each DOM represents a D dimensional feature vector, the input *attribute matrix*

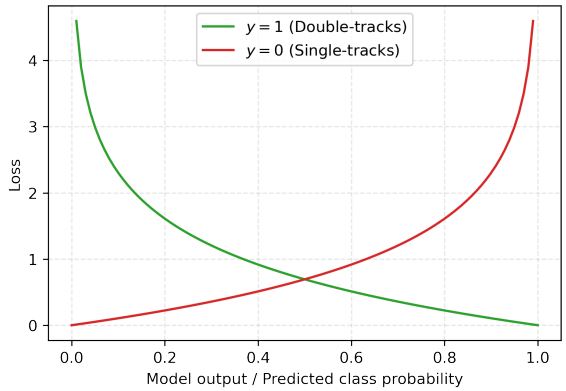


Figure 5.5: The binary cross-entropy loss as a function of the model output for class 0 (single-track) and class 1 (double-track) events.

$X \in \mathbb{R}^{M \times D}$ is formed by using the DOM features ($\overrightarrow{x}$) as row vectors and stacking them in M number of rows. The edge values for each pair of DOMs are used to construct the *adjacency matrix*, $A \in \mathbb{R}^{M \times M}$, representing all the DOM connections for the event. The attribute matrix X and the adjacency matrix A describe each event as graph-structured data and form the input space for a GCNN model.

In recent years, there has been a significant development in graph neural network algorithms, as discussed in Ref. [187]. The GCNN model used in the dimuon classifier follows the algorithm developed in Ref. [183, 188] and performs the following steps as part of the graph convolution operation.

- **Feature Transformation**

  A linear transformation is applied to the attribute matrix by multiplying it with a *weight matrix* W. The elements of W form the model's parameters (elements of the set $\boldsymbol{\theta}$ as introduced in the previous section) and are optimized during model training. The transformation converts the D component input feature vector into a D' component output vector for each node in the graph and is expressed as,

  $$X \cdot W \quad \in \mathbb{R}^{M \times D'}, \quad \text{where} \quad X \in \mathbb{R}^{M \times D}, \quad W \in \mathbb{R}^{D \times D'} \quad . \tag{5.3}$$

  For a single node i and an element $d' \in \{1...D'\}$ of the output vector, the above expression can be rewritten using the element-wise notation as,

  $$\sum_{d=1}^{D} X_{id} \cdot W_{dd'} \quad . \tag{5.4}$$

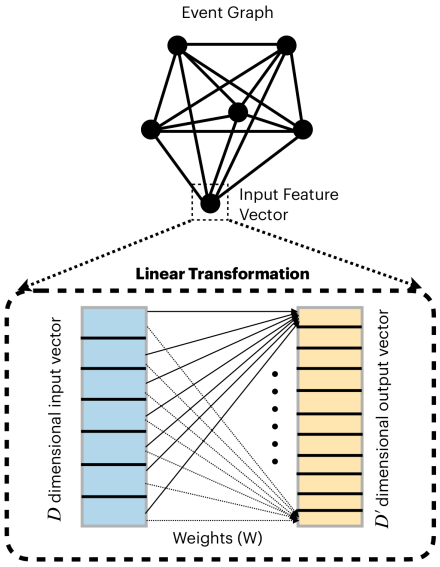


Figure 5.6: Illustration of the linear transformation on a node in the graph. Each arrow (solid or dotted) between the input and output vector represents an element in the weight matrix W.

An illustration of the transformation is shown in Figure 5.6.

- **Message Passing**

  In this step, each node in the graph updates its transformed features by aggregating features from the neighbouring DOMs. The aggregation process performs a weighted sum of the neighbouring nodes, with the edge connections in the adjacency matrix A dictating the weights. This operation enables all the nodes to pass their learned feature representation to neighbouring nodes and is referred to as a message passing framework. The message passing step can be expressed by updating Equation (5.3) as,

$$N^{-\frac{1}{2}} A N^{-\frac{1}{2}} \cdot (X \cdot W) \in \mathbb{R}^{M \times D'} \quad , \tag{5.5}$$

  where $N_{ii} = \sum_j A_{ij}$ is a diagonal matrix to normalize the edge weights for each DOM i from its neighbours $\{j\}$. The graph is fully connected (including self-connections) and undirected (equal weights in both directions of the edge) in this algorithm. Therefore, the adjacency matrix is symmetric (i.e. $A_{ij} = A_{ji}$) with non-zero elements. The element-wise form of the above equation for a

single node i and $d'^{\,th}$ element of the output feature vector can be expressed as,

$$\frac{1}{N_{ii}} \sum_{j=1}^{M} A_{ij} \cdot \left[\sum_{d=1}^{D} X_{jd} \cdot W_{dd'} \right] \quad . \tag{5.6}$$

The edge weights in the adjacency matrix represent the strength of the connections between pairs of DOMs and can be optimized via a set of learnable parameters during model training. This optimization mechanism comes from an additional algorithm called *graph attention* [189]. The construction of the adjacency matrix using graph attention for the dimuon classifier is discussed in Section 5.2.5. Figure 5.7 illustrates the message passing framework for a single node in the graph.

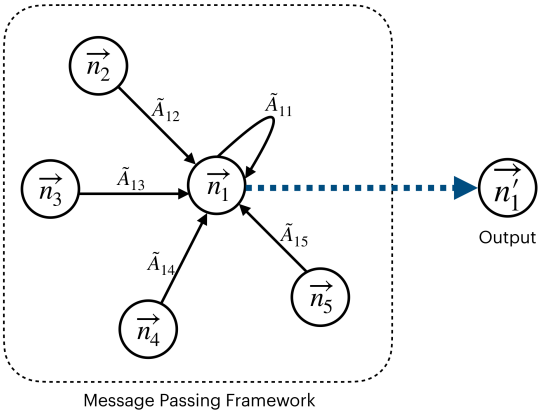


Figure 5.7: Diagram showing the message passing framework for a node in the graph. The vectors $\{\vec{n}_i\}$ represent the linear transformed feature vectors for respective nodes and $\{\tilde{A}_{ij}\}$ represents the normalized edge wights (adjacency matrix elements) between pair of nodes. The output of the message passing operation for node n_1 is denoted as $\vec{n'}_1$.

- **Nonlinear Transformation**

  In the final step, a nonlinear function, called the *activation function* (denoted as σ) is applied element-wise to the output matrix from the previous step. The nonlinear transformation, along with the two previous steps (Equation (5.3) and Equation (5.5)) can be expressed as,

$$\sigma(N^{-\frac{1}{2}} A N^{-\frac{1}{2}} \cdot (X \cdot W)) \in \mathbb{R}^{M \times D'} \tag{5.7}$$

An activation function called *Rectified Linear Unit (ReLU)* is used in this framework. Two other activation functions, *Sigmoid* and *Softmax*, are also used in the later part of the dimuon classifier. The sigmoid function converts a real-number input into a classification score in the range $(0, 1)$. The softmax function is the generalized form of the sigmoid function and converts a vector of K real numbers ($\overrightarrow{z}$) into score values of K possible classes. The three nonlinear functions are defined as,

$$\begin{aligned} \textbf{ReLU:} \quad & \sigma(z) = \max(0, z) \\ \textbf{Sigmoid:} \quad & \sigma(z) = \frac{e^z}{1 + e^z} \\ \textbf{Softmax:} \quad & \sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}, \quad \text{for} \quad i = 1, ..., K \quad \text{and} \quad \overrightarrow{z} \in \mathbb{R}^K \end{aligned} \tag{5.8}$$

The use of sigmoid and softmax functions is discussed in the later sections of the model description. Figure 5.8 shows the output of the ReLU and Sigmoid functions.

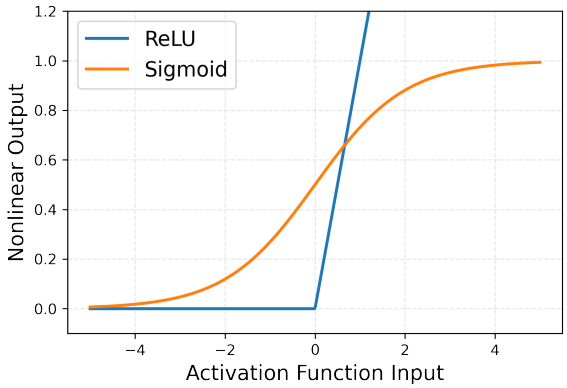


Figure 5.8: The output of the activation functions, ReLU and Sigmoid.

The execution of the above three steps completes a full graph convolution operation (also alternatively called 'layer'). In a GCNN, multiple graph convolution can be applied subsequently on the output of the previous layer to increase the complexity of the model. The input feature matrix for each layer $L+1$ can be expressed in terms of the output of the previous layer L using the following iterative equation,

$$X^{L+1} = \text{ReLU}\left(N^{-\frac{1}{2}} A N^{-\frac{1}{2}} \cdot \left(X^L \cdot W^L\right)\right) \quad . \tag{5.9}$$

At the first iteration, $X^{L=0}$ is the input attribute matrix from all the hit DOMs in an event. All the intermediate layers between the input attribute matrix and the final

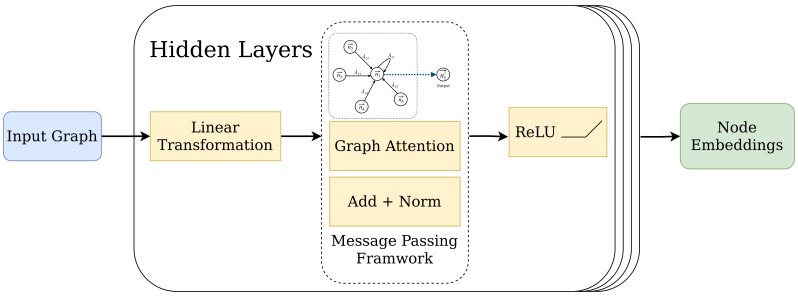


Figure 5.9: Diagram showing the steps of a full GCNN model, which outputs the node embeddings while keeping the same graph structure. The stack of hidden layers indicates the application of multiple graph convolutions. 'Add+Norm' in the message passing step denotes the aggregation operation discussed in the text.

output are called *hidden layers*. Figure 5.9 shows the flow diagram of the intermediate steps of a GCNN used in this work. After multiple graph convolutions, the model outputs the same graph structure with a different feature representation on the nodes, called an *embedding*. After model training, the node embeddings contain a learned representation of the input features that should discriminate between double- and single-track events. The graph (alternatively, event) classification is then achieved by using a process called *pooling operation* which condenses the information of the node embeddings into a single class score indicating the double-trackness of the event.

5.2.3 Hierarchical Graph Pooling

The most simplistic way to perform pooling for the graph classification is to use a global pooling function such as 'Average', 'Sum', and 'Max'. A global pooling function, e.g. 'Average' takes the mean values of each feature (in the embedding space) across all the nodes in the graph. After the application of the pooling method, all node embeddings are reduced to a single vector with a length equal to that of the output vector of each individual node embedding (denoted as D^{out}), which is expressed as,

$$\left[X_G^{pool}\right]_j = \frac{1}{M}\sum_{i=1}^{M} X_{ij}^{out} \quad \text{for} \quad j = 1, ..., D^{out} \quad , \tag{5.10}$$

where X^{out} is the final node embeddings of the graph and $\left[X_G^{pool}\right]_j$ is the j^{th} element of the output vector $\overrightarrow{X}_G^{pool} \in \mathbb{R}^{D^{out}}$. In the dimuon event classification context, most of the DOM hits in a double-track event show very similar patterns to the single-track

events due to the small track separation of the dimuons and the scattering of light in the ice. Therefore, only a few DOMs in dimuon events are expected to contain high-level features constructed from these hits that show distinctive deviations from single-track features. Equation (5.10) shows that the global pooling methods operate on all the nodes in the event graph equally and so ignore the structural information of the graph. As a result, the majority of the contribution in the construction of graph-level representation, i.e. the output of the pooling operation, will come from DOMs that provide little to no discrimination, diminishing the effectiveness of the dimuon classifier. This type of issue is a common graph classification problem in many fields where a faint signal in the data is accompanied by a substantial amount of noise, as discussed in Ref. [190].

Developing an efficient graph pooling method is still a highly active field of research in computer science. This work implements a complex, powerful, and state-of-the-art pooling method called *Differential Pooling (DiffPool)* [191] as an alternative to the global pooling methods. During the training phase, in addition to graph convolution, the DiffPool pooling function learns to map the nodes of the original graph to a smaller number of clusters, forming a reduced graph representation. One application of the DiffPool operation forms a *level* of node reduction by combining subsets of input graph nodes into what is effectively a single node in the smaller output graph. The same node pooling process can be repeated on the newly formed representation to achieve the next level of reduction. Therefore, multiple DiffPool operation hierarchically condenses the original graph embedding into a single node with a final $\mathbb{R}^{D^{out}}$ dimensional pooled vector equivalent to the global pooling methods, i.e. the output of Equation (5.10). An example diagram in Figure 5.10 illustrates three levels of DiffPool operation to achieve the final pooled vector used for graph classification.

The objective of the DiffPool method is to compute a cluster assignment matrix S^l at level l, which performs the pooling operation on the input nodes and transforms them into fewer clusters to make a coarsened graph at level $l+1$. Let us denote a multiple graph convolution operation on an input graph with M nodes (described by the attribute matrix X and adjacency matrix A) as $GCN(A, X)$. The final output can be expressed as,

$$Z = GCN(A, X) \in \mathbb{R}^{M \times D'} \quad , \tag{5.11}$$

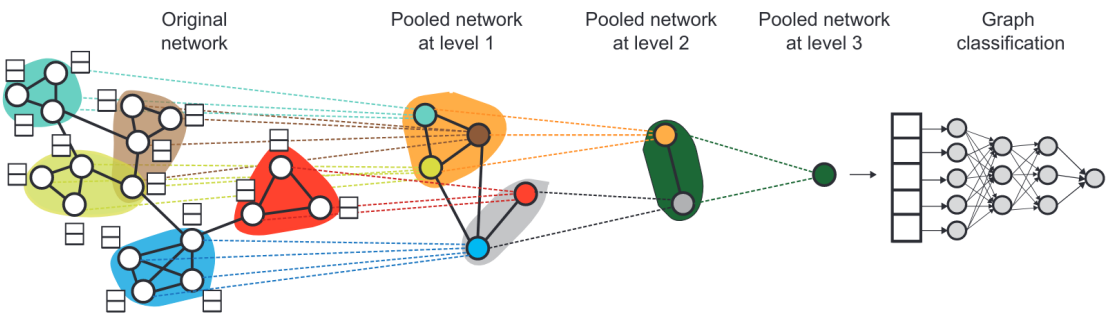


Figure 5.10: Illustration of the DiffPool method from Ref. [191]. At each hierarchical level a GCNN is run to learn the cluster assignment of the input network nodes to output a condensed graph representation. The colours and the dotted lines represent the clustering of the nodes to form a new node for the next hierarchical level.

where D' is the dimension of the output embedding vector of each node. In the DiffPool method for each level l, two separate GCNNs are trained in parallel. One GCNN is trained for learning on the usual embedding space and is called GCN_{emb}. The other GCNN is trained to generate the cluster assignment matrix and is called GCN_{pool}. For any level l, let us denote the number of nodes in the input graph as M^l, the attribute and the adjacency matrix as X^l and A^l, respectively. The input for both GCN_{emb} and GCN_{pool} is the same, but the output is different and serves two different purposes. The output of GCN_{emb}, Z^l_{emb} is the node embedding of the input graph and has the dimension $M^l \times D'$. The output of GCN_{pool}, Z^l_{pool} has the dimension $M^l \times N^l_C$, where N^l_C is the number of output clusters. Therefore, Z^l_{pool} can be used to compute a probabilistic assignment of the M^l input nodes into N^l_C output clusters with $N^l_C < M^l$. The following expression summarizes the two GCNN operations,

$$\begin{aligned} Z^l_{emb} &= GCN^l_{emb}(A^l, X^l) \quad \in \mathbb{R}^{M^l \times D'} \\ Z^l_{pool} &= GCN^l_{pool}(A^l, X^l) \quad \in \mathbb{R}^{M^l \times N^l_C} \quad . \end{aligned} \tag{5.12}$$

The Softmax function is then applied on each row of Z^l_{pool} to convert it into the cluster assignment matrix S^l,

$$S^l = \text{Softmax}(Z^l_{pool}) \in \mathbb{R}^{M^l \times N^l_C} \quad . \tag{5.13}$$

The pooling for the current level can now be performed on the node embeddings of the input graph, Z^l_{emb} using S^l. The DiffPool operation to generate the reduced graph

representation at level $l+1$ is expressed as,

$$\begin{aligned} X^{l+1} &= S^{l^T} \cdot Z^l_{emb} \quad \in \mathbb{R}^{N^l_C \times D'} \\ A^{l+1} &= S^{l^T} \cdot A^l \cdot S^l \quad \in \mathbb{R}^{N^l_C \times N^l_C} \quad . \end{aligned} \tag{5.14}$$

The input graph at level $l+1$, (X^{l+1}, A^{l+1}) can be treated the same way using Equation (5.12), Equation (5.13), and Equation (5.14) to get the input graph for $l+2$. In this iteration, the number of input nodes is $M^{l+1} = N^l_C$, and the number of target clusters for the next level is $N^{l+1}_C < N^l_C$. The number of pooling levels $\{l\}$ and the number of clusters at each level $\{N^l_C\}$ in the model are part of the model configuration parameters and are discussed in Section 5.2.6. The last pooling operation results in a single output node that has the embedding vector with predetermined length D^{out}, denoted as $X^{pool}_G \in \mathbb{R}^{D^{out}}$, similar to the output of Equation (5.10). Since the number of target clusters in the last level, $N^{last}_C = 1$, the last pooling operation using Equation (5.14) is essentially a global 'Sum' operation discussed at the beginning of the section.

5.2.4 Multilayer Perceptron

The graph neural network discussed above operates on the DOM-level features. An MLP component is implemented in the dimuon classifier to utilize a set of event-level features to further distinguish dimuons from single muon events. Let us denote the event-level feature vector with F number of features as $X_F \in \mathbb{R}^F$. Each layer of the MLP network then performs the following operation to propagate the representation from layer L (denoted as X^L_F) to $L+1$ (denoted as X^{L+1}_F),

$$X^{L+1}_F = \text{ReLU}\left(X^{L^T}_F \cdot W^L_F\right)^T \quad , \tag{5.15}$$

where W^L_F is the weight matrix for the layer L. W^L_F has the dimension $F \times F'$, where F' is the feature dimension of the next layer $L+1$. After multiple layers, the output of the MLP is concatenated with the final pooled output of the graph neural network model. If the output vector of the MLP network is denoted as $X^{out}_F \in \mathbb{R}^{F^{out}}$ with F^{out} as the dimension of the last layer, the concatenated vector $X_{G+F} = \text{concat}(X^{pool}_G, X^{out}_F)$ contains the complete event-level representation of both the graph neural network and the MLP module and has a length equal to $D^{out} + F^{out}$.

Finally, the vector X_{G+F} is the input for another set of MLP layers with the same operation shown in Equation (5.15), but with X_{G+F} as the input vector, before producing a final classification output. This part of the model is referred to as *fully connected* as the data processed in this step connects all the previous model components. In the last fully connected layer, the ReLU activation function is replaced by the Sigmoid function to output a class score of the events,

$$\text{Class Score} = \text{Sigmoid}\left(X_{G+F}^{L_{FC}^T} \cdot W_{G+F}^{L_{FC}}\right) \quad , \tag{5.16}$$

where L_{FC} is the last layer in the fully connected unit. The weights $W_{G+F}^{L_{FC}}$ in the last layer is a column vector with the same length as $X_{G+F}^{L_{FC}}$. Therefore, the matrix multiplication in the above equation outputs a single real number used in the Sigmoid function to convert it into a classification score in the range $[0, 1]$, representing the double-trackness of a given input event. A flow diagram of all the components of the hybrid neural network model for the dimuon classifier is shown in Figure 5.11.

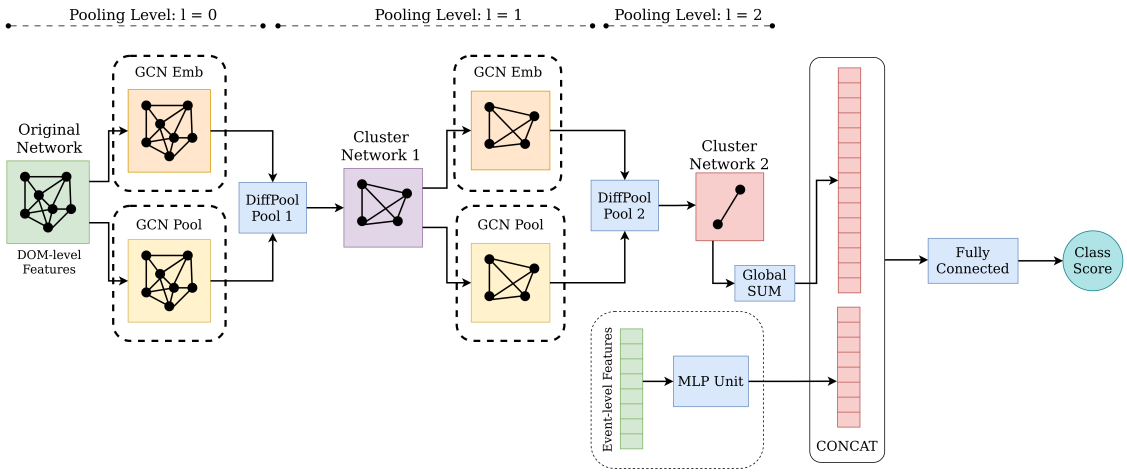


Figure 5.11: Framework of the hybrid neural network for the dimuon classifier. The GCN modules in each dashed box are a full GCNN operation on the respective input graph. Each DiffPool operation generates the graph with a reduced number of nodes (denoted as *cluster network*), which is used as the input graph for the next pooling level. The last pooling operation in the diagram is denoted as the 'Global SUM', equivalent to the DiffPool operation with a single target cluster, as discussed in the text.

5.2.5 Model Input Space

The DOM-level features that form the attribute matrix X, the edge connections between DOMs that form the adjacency matrix A, and the event-level features X_F for the input of the MLP module collectively create the model input space for the dimuon classifier. The high-level properties for these three objects are constructed from the single-track reconstruction MuEx and the noise-cleaned pulse series TTPulses (discussed in Section 4.3) of each event.

Attribute Matrix from DOM-level Features

The construction of the DOM-level features is motivated by a qualitative inspection of the DOM hit patterns from dimuon events. Let us call the plane containing the two outgoing true muon tracks as *Track-plane*, T. The plane perpendicular to T that contains the points of intersection of the Cherenkov cones from the two muons is called the *Intersection-plane* I, as shown in Figure 5.12. The single-track reconstruction (MuEx) of the dimuon events is assumed to return an approximately average direction of the two muons and is illustrated in Figure 5.12. DOM II in the right diagram of

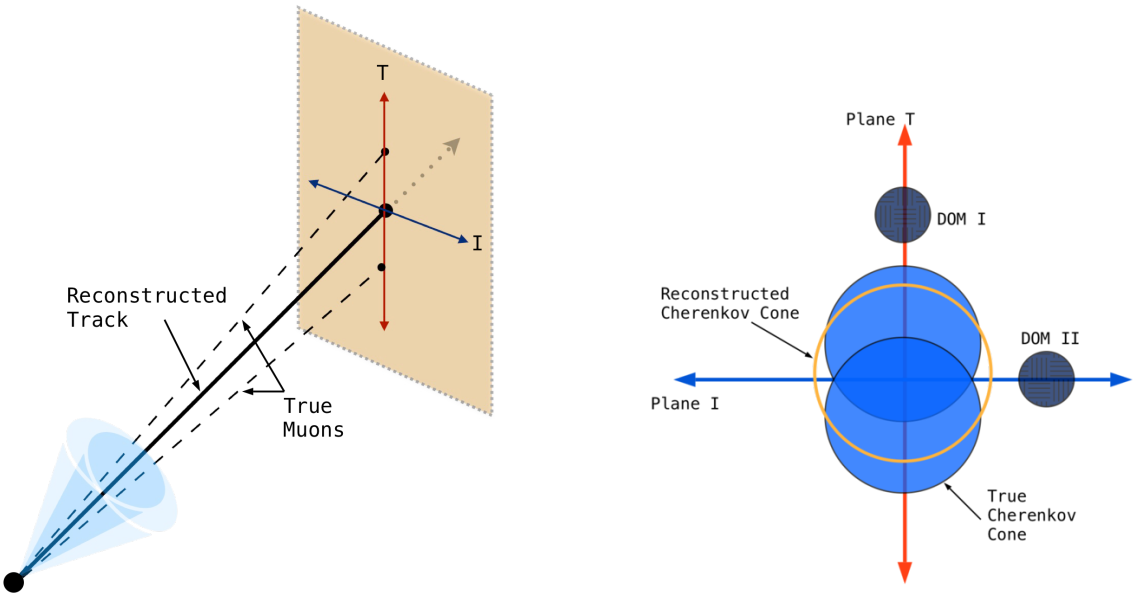


Figure 5.12: Left: Illustration of a dimuon event with a reconstructed track and the visualization of Plane T, I perpendicular to the average track direction. Right: Cross-sectional view of the two Cherenkov cones produced by the dimuon events and expected hit pattern under single-track reconstruction.

the figure expects a delayed hit from the two true muon tracks compared to the single-track assumption for the event. In addition, all the photon hit times in DOM II are expected to form a single peak distribution as the Cherenkov photons from both the muons arrive at the DOM simultaneously. In contrast to DOM II, DOM I in the same figure expects an earlier and a later hit from the true Cherenkov cones compared to the reconstructed Cherenkov cone. If the separation of the two true Cherenkov cones is above the timing resolution of the detector ($\sim$ 3 ns), DOM I is expected to see a double peak signature in its pulse time distribution as the two Cherenkov cones arrive at different times. These expected dimuon signatures rely on the ideal conditions of no scattering and absorption of photons in the ice. In a more realistic scenario with the scattering and absorption of the photons, the dimuon hit signature is expected to be smeared out for most of the DOMs in an event, and one or both tracks may not generate a hit in a nearby DOM. To validate our expectation of dimuon signatures, a toy MC dataset of class A dimuons is produced by injecting the events horizontally into the detector at depths with DeepCore strings. This configuration simulates a best-case scenario of dimuon event signatures in the detector. An MC set of single muons with equivalent energies, i.e. $E_\mu = E_{\mu_1} + E_{\mu_2}$ are also simulated with the same configuration to compare with the dimuon events. The plots in Figure 5.13 show the expected photon arrival time in a string from the reconstructed and true, simulated tracks with an overlay of the first hit times in each DOM. The comparison of the double-track and single-track events in the plots confirms the delayed timing effect of the dimuon signature for DOM II (refers to the DOMs near the tip of the reconstructed Cherenkov light curve in the plots) discussed above. Based on this initial investigation, twelve DOM-level features are constructed for the graph neural network component of the dimuon classifier. The expected Cherenkov photon properties computed from the track reconstruction in some of the features ignore the scattering and absorption effects in the ice. The definitions of the twelve features are as follows:

- **Residual Time (*TRes*) :** The difference between the first hit time in the DOM and the expected hit time from single-track reconstruction,

$$TRes = 1^{st} \text{ Pulse Time} - \text{Expected Cherenkov Photon Hit Time} \quad . \tag{5.17}$$

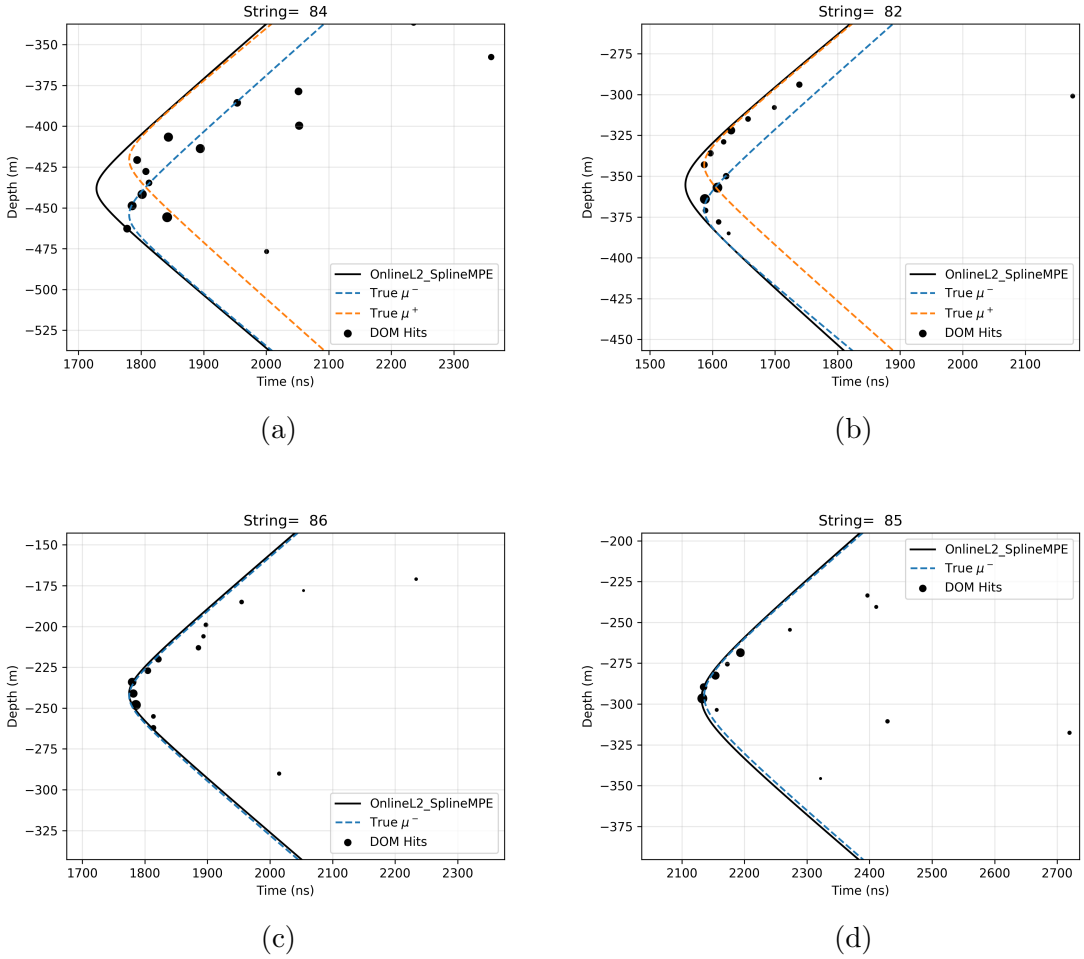


Figure 5.13: The simulated events visualized in individual strings. The first pulse times of each hit DOM in a string and the expected arrival time of light from the tracks are plotted on the x-axis. The z-distance of the strings (along with the DOM positions) in the IceCube coordinate is plotted on the y-axis. The size of the scatter points represents the total charge detected by the DOMs. The top two (a and b) plots are from simulated dimuon events, and the bottom two plots (c and d) are from simulated single muon events. A basic L2 track reconstruction, Online_L2SplineMPE is used in this plot.

- **Cherenkov Distance (*ChDist*) :** The distance travelled by the Cherenkov photon from the point of origin on the reconstructed track (referred to as *Cherenkov position*) to the DOM,

$$ChDist = \|\text{DOM Position} - \text{Cherenkov Position}\| \quad . \tag{5.18}$$

- **Distance-on-track (*TDist*) :** The distance between the Cherenkov position and the 'centre of gravity' (CoG) of the track's total charge deposition. This

feature represents a temporal correlation of the DOMs as the Cherenkov position is directly connected to the propagation of the track in time,

$$TDist = \frac{\text{Cherenkov Position} - \text{Track CoG}}{\text{Track Direction (Unit Vector)}} \quad . \tag{5.19}$$

For a reconstructed track of length L, the value of $TDist$ is approximately in the range $[-L/2, +L/2]$.

- **First Pulse Charge (*QFirst*)** : The charge of the first pulse in the hit DOM.

- **Total Charge (*QTot*)** : The total charge in a given Hit DOM,

$$QTot = \sum_{i=1}^{N} q_i \quad , \tag{5.20}$$

where N is the number of pulses in the DOM and $\{q_i\}$ is the set of pulse charges.

- **Max Pulse Charge (*QMax*)** : The maximum charge among all the pulses in a hit DOM,

$$QMax = \max\{q_i\} \quad i \in \{N\} \quad . \tag{5.21}$$

- **Max Pulse Time (*TMax*)** : The relative time (with respect to the first pulse) of the pulse with the maximum charge in the hit DOM,

$$TMax = \text{time}(\text{argmax}\{q_i\}) - time(1^{st}\text{Pulse}) \quad \text{for} \quad i \in \{N\} \quad . \tag{5.22}$$

- **Standard Deviation of Time Differences (*TDiffSig*)**: The standard deviation of the time differences between the consecutive pulses in the hit DOM. This is a measure of the temporal spread of the pulses in a DOM and is defined using the following two expressions,

$$TDiffMean = \frac{\sum_{i=1}^{N-1} \text{time}(i+1) - \text{time}(i)}{N-1} \quad , \tag{5.23}$$

$$TDiffSig = \sqrt{\frac{\sum_{i=1}^{N-1} [\text{time}(i+1) - \text{time}(i) - TDiffMean]^2}{N-1}} \quad .$$

For $N = 1$, the feature value is set to 0.

- **Mean Residual Time (*TMeanRes*)** : The residual time constructed at the beginning only considers the first hit in a DOM. This feature computes a charge-weighted average of the residual times from all the hits in the DOM,

$$TMeanRes = \frac{\sum_{i=1}^{N} TRes(i) \cdot q_i}{N \cdot QTot} \quad , \tag{5.24}$$

  where $TRes(i)$ is the residual time of the i^{th} pulse.

- **Number of Pulse Hits (*NHits*)**: The total number of pulses in a hit DOM.

- **Hit Time Window (*HitWindow*)** : The difference between the first and last hit times in the DOM pulses,

$$HitWindow = \text{time}(i_{last}) - \text{time}(i_{first}) \quad . \tag{5.25}$$

- **Mean Charge per Hit (*QMean*)** : The average charge per hit is calculated using the total charge in a DOM and dividing it by the number of hits ($NHits$),

$$QMean = \frac{QTot}{NHits} \quad . \tag{5.26}$$

These DOM-level features form the $D = 12$ dimensional input feature vector for each node and are used to construct the attribute matrix in the event graph. A complete correlation matrix of the DOM input features is constructed and discussed in Appendix A.

Adjacency Matrix from DOM Connections

Figure 5.12 shows that the double-track signatures in DOM I and II are complementary to each other, and their positions are in the perpendicular planes (i.e. plane T and plane I, respectively). Therefore, the hit DOMs, which lie in the same plane along the track and have their Cherenkov positions close to each other, are expected to see similar features described in the previous section. Based on this assumption, an edge connection is constructed as the 2-D Gaussian kernel, which is a function of the relative angle between two DOMs with respect to the reconstructed track and the distance between the Cherenkov positions of the same two DOMs, as shown in

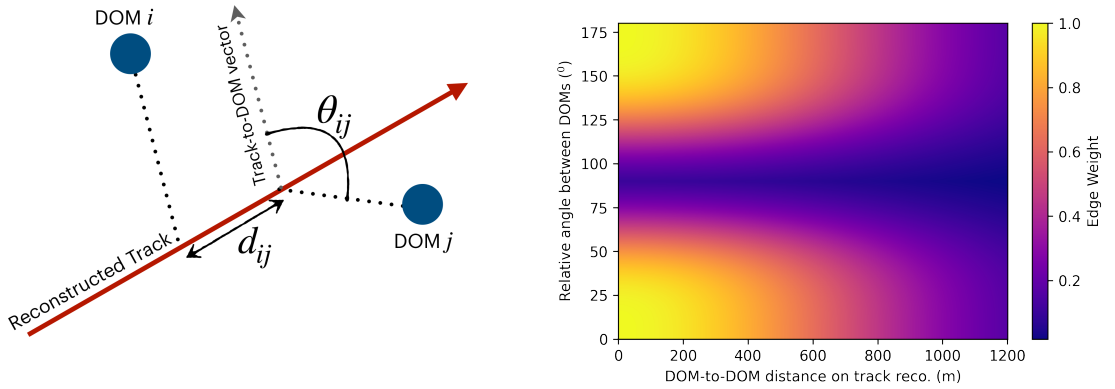


Figure 5.14: (Left) An illustration of the two distance parameters (d_{ij} and θ_{ij}) for a pair of DOMs. (Right) The edge weight distribution as a function of DOM-to-DOM distance on track (d_{ij}) and the relative angle between two DOMs on the track (θ_{ij}).

Figure 5.14. The output of the kernel is the edge weight e_{ij} between DOM i and j and is expressed as,

$$e_{ij} = N \cdot \exp\left[-\left(\frac{d_{ij}^2}{\sigma_d^2} + \frac{(1-|\cos\theta_{ij}|)^2}{\sigma_\theta^2}\right)\right] \quad , \tag{5.27}$$

where d_{ij} is the distance between the Cherenkov positions and θ_{ij} is the angle between the track-to-DOM vectors of the two DOMs. N is the normalization factor computed during the normalization of the full adjacency matrix A, shown in Equation (5.5). The standard deviations, σ_d (for distance along the track) and σ_θ (for relative angle) dictate the locality of the neighbourhood strength for the DOMs. They are considered to be the learnable parameters of the model during training, i.e. part of the model parameter space $\boldsymbol{\theta}$ discussed in Section 5.2.1. The optimization of the edge weights by tuning the parameters σ_d and σ_θ refers to the graph attention mechanism introduced in Section 5.2.2. The plot in Figure 5.14 shows an example 2-D kernel describing the edge weights as a function of d_{ij} and θ_{ij}. The edge weights $\{e_{ij}\}$ generated using Equation (5.27) are directly used to form the adjacency matrix A, i.e. $A_{ij} = A_{ji} = e_{ij}$.

Event Features

The six event-level features used as the input vector of the MLP unit is defined below.

- **Event Charge (*QEvent*) :** The sum of the pulse charges from all the hit

DOMs in the event,

$$QEvent = \sum_j \sum_i \{q_i\}_j \quad , \tag{5.28}$$

where $\{q_i\}$ is the set of pulses in a DOM and the index j runs on all the hit DOMs for the event.

- **Event Tracklength (*LEvent*) :** The total track length of the event calculated using the Cherenkov position of the DOMs. This feature is computed using a subset of the hit DOMs called *direct-hit DOMs*. These DOMs are considered to be close enough to the reconstructed track for detecting unscattered photons from the muon. This property is implemented in the platinum selection method and, thus, is reused here.

- **Initial Track Intensity (*ITInt*) :** The light intensity of the first 100 m of the track segment. If L_i is the subset of the hit DOMs with their Cherenkov position in the first 100 m of the track segment, the intensity parameter is expressed as,

$$ITInt = \frac{\sum_{i \in L_i} Q_i \cdot d_i \cdot \exp(d_i / l_{attn})}{N_{phasespace}} \quad , \tag{5.29}$$

where Q_i and d_i are the total charge and Cherenkov distance of the DOMs, respectively. $l_{attn} = 100$ m is an attenuation length parameter, and $N_{phasespace}$ is the number of DOMs in the entire detector (including no-hit DOMs) with the projected Cherenkov position inside the track segment. To avoid any undefined value for this feature (e.g. events with $N_{phasespace} = 0$), a minimum track length requirement is imposed, as discussed in Section 5.2.7.

- **Final Track Intensity (*FTInt*) :** The light intensity parameter for the final 100 m of the track segment. This feature uses the same calculation shown in Equation (5.29) but accounts for the final 100 m segment of the reconstructed track.

- **Track Smoothness (*TSmooth*) :** This feature comes from a commonly used track property within the IceCube analysis software and measures the smoothness of the light distribution along the track. It is defined as the maximum

relative deviation of the pulses from a uniform distribution along the track [192],

$$TSmooth = \max\left(\left|\frac{i}{n} - \frac{l_i}{l_n}\right|\right) \quad \text{for} \quad i = 2, 3, ..., n-1 \quad , \tag{5.30}$$

where i is the index of the hit DOMs sorted in distance along the track with total n hit DOMs. l_i is the length between the Cherenkov positions of the first hit DOM and the i^{th} hit DOM in the sorted distance list. The definition of this feature requires at least three hit DOMs in an event and is also included in the event selection criteria discussed in Section 5.2.7.

- **Direct Smoothness (*DSmooth*)** : The same track smoothness parameter, but computed using the subset of direct-hit DOMs in the event.

These six features for each event form the $F = 6$ dimensional input feature vector for the MLP unit of the dimuon classifier. The correlation among the event-level features is also discussed in Appendix A.

Feature Scaling

Both the DOM- and event-level features used in this classification problem come from different physics properties with different units. Directly using these features in the attribute matrix and the input vector results in a wide range of required values for the model parameters associated with the corresponding features, affecting the loss function optimization process. The issue is avoided by performing a *feature scaling* process on DOM- and event-level features. The scaling of each feature f can be expressed in terms of its mean μ_f and standard deviation σ_f computed on the entire dataset,

$$f_{scaled} = \frac{f - \mu_f}{\sigma_f} \quad . \tag{5.31}$$

After the feature scaling, all the feature distributions are centred approximately around zero with a range of $\sim \pm 10$. This process preserves the shape and performs only scaling and translation of the distributions such that all input features of the model have a common scale.

5.2.6 Model Architecture

The dimuon classifier in this work is implemented using the `PyTorch Geometric (PyG)` library [193], which is developed on the platform of the `PyTorch` software framework [194]. Several configurations of the model are investigated during the model development and are part of the model hyperparameters, which are discussed in detail in Section 5.2.9. The final configuration of all the components of the dimuon classifier is discussed below.

Graph Neural Network for DOM Features

The discussion of the graph neural network in Section 5.2.3 indicates the computation of multiple GCNN operations (referred to as GCNN modules for each GCN_{emb} and GCN_{pool} operation) in a combination of parallel and series for the model. Therefore, a stack of GCNN modules is used to build the complete graph neural network. The `DenseGCNConv` operator [195] in `PyG` is used to perform the graph convolution operation on each layer of the GCNN modules. The following terminology is used to provide the configuration details of each GCNN module:

- *num_nodes* : Number of nodes in the input graph
- *in_channels* : Dimension of the input node features
- *out_channels* : Dimension of the output node embedding
- *hidden_channels* : Dimension of the node output in each hidden layer
- *num_hidden_layer* : Number of hidden layers

In addition, the configuration related to the DiffPool method requires the following details,

- *num_clusters* : Number of output clusters at a pooling level
- *num_levels* : Number of hierarchical DiffPool pooling levels

The model implemented in this work has two hierarchical pooling levels (*num_levels*). For level 0 and 1, the numbers of target output clusters for the smaller graph (*num_clusters*) are set to 8 and 4, respectively. The details of all the GCNN components

	Pooling Level: $l = 0$	Pooling Level: $l = 1$
Node Embedding	$\mathbf{GCN^0_{emb}}$	$\mathbf{GCN^1_{emb}}$
	num_nodes : M	*num_nodes* : 8
	in_channels : 12	*in_channels* : 32
	hidden_channels : 32	*hidden_channels* : 32
	out_channels : 32	*out_channels* : 32
	num_hidden_layer : 2	*num_hidden_layer* : 2
	Output Dimension : $M \times 32$	**Output Dimension** : 8×32
Node Clustering	$\mathbf{GCN^0_{pool}}$	$\mathbf{GCN^1_{pool}}$
	num_nodes : M	*num_nodes* : 8
	in_channels : 12	*in_channels* : 32
	hidden_channels : 32	*hidden_channels* : 32
	out_channels : 8	*out_channels* : 4
	num_hidden_layer : 2	*num_hidden_layer* : 2
	Output Dimension : $M \times 8$	**Output Dimension** : 8×4

Table 5.2: Architecture details of the GCNN modules used in the graph neural network. At level 0, M is the number of hit DOMs and varies for each event. Also, *in_channels* at level 0 is the number of the DOM-level input features. The node dimension, *hidden_channels* is the same for all the hidden layers in a GCNN component. The values shown in this model configuration are optimized via hyperparameter tuning, discussed in Section 5.2.9.

in the complete graph neural network architecture are provided in Table 5.2 with a notation of each corresponding module as $\mathbf{GCN^{<l>}_{emb/pool}}$ (as introduced in Equation (5.12)). After each level of two parallel GCNN (GCN_{emb} and GCN_{pool}) operations, the `dense_diff_pool` operator [196] is used to transform the input graph into a reduced graph representation for the next level, as shown in Equation (5.14). At the end of two subsequent DiffPool operations (two levels), the model outputs a graph with four nodes and applies the `SUM` operation (global pooling, equivalent to DiffPool operation with single node output) to get the final embedding vector with fixed dimension, 1×32.

MLP Architecture
in_channels : 6
hidden_channels : 16
out_channels : 8
num_hidden_layer : 2
Output Dimension : 1×8

Table 5.3: Details of the MLP module configuration in the dimuon classifier.

Given a large number of parameters ($\mathcal{O}(10^3)$) in the model, it can often suffer from overfitting of the data. Two regularization methods, called *Batch Normalization* and *Dropout*, are used to prevent the issue of overfitting. The Batch Normalization technique normalizes the values of the processing layer at the step before applying the activation function using two free (learnable) parameters, *running mean* and *running variance*. This method helps stabilize the model training, as discussed in Ref. [197]. In the Dropout technique [198], a certain fraction of the connections between the layers (weight connections as illustrated in Figure 5.6) is randomly ignored in each iteration of the parameter optimization process to prevent the overfitting of the weights. The `BatchNorm` and `Dropout` methods in `PyG` are used to impose these regularization steps between the hidden layers of the model.

MLP Module for Event Features

The architecture details of the MLP component are summarized in Table 5.3. Without the loss of generality, the same terminology defined for a node in the graph is reused to describe the MLP configuration in the table. `BatchNorm` and `Dropout` are applied between the hidden layers of the network, similar to the graph neural network component.

The output of the graph neural network (32-component vector) and the MLP network (8-component vector) are concatenated to form the 40-component input vector for the fully connected neural network. In this final network of the model, one hidden layer is added with 16 *hidden_channels*, followed by the final output layer with one output channel (*out_channels*) that contains the classification score.

5.2.7 Model Training

Training Datasets

For the model to identify class A dimuon events (as discussed in Section 5.1), the model training requires class A events as the signal and single-track events as the background. This is achieved by using an MC dataset called *DatasetA* that contains only class A dimuons as class 1 events and single muons from ν_μ CC DIS as class 0 events. In addition, a second model with identical architecture is trained to identify class B dimuons with *DatasetB*, which contains class B dimuons as the double-track events. The two models trained with DatasetA and DatasetB are referred to as *GNetA* and *GNetB*, respectively. The class scores from both models are then used to select a final signal-rich region with minimum background contamination for performing the search analysis.

As discussed in Section 5.1, the class A dimuons are only a small fraction ($\sim 3\%$) of all dimuon events resulting in low statistics in the MC DatasetA. However, the model training requires a large number of events ($\sim 5 \times 10^4$) from both classes (dimuon and single muon) due to the high complexity and large number of parameters in the model. Since it is too computationally expensive to produce enough class A events via the event generation of all dimuons (class A, B, and C), a geometry resampling technique is developed in this work to address the low-statistics issue. The technique moves the already generated class B and C dimuon events (output of `Charm` and `Trident Dimuon Generator`) around the detector volume and converts them into high-quality class A double tracks. The resampling method is described in detail in Appendix B and is used to generate the training dataset, DatasetA, with a higher number of class A dimuons compared to the original simulation.

Pulse and Event Selection Cuts

Although the platinum selection process removes the majority of the low-quality track events and noise in the pulse series, the sample still contains inessential events and pulses in the context of dimuon classification. Therefore, a set of pulse and event selection cuts (for both training and final analysis) is applied before processing the events for the classification models' inputs. These cuts are referred to as *preprocessing*

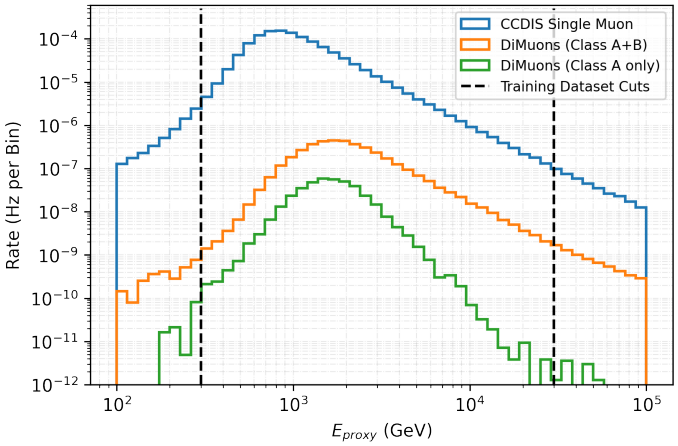


Figure 5.15: Expected event rates as a function of the reconstructed energy for different classes of tracks. The vertical lines show the energy cuts imposed before performing the model training.

filters. A minimum charge threshold (q_{min}) and a residual time window ($TRes$) for each of the pulses are implemented to reduce the noise hits (mainly from afterpulses and dark noise) while constructing the DOM-level features for the event graph. An additional cut on the Cherenkov distance ($ChDist$) is included to select the hit DOMs in the GNetB model. This is implemented to limit the graph size of high energy ($> \mathcal{O}(10$ TeV)) events in DatasetB. Two event-level cuts on the minimum number of hit DOMs ($NHitDOMs$) and track length are imposed to avoid undefined feature values for the events (discussed in Section 5.2.5). The dimuon events are simulated with E_ν in the range 100 GeV $-$ 1 PeV. However, Figure 5.15 shows that class A dimuons are expected to be mainly in the reconstructed energy (E_{proxy}) range 300 GeV $-$ 30 TeV, resulting in an energy cut for events in both DatasetA and DatasetB. Later, in the final search analysis, an updated reconstructed energy cut will replace the current reconstructed energy criterion for the training datasets, which is discussed in Section 6.3.1. The summary of the pulse- and event-level cut details at this step is shown in Table 5.4. After applying the above cuts, the distributions of the DOM-level and event-level scaled features are shown in Appendix A.2.

5.2.8 Treatment of Imbalanced Training Dataset

Ideally, the training dataset should contain an equal number of events in both event classes (single- and double-tracks). However, the expected number of background

Pulse Selection Cuts	**Event Selection Cuts**
$q_{min} \geq 0.25$ PE	NHitDOMs ≥ 3
$TRes \in [-500 \text{ ns}, 1500 \text{ ns}]$	Track Length ≥ 125 m
$ChDist \leq 500$ m	$E_{proxy} \in [300 \text{ GeV} - 30 \text{ TeV}]$

Table 5.4: The preprocessing filters containing the pulse and event selection cuts before processing the events for the classification model input. The cut criteria on the Cherenkov distance $ChDist$ is imposed only for the model GNetB. All the preprocessing filters except the reconstructed energy (E_{proxy}) are used for the final analysis dataset.

events is two and three orders of magnitude higher than the class B and class A signal events, respectively. Therefore, the model training process runs into the issue of highly imbalanced classes in both DatasetA and DatasetB. In addition, the MC datasets for different physics processes are produced with different generation spectra (shown in Table 4.1), leading to a mismatch of the raw energy distributions of the events between single-track and double-track classes. Mismatch of the underlying energy distributions in the training dataset can affect the model training as the input features such as *QTot*, *QMax*, and *QEvent* are highly correlated to the track's energy. As a result, the trained model can be susceptible to unwanted machine bias. There are widely used class and sample weighting methods applicable to simple ML methods such as boosted decision trees and MLPs. These methods use the entire imbalanced dataset during training but weight the loss output of each event according to the class size (*class weights*) or sample distribution (*sample weights*). However, such weighting techniques do not work well with graph neural networks due to their complicated architecture. The development of efficient techniques in handling the class imbalance issue in graph neural networks is still an active field of research, as discussed in Ref. [199, 200]. Another brute force method of handling class imbalance is either undersampling or oversampling of one of the classes to get a balanced dataset. This work randomly undersamples single muon background events from the full training dataset for the training of the dimuon classifiers. After the sampling process, the number of single muons in the sampled subset is equal to the number of dimuons for the training and has a matching reconstructed energy distribution,

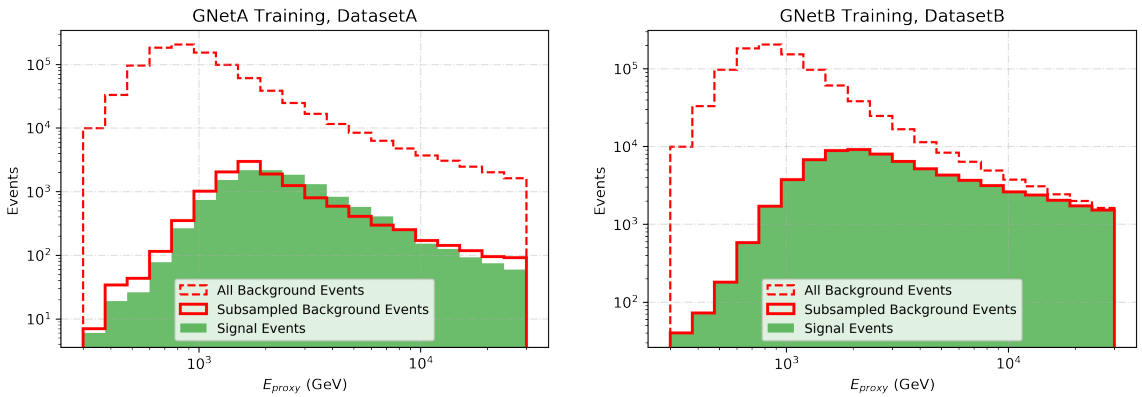


Figure 5.16: Raw event distributions, i.e. events without any event weights of the reconstructed energy for both signal and background classes in DatasetA (left) and DatasetB (right). The solid red line shows an example random subsample set of the background with a number of events equal to the signal events.

as shown in Figure 5.16. While the undersampling method solves the class imbalance and sample energy distribution issues, it results in lower statistics in the final training sample. This limits the improvement of the parameter optimization and is addressed by generating multiple random subsample sets of the background events and using them in successive iterations during model training. The detailed steps of the technique are as follows:

1. Generate a random background subsample (denoted as S_1) with a distribution shown in Figure 5.16 and set it as the background events for the training dataset.

2. Perform the model training with N_e number of epochs, where one epoch is defined as one complete cycle of running the parameter optimization process, i.e. the training or learning process of the algorithm on the entire training dataset.

3. At the end of N_e epochs, pause the current training, generate a new random S_2 subsample of the background events, and update the training dataset by replacing S_1 with S_2.

4. Repeat steps 2 and 3 for subsequent subsample sets to continue the training process for upto subsampling S_N.

Each subsampling S_i (with $i = 1, 2, ..., N$) continues the training process for N_e epochs

individually, but globally the model training progresses for a total $N \times N_e$ number of epochs. A total of $N = 10$ and 5 random subsample sets are used to train GNetA and GNetB, respectively.

5.2.9 Hyperparameter Tuning and Training Settings

The model's hyperparameters include the choice of *out_channels*, *hidden_channels*, *num_hidden_layer*, *num_levels*, and *num_clusters*. Exploring the hyperparameter space to optimize the model performance for a complex model like the dimuon classifier is computationally expensive as each configuration of the settings requires retraining of the model from scratch. Therefore, only a limited set of hyperparameters are investigated to optimize the model performance. *out_channels* and *hidden_channels* for both GCN_{pool} and GCN_{emb} are varied from the set $\{8, 16, 32, 64\}$ at all DiffPool levels $0, 1$. For the MLP module, *out_channels* and *hidden_channels* are varied from the set $\{5, 8, 12, 16\}$. *num_levels* is varied using the values from the set $\{1, 2, 3\}$, and *num_clusters* is varied from the set $\{16, 8, 4, 2\}$. For all GCN_{emb}, GCN_{pool}, and the MLP modules, *num_hidden_layer* is run from the set $\{1, 2, 3\}$. From this tuning process, the choice of the hyperparameters for the best-performing model is used in the final model configuration, as discussed in Section 5.2.6. The regularization steps, i.e. Batch Normalization and Dropout, also have a few associated hyperparameters. For Batch Normalization, a hyperparameter called *momentum* is used in the optimization process of the running mean and variance [201]. A typical value of 0.1 for the momentum is used in this work. The Dropout rate is varied from the set $\{0.1, 0.2, 0.3, 0.4\}$, and the final value of 0.3 is used for both GNetA and GNetB. Therefore, a random 30% of the connections between the hidden layers are dropped during each iteration of the training process. Both GNetA and GNetB are trained for 50 epochs for each subsample set. Both models use a set of ∼150K events (sum of single- and double-tracks) in the training dataset.

The model parameters are optimized using *gradient descent* [202] and backpropagation [203] algorithms during the model training. The `PyTorch` software framework constructs a map of all the parameter values and their interdependent gradients. An optimization software then operates on the gradient map to search for the best parameter values of the model by minimizing the loss function shown in Equation (5.2).

The `ADAMAX` software [204] is used for this purpose and requires the assignment of a parameter called *learning rate* which refers to the step size for scanning the parameter hyperplane. The optimizer starts with an initial learning rate of 0.01 and varies the rate during optimization over multiple epochs using a method called *ReduceLROnPlateau* [205]. In this mechanism, the progress of the optimization process is monitored by checking the model performance (based on the evaluation metrics defined in the next section) on a dataset separate from the training dataset, called the *validation set* at the end of each epoch. Every time the model performance plateaus over a few epochs (a decision made by the algorithm), the learning rate is reduced by a factor of 2 to enable a finer scanning of the parameter space until it reaches a lower limit of 10^{-6}.

5.2.10 Model Performance and Classification Results

Several model evaluation metrics are defined to quantify the model's performance during training. A test of the model's ability to identify each class (single- and double-track) can be constructed by defining a classification score threshold. If there are total P number of class 1 (double-track) events, and the model correctly identifies TP number of double-tracks, i.e. the signal events above the score threshold, the metric called *true positive rate* is defined as,

$$\text{True Positive Rate} = \frac{TP}{P} \quad . \tag{5.32}$$

Similarly, if there are total N number of class 0 (single-track) events, and the model misidentifies FP number of single-tracks as double-track events, the metric *false positive rate* is defined as,

$$\text{False Positive Rate} = \frac{FP}{N} \quad . \tag{5.33}$$

A curve, called *receiver operating characteristic (ROC)* curve is constructed by plotting the false positive rate in the x-axis and the true positive rate in the y-axis for varying class score thresholds from 1 to 0. A model evaluation metric, *area under the curve (AUC)*, is then computed from the area under the ROC curve and ranges between 0 and 1, with 1 indicating the perfect classification of all the events. In

addition, the *accuracy* of the model is also used to monitor the training performance and is defined using a class score threshold of 0.5 as,

$$\text{Accuracy} = \frac{TP + TN}{P + N} \quad , \tag{5.34}$$

where TN is the number of correctly classified single-track events. The model evaluation metrics - AUC, accuracy, and loss as a function of the epochs during the model training are shown in Figure 5.17. The spikes and dips in the metrics indicate the transition from one subsample set to the next subsample set. The plots in the figure also show that the performance of GNetA and GNetB saturates after 6 and 3 subsample sets, respectively.

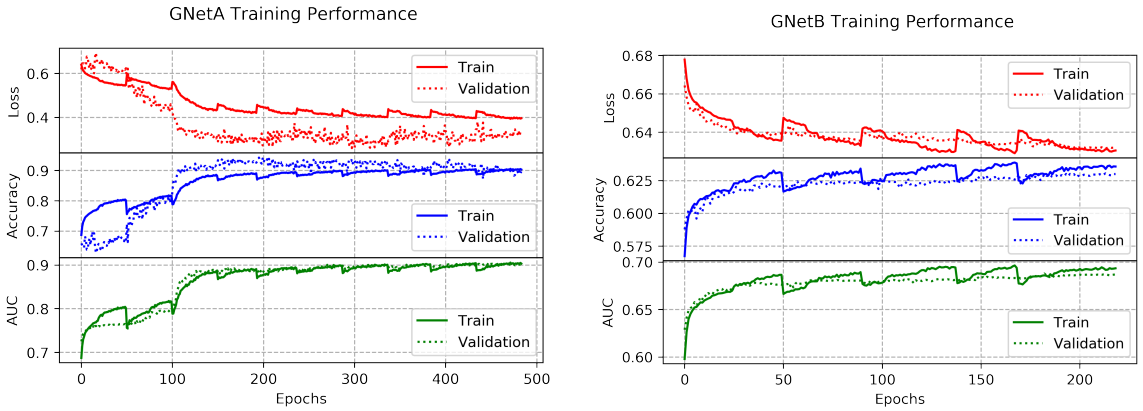


Figure 5.17: Training performance measured in loss (output of the binary cross-entropy loss function), accuracy and AUC score as a function of the epochs for GNetA (left) and GNetB (right).

Using the final trained model for both GNetA and GNetB, the normalized class score distributions of the training and validation sets for signal-track and double-track events are shown in Figure 5.18. In addition, the plots report a reduced chi-square test result (χ^2_η) calculated by comparing the model's performance on the training and validation set. χ^2_η values of 0.95 and 1.18 for GNetA and GNetB, respectively, indicate that both models have no significant overfitting or underfitting during the training. The comparison of the distributions from GNetA and GNetB in Figure 5.18 also shows that GNetA performs significantly better in correctly classifying the class A dimuons from the single muon backgrounds than the classification of class B dimuons in GNetB. This is an expected outcome, as the class A dimuons are postulated to be

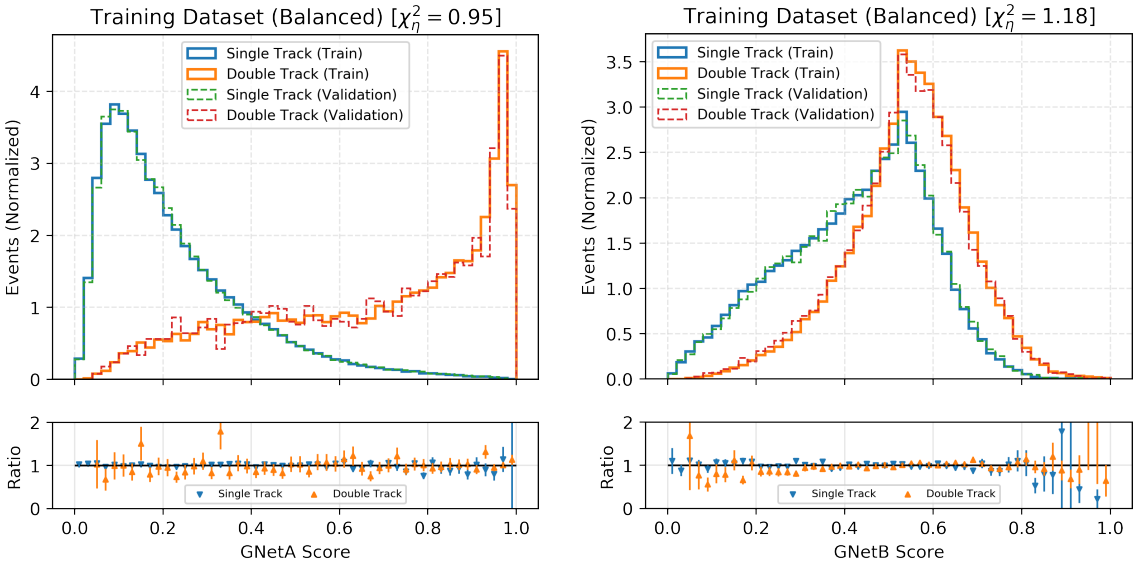


Figure 5.18: Class score distribution of GNetA (left) and GNetB (right) is shown for the training and validation set. The error bars in the training/validation ratio (shown at the bottom) are due to the statistical uncertainties of each bin.

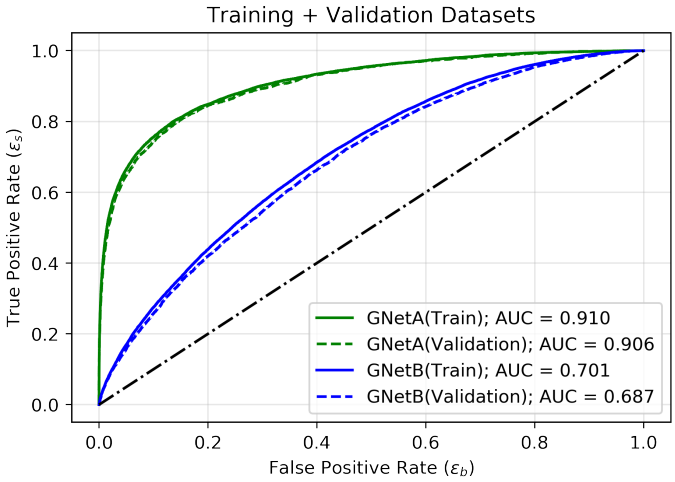


Figure 5.19: AUC curves for GNetA and GNetB.

readily identifiable events in the development of the dimuon classifiers. The plot in Figure 5.19 shows the ROC curves for both models. The plot also reports the AUC score of GNetA and GNetB to be 0.91 and 0.69, respectively.

Finally, the event rates of individual classes of dimuons as a function of the class score for both models and dimuon channels are shown in Figure 5.20. The distributions of the GNetA score indicate that the peak of the class A dimuon distributions

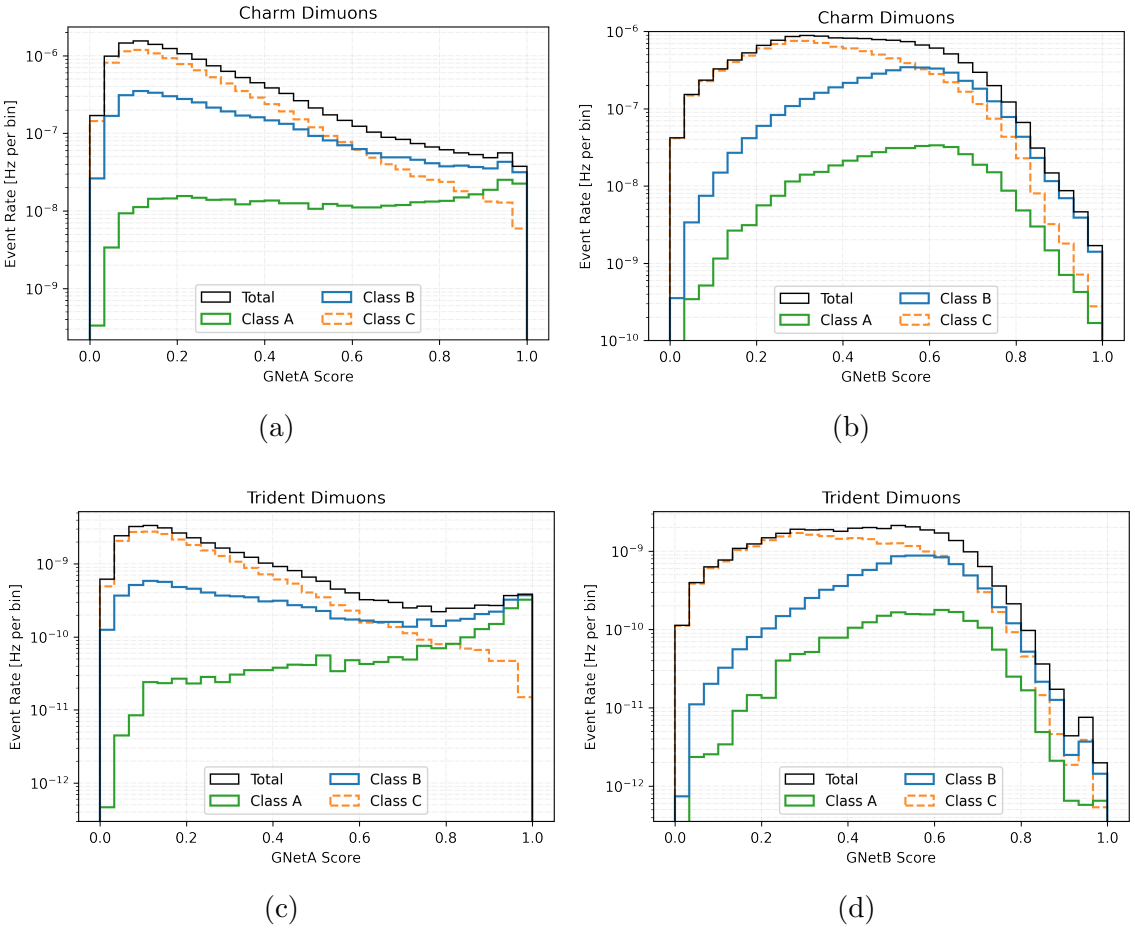


Figure 5.20: The expected event rate of the individual classes of charm dimuon events in the platinum sample for both charm (a and b) and trident (c and d) dimuon productions as a function of the GNetA score (left) and GNetB score (right). The total event rates from all three classes of dimuons are also shown in solid black distributions.

has high double-trackness, i.e. class score ~ 1. The distributions of class B and C dimuons in the same plot have peaks near the low class score, with a steeply falling event rate at the higher scores for class C dimuons which are expected to be aligned with single-track like event signature. The GNetB score distributions in Figure 5.20 show that most class A and B dimuons are predicted to have high GNetB scores, with class C dimuons following the expectation profile similar to the one observed in GNetA. The 2-D correlation of the two score event rate distributions for both the double-track and single-track events are shown in Figure 5.21. The next step in the analysis chain is to define a region on the 2-D score space that offers an expected high sensitivity for extracting the signal events. The plots in Figure 5.21 show that

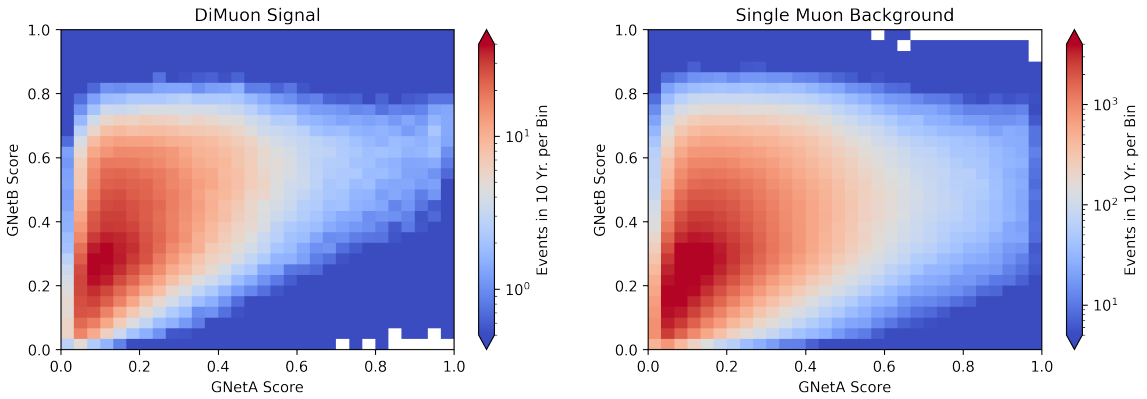


Figure 5.21: Expected event rate distributions in the 2-D score space for signal (left) and background (right) before the final analysis cuts.

the two model scores are not highly correlated. This indicates that a selection cut based on both model scores can achieve a better background rejection than cutting on any individual model score. The construction of such a signal-rich region is part of the final analysis method discussed in the next chapter.

Chapter 6

Analysis Method and Results

The classification scores for each event from the two dimuon classifiers (GNetA and GNetB) form the final observable space of the search analysis. This chapter provides details of the analysis cuts to define the signal regions and the description of the statistical hypothesis test to quantify the identification of the observed events in the signal regions. Finally, the results from the measurement of the dimuon event identification are reported.

6.1 Data Selection

The IceCube data used in this analysis spans from May 13, 2011 to June 07, 2022 with a total livetime of 10.67 years. The dataset's start time refers to the beginning of the detector operation with the complete 86 strings configuration. The data acquisition of the detector typically operates for $\sim 8-$hour segments (referred to as *runs*), and the collection of all the runs for approximately a year comprises the data of the *run season* for the corresponding starting year. Every run is monitored for data quality and is marked as 'good run' if the detector performance during the run is within normal operating conditions, i.e. no major hardware or software failures. The set of good runs, thus, excludes any data that can be attributed to runs with data acquisition issues and detector calibration and forms the primary dataset for all physics analyses in IceCube. In addition, the good runs are required to have all 86 active strings and at least 5000 active in-ice DOMs to prepare the final dataset for this analysis. These additional criteria further enhance the data quality and maintain the uniformity of the active detector geometry throughout the total duration of the dataset. The events

in the final dataset are then processed through the same platinum selection process as the MC dataset (discussed in Section 4.3). Table 6.1 summarizes the individual and total events in the platinum sample of the IceCube data for the $2011 - 2021$ run seasons. In addition, the event rates of the platinum sample for each season are

Run Season	**Number of Events**	**Livetime (s)**	**Live-Fraction (%)**
2011	39988	28777782.50	90.56
2012	39432	28037343.06	87.53
2013	41774	29844500.64	90.19
2014	43369	31293410.01	89.98
2015	43924	31325562.97	92.38
2016	42804	30611916.66	97.87
2017	49344	35161189.04	97.44
2018	44180	31617466.26	93.11
2019	37107	26491018.04	96.91
2020	43446	31145444.57	93.78
2021	45027	32147088.81	93.08
Total	470395	336452722.55	96.32

Table 6.1: The number of events in the platinum sample along with the livetime and live fraction for each season runs. The livetime refers to the total duration of the selected runs for the season after applying the additional criteria discussed in the text. The live-fraction is the livetime as a fraction (%) of the full season duration. The last row in the table also shows the total number of observed events, total livetime, and overall/average live-fraction in the platinum sample for the full duration of the dataset.

plotted in Figure 6.1 and show no significant deviation from the average data rate of the total duration.

Before further processing of the data, a complete analysis framework is developed based on the MC events by following the blind-analysis method to avoid confirmation bias. The framework consists of treating the systematic uncertainties, applying the final analysis cuts, and developing the hypothesis tests discussed in the following

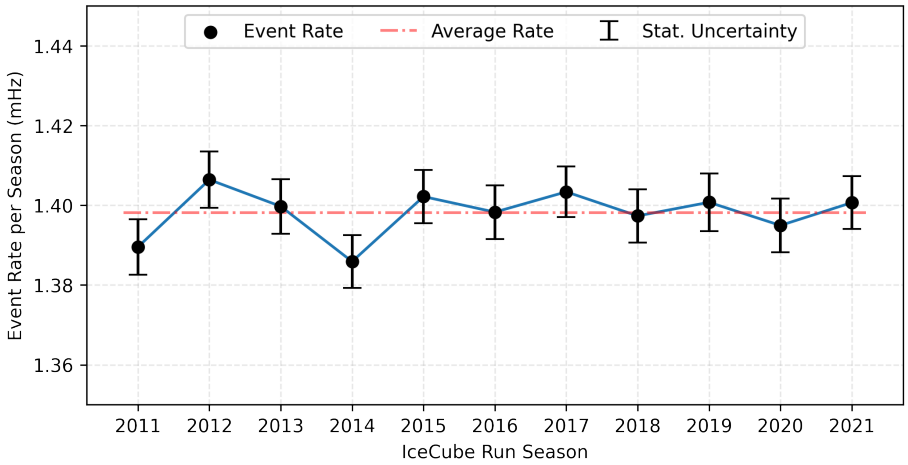


Figure 6.1: The event rate of the platinum selection in IceCube data for each run season. The error bars indicate the statistical uncertainty.

sections.

6.2 Treatment of Systematic Uncertainties

The uncertainties related to the detector modelling, neutrino fluxes, and the interaction cross section can lead to the variation of the expected events in the MC simulation. Therefore, a set of systematic effects on the simulated events is considered in the analysis to account for the uncertainty in the measured values of the model parameters and is discussed below. For each systematic effect, a pair of parameter values (upper and lower bounds) around the central (nominal) value of the model is considered to be the $\pm 1\sigma$ uncertainty based on the prior knowledge from past IceCube analyses [134].

6.2.1 Detector Systematics

Understanding the effects of the detector systematics requires re-simulating the Cherenkov photons for each value of the model parameters. The DOM efficiency parameter (denoted as DOM Eff.) describes the overall scaling factor of the modelled photon detection efficiency of the entire detector as discussed in Section 3.2.3. The nominal MC datasets (discussed in Chapter 4) are produced with a central DOM eff. value of 0.97. Two separate MC datasets are produced with DOM Eff. values of 0.93

and 1.03 to account for the lower and upper variation of the light detection efficiency, respectively. The simulation of $\nu_\mu + \bar{\nu}_\mu$ CC DIS interactions producing both the single muons (primary background) and charm dimuons (primary signal) are considered for these systematic sets.

The uncertainties in modelling the bulk ice are implemented by varying the scattering and absorption coefficients. The nominal MC datasets are produced by simulating the photon propagation using SPICE3.2, as discussed in Section 4.2.2. Four systematic datasets for the primary signal MC are generated by re-simulating the photon propagation using a combination of $\pm 5\%$ scattering and $\pm 5\%$ absorption variation around the SPICE3.2 parameters. A single bulk ice systematic set for the primary background was produced in the MEOWS analysis using a technique called *Multisim* and is reused in this work. The Multisim method performs a Fourier decomposition of the scattering and absorption coefficients of each ice layer. It then generates a variation of the ice model from the perturbation of the amplitudes and phases of the Fourier series, as discussed in detail in Ref. [206].

The uncertainty related to the hole ice primarily comes from the different scattering properties (compared to bulk ice) of the refrozen ice columns that contain the DOMs [207]. The scattering of the photons in the bulk ice occurs over a large propagation length and is implemented by explicitly simulating the photon propagation from the light source to the DOM positions (as discussed in Section 4.2.2). However, the scattering due to the accumulated air bubbles in the hole ice is a local phenomenon around the DOMs and primarily affects the detection of the photons (that reach the DOM positions) based on their incoming direction. Therefore, the hole ice systematic effect is modelled by constructing a probabilistic estimation of the simulated photon's detection in the DOMs as a function of incoming photon direction. This is referred to as DOM angular acceptance and has the following parametric form [134],

$$A(\eta) = 0.34 \left(1 + 1.5\cos(\eta) - \frac{\cos^3(\eta)}{2}\right) + p_1 \cos(\eta) \left(\cos^2(\eta) - 1\right)^3 + p_2 \exp[10\,(\cos(\eta) - 1.2)] \quad , \tag{6.1}$$

where η is the angle of the incoming photon with respect to the DOM, i.e. $\eta = 0$ corresponds to photons travelling vertically upwards towards the face of the PMT, and p_1 and p_2 are the free parameters to account for the systematic variation. As

discussed in Ref. [134], the parameter p_1 is found to have a negligible impact on the angular acceptance of the photons and a fixed value of 0.30 is used for all the MC datasets. $p_2 = -1.0$ is used for the simulation of nominal MC events. $p_2 = +2.0$ and -5.0 are used to describe the $\pm 1\sigma$ variation of the hole ice in producing the systematic datasets. Similar to the DOM Eff. and bulk ice sets, the primary signal and background datasets are considered for the hole ice variation. The plot in Figure 6.2 shows the angular acceptance ($A(\eta)$) models as a function of $\cos(\eta)$ used for the nominal and systematic datasets.

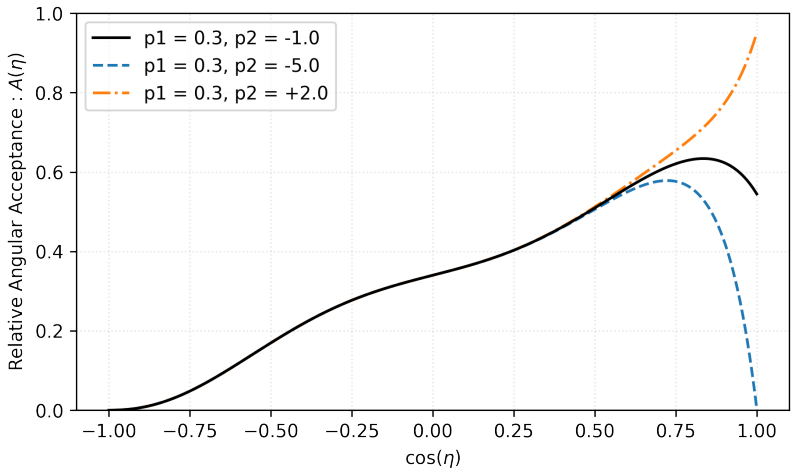


Figure 6.2: The variation in the angular acceptance of different hole ice models used to produce the nominal (solid black line) and systematic (dashed and dash-dotted lines) MC datasets.

6.2.2 Neutrino Flux and Cross Section Systematics

In addition to the detector systematics, the uncertainties in the neutrino interaction cross section and the flux models are also considered in the analysis. Since the modelling of the cross sections and the fluxes are only involved in the calculation of the MC event weights, the related systematic effects can be implemented on the nominal MC datasets with a variation of the event weights. The $\pm 1\sigma$ variation of the total cross section for $\nu_\mu + \bar{\nu}_\mu$ CC DIS interaction is taken to be $\pm 10\%$ of the nominal cross section model (CSMS) reported in Ref. [84].

The atmospheric neutrino flux has two components, as discussed in Section 3.4.1. The nominal dataset uses the H3a cosmic ray model and SIBYLL2.3c hadronic in-

teraction model to build the conventional flux [141, 142] and the BERSS model [144] as the prompt flux. Considering the conventional atmospheric flux to be a power law spectrum (as shown in Equation (3.5)), the systematic uncertainties are implemented using ±30% of the nominal flux normalization (denoted as $\Delta\Phi_{conv}$) and ±0.01 around the nominal spectral index (denoted as $\Delta\gamma_{conv}$) as the $\pm1\sigma$ variations reported in Ref. [208, 209]. The similar flux uncertainties (i.e. normalization and spectral index denoted as $\Delta\Phi_{astro}$ and $\Delta\gamma_{astro}$, respectively) from the astrophysical neutrino flux measurement in Ref. [146] provide the corresponding $\pm1\sigma$ variations. The measurement reports the uncertainties in $\Delta\Phi_{astro}$ and $\Delta\gamma_{astro}$ to be ±18% and ±0.09, respectively. Table 6.2 summarizes the treatment of all the systematic effects discussed above.

Systematic Parameters	**MC Datasets**	
	Nominal	Systematic Variation
DOM Eff.	0.97	0.93, 1.03
Bulk Ice	SPICE3.2	±5% Scatt., ±5% Abs. (signal) Multisim (background)
Hole Ice	$p_1 = 0.30, p_2 = -1.0$	$p_1 = 0.30, p_2 = -5.0, +2.0$
CC DIS Cross Section	CSMS	±10%
$\Delta\Phi_{conv}$	H3a+SIBYLL2.3c	±30%
$\Delta\gamma_{conv}$	H3a+SIBYLL2.3c	±0.01
$\Delta\Phi_{astro}$	Astro. $\nu_\mu + \bar{\nu}_\mu$ Fit	±18%
$\Delta\gamma_{astro}$	Astro. $\nu_\mu + \bar{\nu}_\mu$ Fit	±0.09

Table 6.2: A summary of all the systematic effects considered for the analysis. The effects are evaluated for the primary signal (charm dimuon production) and primary background (single muon from ν_μ CC DIS) simulation. The variation of the total event rate distributions due to the systematic effects in the sub-dominant signal and background processes is considered negligible and, thus, is not implemented in this work.

6.2.3 Systematic Variations of the Classification Scores

Each systematic parameter value reported in the previous section produces a separate representation of the MC simulation. All the systematic datasets are then processed through the same event selection and classification processes as the nominal MC. A comparison of the event rate distributions of the class scores (GNetA and GNetB scores) between the nominal and systematic MC can be computed by defining the systematic variation as,

$$[\text{Systematic Variation}]_i^p = \frac{S_i^p - N_i}{N_i} \times 100\% \quad , \tag{6.2}$$

where the index p refers to each parameter of the systematic set and the index i corresponds to each bin in the class score distribution. S_i^p and N_i are the expected number of events in bin i for the systematic and nominal MC, respectively. The plots in Figure 6.3 show the variations for all the systematic effects in the GNetA and GNetB score distributions for both primary signal and background processes. The plots show that the most significant variation in the expected event rates comes from the conventional atmospheric flux normalization ($\Delta\Phi_{conv}$) with an estimated $\sim 30\%$ variation. The detector systematics (DOM Eff., hole ice, and bulk ice) vary the total detected light for each simulated event, resulting in a maximum variation of $\sim 20\%$ of the class score distributions, with the largest contributions coming from the DOM Eff. and bulk ice effects. An increase in the detector systematics variation in the high GNetB scores (between 0.8 and 1.0) is due to the low MC statistics in those bins i.e. a large statistical uncertainty in those bins contaminates the systematic variation calculation.

The variation due to the detector systematics in Figure 6.3 shows the change in the event rate due to the impact of reconstruction, selection, and classification of the events. A robustness test is developed to assess the performance of only the classification models under the detector's systematic uncertainties. The test computes the variation in the AUC score of the final classification models between the nominal and systematic MC, which represents the change in the model's ability to correctly classify the events under the corresponding systematic uncertainty. The AUC variation

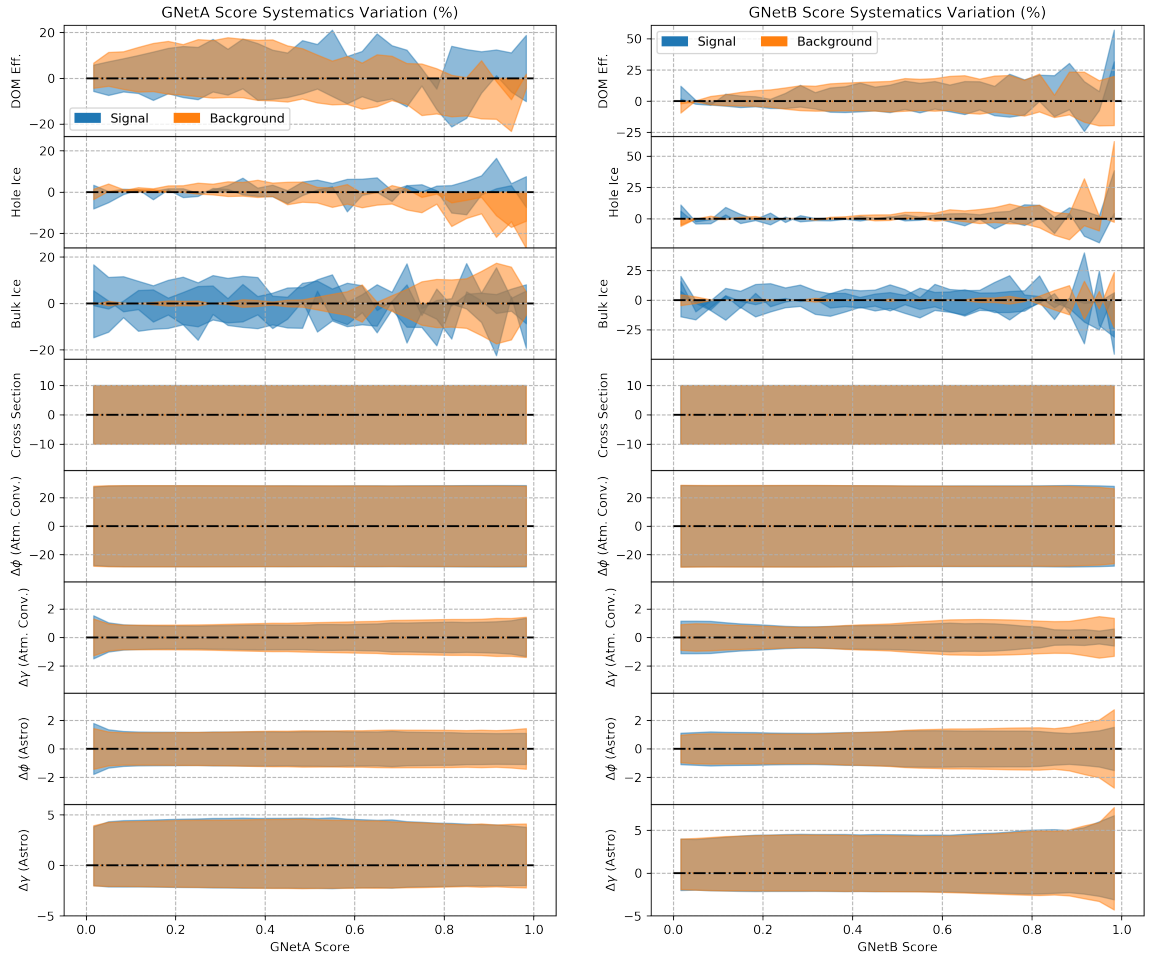


Figure 6.3: Variation in score distributions for GNetA (left) and GNetB (right) due to each systematic effects. The largest uncertainty contribution comes from the conventional atmospheric flux normalization ($\Delta\Phi_{conv}$) with $\sim \pm 30\%$.

is defined as,

$$\text{AUC Variation} = \left(\frac{AUC_{nom.} - AUC_{sys.}}{AUC_{nom.}} \right) \times 100\% \quad , \tag{6.3}$$

where $AUC_{sys.}$ and $AUC_{nom.}$ are the AUC score computed on each detector systematic MC and the nominal MC, respectively. Figure 6.4 shows that the largest variation of the classification performance ($\sim$ 2.2% for GNetA scores) comes from the bulk ice uncertainty, and the rest of the effects contribute $\lesssim$ 1% variation in the model performance for both GNetA and GNetB.

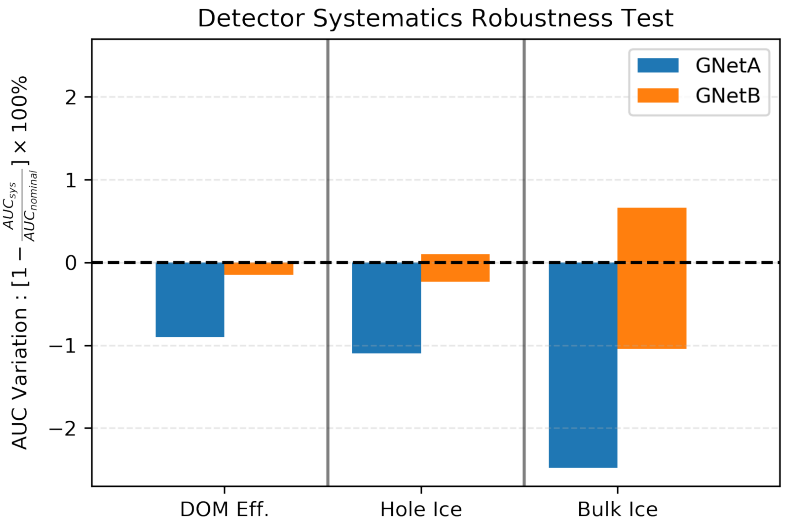


Figure 6.4: AUC variation due to the detector systematics.

6.3 Analysis Cuts

A set of selection criteria (referred to as analysis cuts) based on the four observable (reconstructed energy, zenith, GNetA, and GNetB scores) is implemented to define the final analysis regions used to search for dimuon events.

6.3.1 Energy and Zenith Cuts

The selection of the up-going track events is implemented by requiring the reconstructed $\cos(zenith) < 0$ and reduces the contamination from atmospheric muons. The remaining selection criteria for the analysis cuts of the signal region are developed based on a preliminary class scores observable subspace with the highest class A dimuon content, referred to as the region of interest (ROI). Figure 6.5 shows the implementation of individual cuts on the GNetA and GNetB scores to define the ROI. A loose threshold cut of 0.8 for the GNetA score is chosen on the normalized event rates to capture the peak of class A dimuon events at high scores. A threshold score of 0.4 for GNetB is set where the normalized distribution of class A dimuons and single muon background event rates are equal. The reconstructed energy distribution and the signal-to-background ratio of the signal (all dimuons) and background events in ROI are shown in Figure 6.6 and are investigated to construct the energy cuts. The low energy threshold cut (E_{min}) of 1.1 TeV is imposed, where the signal-

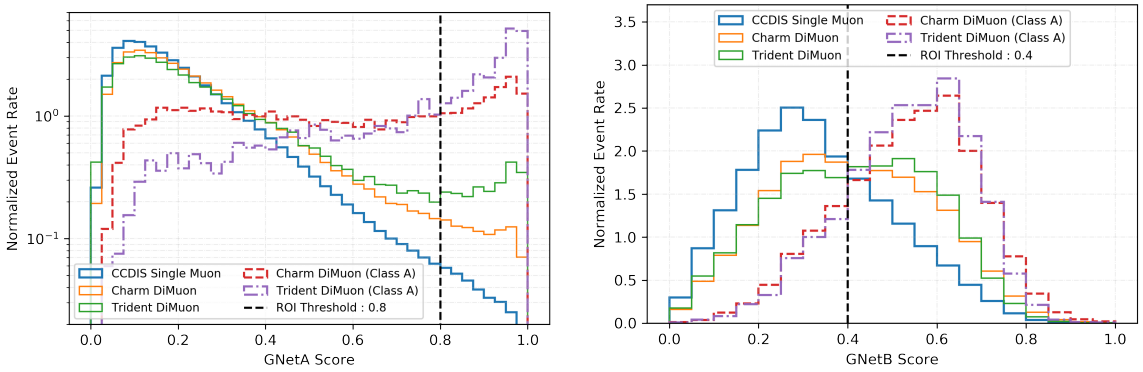


Figure 6.5: Normalized event rate distributions for background and signal (total and class A dimuons) events for GNetA (left) and GNetB (right). The vertical dashed lines are the score thresholds for ROI.

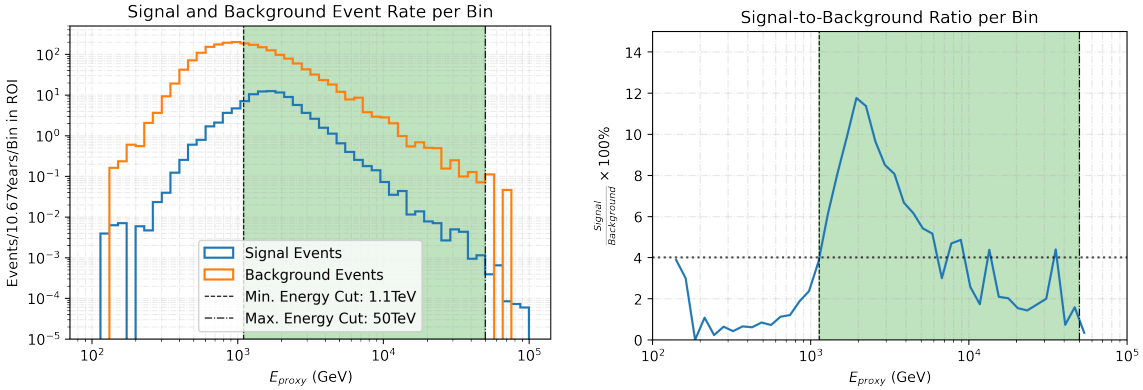


Figure 6.6: (Left) Reconstructed energy distribution of signal and background events in ROI. (Right) Signal-to-background ratio, computed from the left plot. The lower energy cut is set at 1.1 TeV, where the ratio increases above $\sim$ 4%, and the green shaded area is showing the selected region for the energy cuts.

to-background ratio in ROI is 4%, as shown in Figure 6.6. The maximum energy cut (E_{max}) is set at 50 TeV, where the expected number of signal events in 10.67 years becomes vanishingly small (< 0.001 events in 10.67 years). The energy cuts in ROI reduce approximately half of the background events with a small ($\sim$ 15%) sacrifice of the signal events.

The final reconstructed energy and zenith cut developed based on the events in ROI is then extrapolated to the full platinum sample to reduce the background events before analyzing the classification outputs for defining the final analysis regions. The pulse and event selection cuts developed for the model training (preprocessing filters, summarized in Table 5.4), together with the reconstructed energy and direction cuts

discussed above, form the set of *pre-classification analysis cuts*. The expected numbers of events from all signal and background processes in 10.67 years after applying the pre-classification analysis cuts are shown in Table 6.3. The event rate distributions after these cuts as a function of the four observable (reconstructed energy, cosine zenith, GNetA and GNetB scores) for all MC processes are also shown in Figure 6.7.

Physics Processes	**Expected Events in 10.67 Years $\pm 1\sigma$ Stat. Unc.**			
	Total	**Class A**	**Class B**	**Class C**
Charm Dimuon	2,637 $\pm$51	120 (4.6%)	1,231 (46.7%)	1,406 (53.3%)
Trident Dimuon	7 ± 3	1 (14.3%)	3 (42.9%)	4 (57.1%)
$\nu_\mu + \bar{\nu}_\mu$ CC DIS	134,124 $\pm$366	—	—	—
$\nu_\tau + \bar{\nu}_\tau$ CC DIS	69 ± 8	—	—	—
$\nu_e + \bar{\nu}_e$ CC DIS	1 ± 1	—	—	—
Atmospheric Muon	29 ± 5	—	—	—

Table 6.3: Expected number of events from each physics processes after applying the pre-classification analysis cuts. The expectations from individual classes of dimuon events are also shown for signal MC.

6.3.2 Classification Score Cuts

The 2-D score distributions of the expected signal and background events in ROI in 10.67 years are shown in Figure 6.8. The comparison of the distributions between the total signal and background MC shows a region with high GNetA (> 0.91) and GNetB (> 0.6) scores, where a peak of signal events (primarily from class A dimuons) is located near the falling distribution from the background. The signal region must be defined near the peak of the signal events on these plots with minimum background contamination to achieve high sensitivity for the search of dimuon events. Hyperbolic cuts (denoted as function H) with a centre at the highest class scores (i.e. GNetA Score = GNetB Score = 1) are implemented to search for such regions, and the

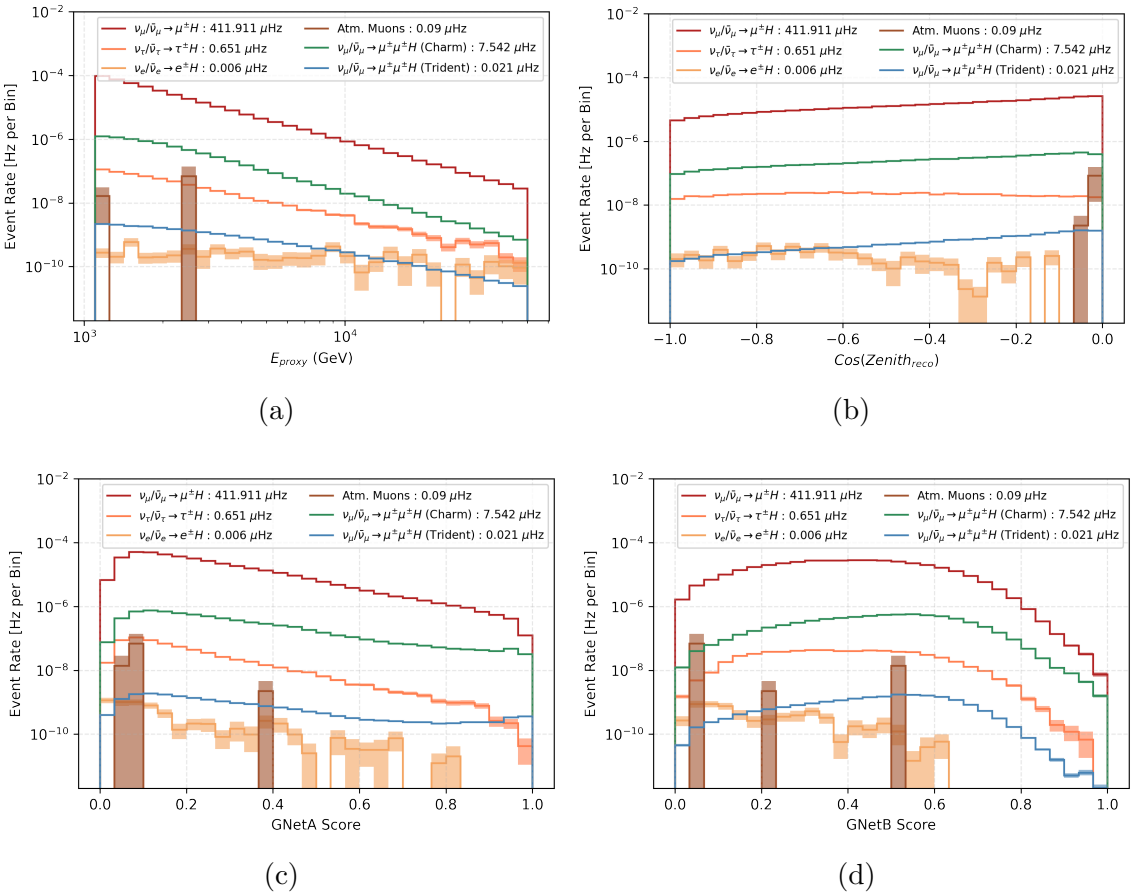


Figure 6.7: Event rate distributions of the final observable (reconstructed energy (a), cosine zenith (b), GNetA (c), and GNet) for each physics process after applying the pre-classification analysis cuts.

parametric form of the cut curve is expressed as,

$$H(S^A, S^B, \alpha, t_{cut}) = \frac{1}{(1.67 + 3.33\alpha)} \big[5\alpha \cdot S^A + 1.67(1-\alpha) \cdot S^B - 20 \cdot (S^A - 1)(S^B - 1)\big] - t_{cut} \ , \tag{6.4}$$

where S^A and S^B are the GNetA and GNetB scores, respectively. α and t_{cut} are the free parameters and dictate the shape and position of the hyperbolic curves. For a given α and t_{cut}, the values of S^A and S^B for which the function H outputs zero form the contour of the corresponding hyperbolic curve. Example hyperbolic curves for different values of α and t_{cut} are shown in Figure 6.9. An optimization process to search for the signal region is implemented by scanning a range of t_{cut} and α values. For each hyperbolic curve defined by an instance of t_{cut} and α, the potential signal region is defined as the region with the class scores S^A, S^B for which the function

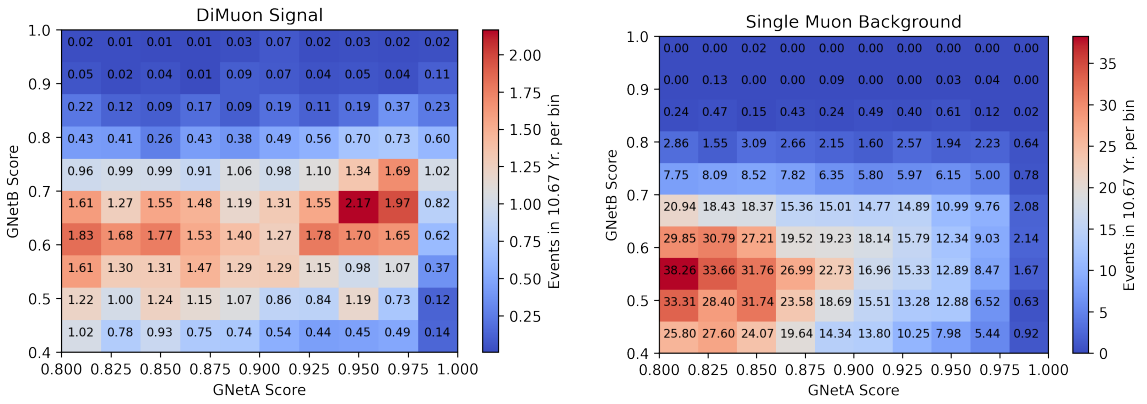


Figure 6.8: Expected event distributions in ROI for signal (left) and background (right). The expected number of events in each bin is also shown in the plots.

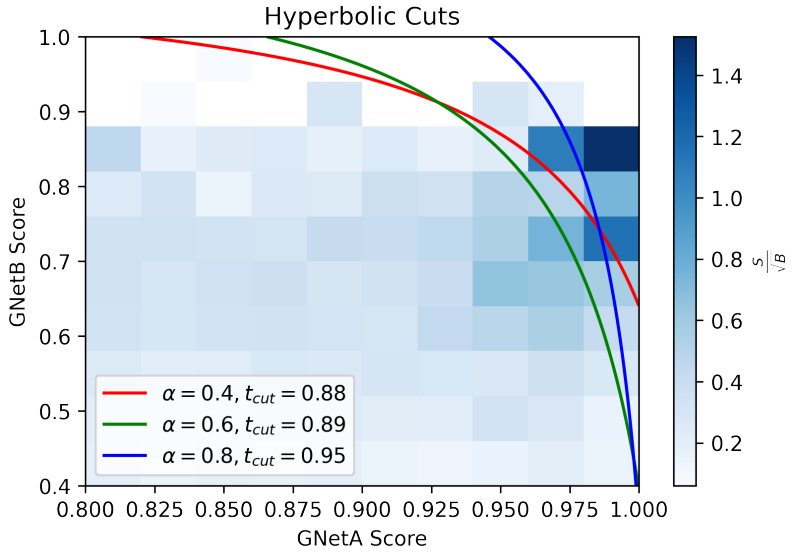


Figure 6.9: Signal/$\sqrt{\text{Background}}$ for each bin on the 2-D score space with an overlay of example hyperbolic cuts for different shape factor α and t_{cut} values.

H in Equation (6.4) outputs ≥ 0, i.e. the region right to the hyperbola in the 2-D score space. A preliminary sensitivity for the presence of signal events and the contamination from the background for each scanned signal region is then defined as,

$$\text{Preliminary Sensitivity} = \frac{S}{\sqrt{B + \sigma_B^2}} \quad , \tag{6.5}$$

where S and B are the expected number of signal and background events in 10.67 years. A constant approximate estimation of the systematic background uncertainty, σ_B, is taken to be $\pm 40\%$ for this optimization process. The preliminary sensitivity scan for different signal regions is performed by varying the hyperbola parameters t_{cut}

and α in the range $[0.7, 1.0]$ and $[0.1, 1.0]$, respectively, as shown in Figure 6.10. The figure also shows the computed sensitivities as a function of the expected number of signal events in the corresponding scanned signal regions, indicating a trade-off between achieving high sensitivity and high statistics on the expected signal events. Two signal regions, SR1 and SR2 are constructed based on the preliminary sensitiv-

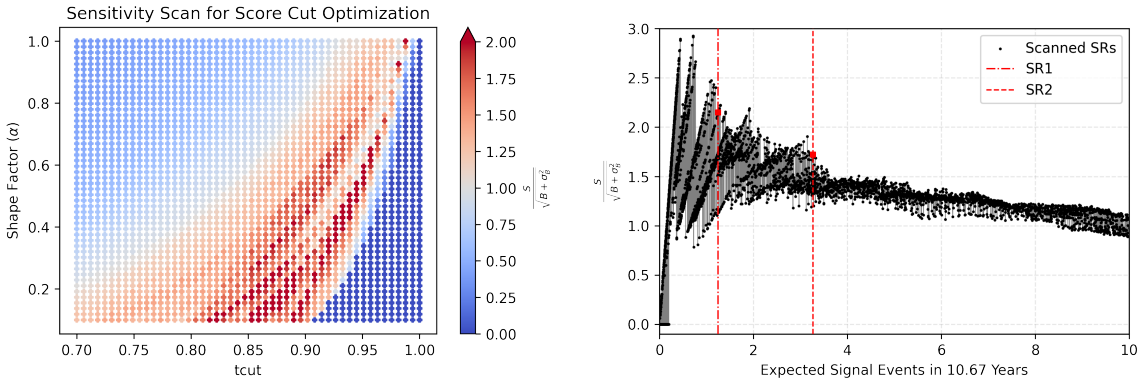


Figure 6.10: (Left) The preliminary sensitivity scan for the hyperbolic cuts defined by the varying shape factor α and contour value t_{cut}. (Right) The computed sensitivities for each scanned signal region as a function of the expected number of signal events. The two red vertical lines and square points show the expected number of signal events and sensitivity of the final signal regions (SR1 and SR2) constructed from the scan. The higher sensitivity points immediately left to SR1 have large statistical fluctuations for small changes in the cut parameters and, thus, are not considered as candidate signal regions.

ities, the expected number of signal events, and the fluctuation of the sensitivities due to slight changes in the parameter values of the scanned signal regions. The hyperbolic cut parameters (t_{cut}, α) for SR1 and SR2 are found to be $(0.861, 0.216)$ and $(0.882, 0.463)$, respectively. The expected numbers of signal and background events in SR1 are 1.25 and 0.21 in 10.67 years, respectively. For SR2, 3.27 signal events and 2.13 background events are expected in 10.67 years.

In addition to the signal regions, two other types of regions, control and validation region, need to be constructed for the search analysis. The control region is used to predict the background contamination in the signal regions using a data-driven method in the analysis and must be defined in a region dominated by the background events. The validation regions validate the background model by comparing the MC expectation and observed data in regions with negligible signal events. The purpose

Analysis Regions	Cut Parameters	Nominal MC Expectation (10.67 Years) Signal Events	Background Events
SR1	$(0.861, 0.216)$	1.25 ± 1.12 (85.6%)	0.21 ± 0.46 (14.4%)
SR2	$(0.882, 0.463)$	3.27 ± 1.81 (60.6%)	2.13 ± 1.46 (39.4%)
CR	$(0.10, 0.10, 0.50, 0.30)$	261 ± 16 (1%)	21343 ± 146 (99%)
VR1	$(0.30, 0.35, 0.50, 0.45)$	102 ± 10 (2%)	5400 ± 73 (98%)
VR2	$(0.55, 0.20, 0.65, 0.40)$	15 ± 4 (2%)	830 ± 29 (98%)
VR3	$(0.45, 0.50, 0.70, 0.60)$	87 ± 9 (3%)	2768 ± 53 (97%)
VR4	$(0.67, 0.20, 0.77, 0.60)$	39 ± 6 (3%)	1271 ± 36 (97%)

Table 6.4: Summary of all the analysis regions showing their definitions in terms of the cut parameter values and the expected number of signal and background events. The cut parameters for SR1 and SR2 refer to the hyperbola parameters (t_{cut}, α). The box cut parameters $(S^A_{min}, S^B_{min}, S^A_{max}, S^B_{max})$ are shown for the control and validation regions. The expected events are also reported in fractions (in parenthesis) to show high signal content in the signal regions and low signal content in the control and validation regions. The errors denote the statistical uncertainties only.

and use of the control and validation regions for searching dimuons in the signal regions are discussed in detail in the next section. Rectangular regions (referred to as box cuts) described by the four boundary parameters $(S^A_{min}, S^B_{min}, S^A_{max}, S^B_{max})$ of the GNetA (S^A) and GNetB (S^B) scores are constructed to define both the CR and VRs. The box cut describing the control region (labelled as CR) is constructed in a low class scores region where the dominating contribution comes from the single muon backgrounds and is defined by the parameters $(0.1, 0.1, 0.5, 0.3)$. Four arbitrary validation regions (labelled as VR1, VR2, VR3, and VR4) are constructed in the 2-D score space between CR and ROI. The box cut parameters related to the validation regions, along with the definitions of CR, SR1, and SR2, are summarized in Table 6.4. The plots in Figure 6.11 also show all the regions developed for analyzing the 2-D class score distributions for both the signal and background MC.

The construction of the analysis regions completes the event selection method for the dimuon search analysis. The expected number of signal and background events

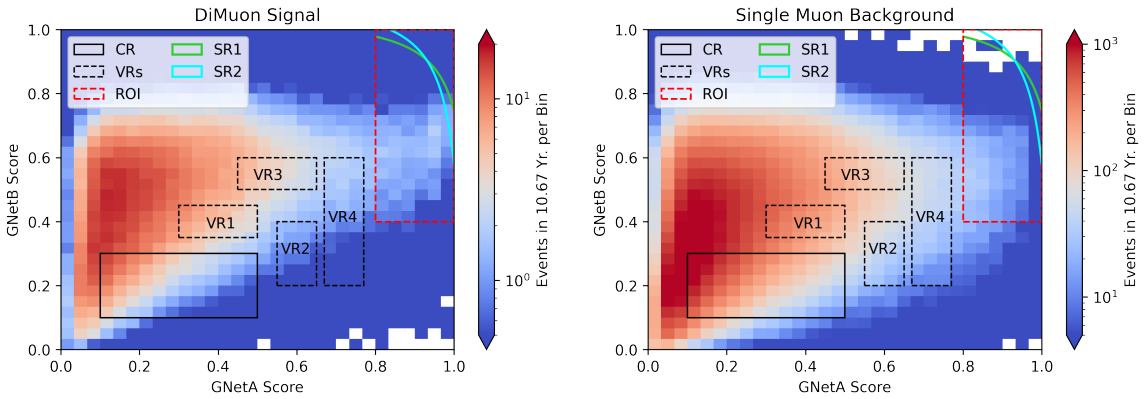


Figure 6.11: 2D Score distributions of signal (left) and background (right) showing the different analyses regions. Solid black box : CR, dashed black boxes : VRs, dashed red box: ROI, solid green and cyan curves: SRs.

with corresponding uncertainties are calculated from the nominal and systematic MC sets. The model expectation is then compared to the observed IceCube data using the statistical data analysis method discussed in the following section. Figure 6.12 shows the full data processing and event selection chain for the MC simulation and IceCube data.

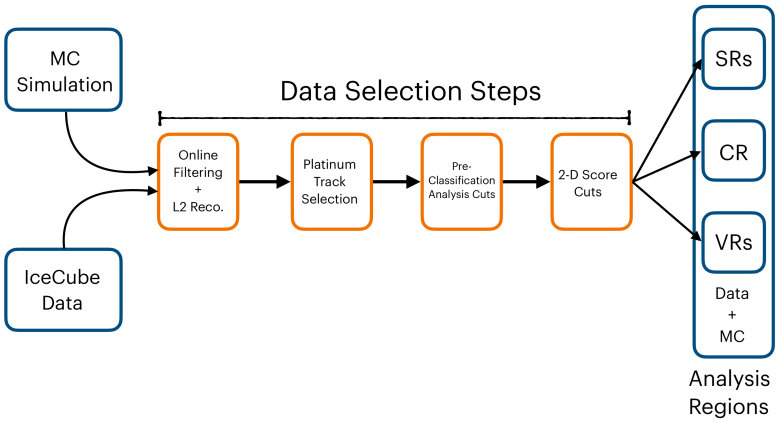


Figure 6.12: Flow diagram showing the complete event selection method for dimuon search analysis. The expected (MC) and observed (data) events of the analysis regions are used in the final statistical data analysis.

6.4 Statistical Analysis Method

The statistical data analysis in this work is performed using a software package called `HistFitter` [210] developed within the ATLAS collaboration for likelihood fits. The search for dimuons is developed using a method called *counting analysis* and consists of three different fits of the model based on the statistical prescriptions described in Ref. [210, 211].

6.4.1 Background-Only Fit

As the name suggests, the background-only fit is used to validate the background event expectation constructed for the model by quantifying the data-MC agreement. Since the data is compared with the MC expectation of the background events, the fit involves only the control and validation regions as the analysis regions where the signal contamination is assumed to be negligible. The fit method performs a measurement on the control region. It then extrapolates the result to predict the background events in the validation region (denoted as n_{VR}^{pred}) using the following expression,

$$n_{VR}^{pred} = \left[\frac{n_{CR}^{obs}}{B_{CR}^{MC}}\right] \times B_{VR}^{MC} = \Phi_N \times B_{VR}^{MC} \quad , \tag{6.6}$$

where n_{CR}^{obs} is the number of observed events in the control region. B_{CR}^{MC} and B_{VR}^{MC} are the MC expectation of the background events in the control and validation regions, respectively. The term within the square bracket in the above equation is referred to as the *scaling factor* Φ_N which extrapolates the control region measurement to the validation regions.

Treatment of Uncertainties in the Fit

Each term in Equation (6.6) also has associated uncertainties used to compute the predicted uncertainty of n_{VR}^{pred} due to the fit, denoted as σ_{pred}. The observed number of events in the control region, n_{CR}^{obs}, carries the Poisson statistical uncertainty. The uncertainties related to the nominal MC expectations, B_{CR}^{MC} and B_{VR}^{MC} come from the systematic effects (discussed in Section 6.2) and the limited MC statistics (i.e. the actual number of simulated events in the corresponding analysis regions). The

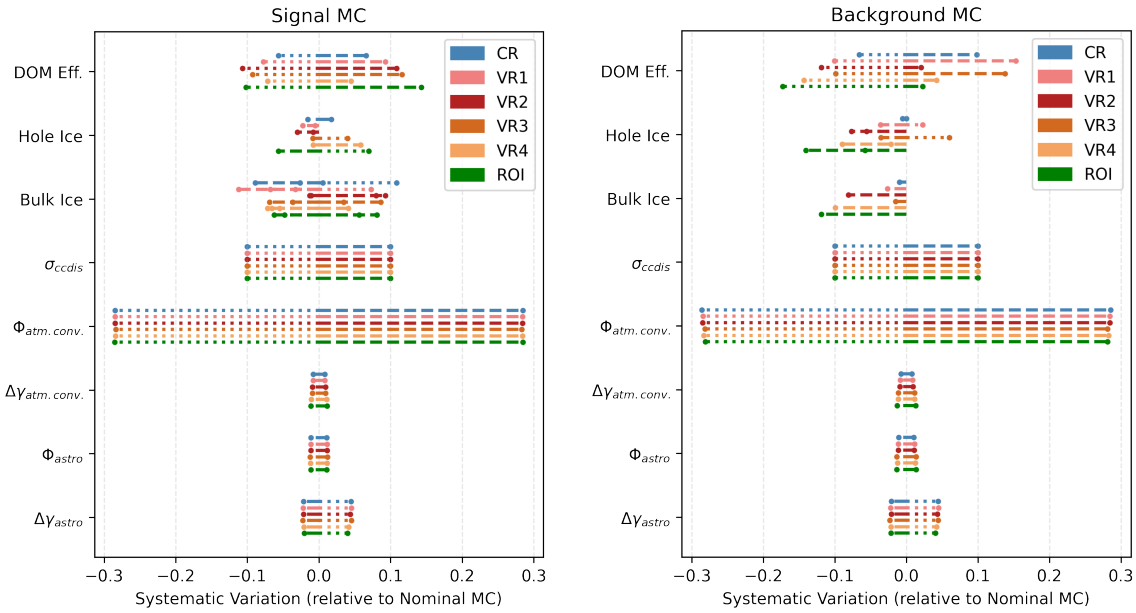


Figure 6.13: The relative variation (around nominal MC) due to each systematic effect in the analysis regions for signal (left) and background (right) MC. The two sides of each systematic effect (upper and lower variation around the nominal) are plotted in separate lines (dashed and dotted) to investigate the correlations of the effects across the analysis regions.

variation in the number of expected events in each analysis region due to the systematic effects are computed using the similar calculation shown in Equation (6.2) (but replacing % with fraction and treating each analysis region as a single bin) and are plotted in Figure 6.13 for the signal and background MC. The plots show that the upper and lower variation (i.e. the variations related to either $+1\sigma$ or -1σ change in the systematic parameter) of DOM Eff., cross section, and the flux uncertainties move in a largely correlated way across the analysis regions for both the signal and background events. Conversely, there are very little or no correlations across the analysis regions for the hole ice and bulk ice effects. Therefore, the changes in the event expectation due to all individual systematic variations are grouped into two overall systematic effects and are referred to as *correlated* (denoted as the parameter α_{corr}, which includes DOM Eff., cross-section, and the flux uncertainties) and *uncorrelated* (denoted as the parameter α_{uncorr}, which includes hole ice and bulk ice) systematics for each analysis region. Conservative estimates of $\pm1\sigma$ priors for these systematic parameters (α_{corr} and α_{uncorr}) are constructed for each analysis region and *sample*

(signal and background) by adding the variations from individual systematics of each group in quadrature. The total systematic effect of each group from the quadrature sum has an underlying assumption of no correlation among the different sources of individual systematic uncertainty. The upper and lower 1σ values of α_{corr} are calculated separately by adding the corresponding variations of the correlated systematic effects in quadrature and are expressed as,

$$\alpha^{\pm}_{corr} = \sqrt{\sum_{i} (var^{\pm}_{i})^2} \quad \text{with} \quad i \in \{\text{Correlated Systematics}\} \quad , \tag{6.7}$$

where α^{+}_{corr} and α^{-}_{corr} correspond to the upper and lower variations, $\{var^{+}_{i}\}$ and $\{var^{-}_{i}\}$, respectively. The fitting process assumes the α_{corr} values of all the analysis regions and samples to be completely correlated.

The uncorrelated systematic α_{uncorr} includes the effects from the hole ice and bulk ice, which have one-sided variations for several analysis regions and samples, as shown in Figure 6.13. To account for a conservative estimate of two-sided variations, α_{uncorr} is treated to be symmetric around the nominal value (0), and $\pm 1\sigma$ width is calculated by taking the maximum variation of each systematic effect in the quadrature sum,

$$\alpha^{\pm}_{uncorr} = \sqrt{\sum_{i} \left(\max\{var^{+}_{i}, var^{-}_{i}\}\right)^2} \quad \text{with} \quad i \in \{\text{Uncorrelated Systematics}\}. \tag{6.8}$$

The α_{uncorr} values for each region and sample are treated as independent systematic effects during the fit. The uncertainties due to the limited MC statistics are included in the analysis as additional Poisson errors (denoted as γ_{mc}) related to the number of simulated events. For each analysis region and sample, this is defined as,

$$\gamma^{\pm}_{mc} = \sqrt{\sum_{i} w_i^2} \quad , \tag{6.9}$$

where w_i is the final weight for each simulated event i in the corresponding analysis region and sample. The fitting algorithm in HistFitter takes the systematic and MC statistics parameters ($\alpha_{corr}, \alpha_{uncorr}$, and γ_{mc}) as the nuisance parameters related to each analysis region and sample as the inputs and treats them correspondingly using the error propagation formula discussed in Ref. [210].

The background-only fit is performed by passing HistFitter the background nominal MC expectation, the related priors for the nuisance parameters, and the observed

events in the data in CR and the validation regions (VR1, VR2, VR3, and VR4). It then returns the estimated number of background events and uncertainty (n^{pred}, σ_{pred}) for each validation region as the fit results. A model validation parameter, called *pull* (denoted as χ), can be defined using the fit results and the observed events for each validation region as,

$$\chi_i = \frac{n_i^{obs} - n_i^{pred}}{\sqrt{\sigma_{pred,i}^2 + \sigma_{stat,i}^2}} = \frac{n_i^{obs} - n_i^{pred}}{\sigma_{tot,i}} \quad \text{with} \quad i \in \{\text{VR1, VR2, VR3, VR4}\} \quad , \quad (6.10)$$

where σ_{stat} is the Poisson uncertainty of the expected number of events in the corresponding validation region. The pulls of the validation regions, thus, provide a measurement of the data-MC agreement in the background-dominated region of the model. The background-only fit and the calculation of the pulls in this work are used as a checkpoint to test for any mismodelling issues, i.e. underestimation or overestimation of the data, before moving to the analysis of the signal regions.

6.4.2 Discovery Fit

The search analysis of dimuons is developed by constructing a likelihood model $\mathcal{L}$ defined by the joint probability distributions of the event counts in the signal and control region and of the nuisance parameter sets $\boldsymbol{\theta} = \{\boldsymbol{\alpha_{corr}}, \boldsymbol{\alpha_{uncorr}}, \boldsymbol{\gamma_{mc}}\}$,

$$\begin{aligned} \mathcal{L}\left(n_{SR}^{obs}, n_{CR}^{obs}, \boldsymbol{\theta^0} | \mu_s, \Phi_N, \boldsymbol{\theta}\right) = \; & \mathbb{P}\text{ois}\left(n_{SR}^{obs} | \lambda_{SR}\left(\mu_s, \Phi_N, \boldsymbol{\theta}\right)\right) \times \\ & \mathbb{P}\text{ois}\left(n_{CR}^{obs} | \lambda_{CR}\left(\Phi_N, \boldsymbol{\theta}\right)\right) \times \\ & \mathbf{C}_{\text{nuisance}}\left(\boldsymbol{\theta^0}, \boldsymbol{\theta}\right) \quad , \end{aligned} \quad (6.11)$$

where $\mathbb{P}$ois denotes the Poisson probability of the event counts for the observed (n^{obs}) and expected (λ) events in the corresponding analysis regions (signal and control). $\mathbf{C}_{\text{nuisance}}$ denotes the product of the functions describing the probability distributions related to the nuisance parameters, with $\boldsymbol{\theta^0}$ being the central values of the parameters related to the nominal MC event counts. The expected number of events in the control region is modelled assuming the contribution from only the background processes. It depends on the scaling factor Φ_N (introduced in Equation (6.6)) and the fluctuation of the nuisance parameters $\boldsymbol{\theta}$. The general form of the expected number of events in

the signal region, λ_{SR} can be expanded as,

$$\lambda_{SR}\left(\mu_s, \Phi_N, \boldsymbol{\theta}\right) = \mu_s \cdot S_{SR}^{MC}(\boldsymbol{\theta}) + \Phi_N(\boldsymbol{\theta}) \cdot B_{SR}^{MC}(\boldsymbol{\theta}) \quad , \tag{6.12}$$

where S_{SR}^{MC} is the expected number of signal events in the signal region, computed from the MC simulation. The background events in the signal region are estimated from the MC simulation (B_{SR}^{MC}) and the measurement in the control region (Φ_N) using the similar method described in the background-only fit. μ_s is the signal strength parameter and dictates the content of the dimuon events in the signal region. The value of $\mu_s = 0$ refers to the model with no dimuon events, and the value of $\mu_s = 1$ refers to the model with a dimuon contribution equal to the Standard Model prediction. As shown in Equation (6.11) and Equation (6.12), the MC expectations are the functions of the nuisance parameters ($\boldsymbol{\theta}$), which govern the fluctuation of the expected events due to the systematic effects and the limited MC statistics. The set of probability density functions in $\mathbf{C}_{\text{nuisance}}$ defines the continuous variations of the expected number of events around the nominal values due to the corresponding nuisance parameters. The variations due to the systematic effects, i.e. the parameters $\{\boldsymbol{\alpha_{corr}}, \boldsymbol{\alpha_{uncorr}}\}$ are constructed to follow the Gaussian probability distributions with $\pm 1\sigma$ prior widths as the values calculated in Equation (6.7) and Equation (6.8) and with the centres at $\boldsymbol{\alpha}_{\boldsymbol{corr}}^{\mathbf{0}} = \boldsymbol{\alpha}_{\boldsymbol{uncorr}}^{\mathbf{0}} = 0$ (nominal MC). The variations due to the limited MC statistics are modelled as the Poisson fluctuation with $\pm 1\sigma$ prior uncertainty defined in Equation (6.9).

The signal regions contain low statistics of the simulated events ($\sim 10 - 100$) for both signal and background processes, resulting in a dominating contribution from the MC statistical uncertainty in calculating the systematic variations for the regions. An extraordinarily large amount of computation power would be required to produce sufficient MC to accurately determine the systematic uncertainties in these regions. Instead, the systematic variations of $\boldsymbol{\alpha_{corr}}$ and $\boldsymbol{\alpha_{uncorr}}$ in SR1 and SR2 are estimated using the systematic variations for the entire ROI which includes both the signal regions. The variations due to the individual systematic effects in ROI (shown in Figure 6.13) are directly used to compute $\boldsymbol{\alpha_{corr}}$ and $\boldsymbol{\alpha_{uncorr}}$ for the signal and background MC in the signal regions. The nominal event expectations in 10.67 years and the corresponding nuisance parameter priors are summarized in Table 6.5 and

Analysis Regions	Sample	Nominal Expectation	$\pm 1\sigma$ Prior		
			$\gamma_{mc}[\pm]$ (Abs.)	$\alpha_{corr}[+,-]$ (Rel.)	$\alpha_{uncorr}[\pm]$ (Rel.)
CR	Signal	261	2	$[0.32, 0.31]$	0.11
	Background	21343	54	$[0.33, 0.32]$	0.01
SR1	Signal	1.25	0.13	$[0.34, 0.33]$	0.11
	Background	0.21	0.13	$[0.31, 0.35]$	0.18
SR2	Signal	3.27	0.21	$[0.34, 0.33]$	0.11
	Background	2.13	0.50	$[0.31, 0.35]$	0.18

Table 6.5: The MC inputs for the likelihood model used in performing the final analysis fits. The priors for the systematic parameters ($\alpha_{corr}, \alpha_{uncorr}$) are reported as the relative error with respect to the nominal MC. The parameter for the limited MC statistics (γ_{mc}) shows absolute errors (number of events in 10.67 years). The systematic parameters for SR1 and SR2 are the same as they are extrapolated from ROI.

form the complete MC inputs for constructing the likelihood model in Equation (6.11).

The likelihood function in Equation (6.11) tests the hypothesized value of the signal strength parameter μ_s, given the observed data (n_{SR}^{obs} and n_{CR}^{obs}). In particular, the hypothesis tests involve constructing a profile likelihood ratio defined as,

$$\Lambda(\mu_s) = \frac{\mathcal{L}(\mu_s, \hat{\hat{\Phi}}_N, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}_s, \hat{\Phi}_N, \hat{\boldsymbol{\theta}})} \quad , \tag{6.13}$$

where the numerator is a conditional maximum likelihood estimator which finds the optimized parameters $\{\hat{\hat{\Phi}}_N, \hat{\hat{\boldsymbol{\theta}}}\}$ for a given fixed value of μ_s. The likelihood function in the denominator is maximized for all the free parameters to find the best-fit values, $\{\hat{\mu}_s, \hat{\Phi}_N, \hat{\boldsymbol{\theta}}\}$. A hypothesis test, referred to as the *discovery fit*, measures a level of disagreement between data and the model with the background-only hypothesis (referred to as the null hypothesis with $\mu_s = 0$). The test statistic for such a test can be defined using the profile likelihood ratio,

$$q_0 = \begin{cases} -2\ln\Lambda(0) & , \quad \hat{\mu}_s \geq 0 \\ 0 \quad , & \hat{\mu}_s < 0 \quad . \end{cases} \tag{6.14}$$

In the context of dimuon search, the test statistic value $q_0 = 0$ refers to the scenario where the observed data is more consistent with the background-only (null, no-dimuon) hypothesis than a model with non-zero dimuon expectation. For a large value of the test statistic, $q_0 \gg 0$ represents a scenario where the background-only model has a large discrepancy in describing the observed data and, instead, indicates a more probable scenario with the presence of a signal (dimuon) contribution. Given the observed data (n_{SR}^{obs} and n_{CR}^{obs}), a level of disagreement between the observation and the null hypothesis can be quantified using the *p-value*, which is defined as,

$$p_0 = \int_{q_{0,\text{obs}}}^{\infty} f(q_0|\mu_s = 0)\ dq_0 \quad , \tag{6.15}$$

where $q_{0,\text{obs}}$ is the test statistic value computed for the observed data, and $f(q_0|0)$ is the PDF of the test statistic q_0 under the assumption of the background-only hypothesis ($\mu_s = 0$). The p-value can be converted into an equivalent *discovery significance* (Z_0) using the formula discussed in Ref. [211]. The measurement from the fit result can then be reported as the discovery significance Z_0 with which the background-only (null) hypothesis can be rejected. The calculation of the p-value and corresponding significance Z_0 requires the construction of the q_0 distribution, $f(q_0|0)$. Multiple pseudo experiments are generated by randomizing the dummy observed number of events from the Poisson distributions and the nuisance parameters from the distributions, $\mathbf{C}_{\text{nuisance}}$, for the null hypothesis to build the test statistic distribution. An additional test statistic distribution can be generated using the model with nominal signal expectation ($\mu_s = 1$), referred to as *alternate hypothesis* and is used to compute the expected sensitivity for the Standard Model dimuon scenario (discussed in Section 6.4.4).

6.4.3 Exclusion Fit

To measure an upper limit on the signal strength parameter μ_s from the observed data, a hypothesis test called *exclusion fit* is performed by constructing a different test statistic,

$$q_{\mu_s} = \begin{cases} -2\ln\Lambda(\mu_s) \quad , & \hat{\mu}_s \leq \mu_s \\ 0 \quad , & \hat{\mu}_s > \mu_s \quad . \end{cases} \tag{6.16}$$

The nominal signal model ($\mu_s = 1$) is considered as the null hypothesis in the exclusion fit. The test statistic q_{μ_s} describes the compatibility between the data and the model with a specific value of μ_s. The p-value for the exclusion fit denoted as CL_s is, thus, a function of μ_s and has a definition similar to the p-value of the discovery fit,

$$\mathrm{CL}_s(\mu_s) = \int_{q_{\mu_s,\mathrm{obs}}}^{\infty} f(q_{\mu_s}|\mu_s)\ dq_{\mu_s} \quad , \tag{6.17}$$

where $q_{\mu_s,\mathrm{obs}}$ is the test statistic value corresponding to the observation, and $f(q_{\mu_s}|\mu_s)$ is the PDF of q_{μ_s} under the model hypothesis μ_s. The CL_s value for $\mu_s = 1$ measures the disagreement between the data and the nominal model expectation, i.e. the increasing CL_s value corresponds to better compatibility of the data with the Standard Model dimuon expectation. The exclusion fit is performed for multiple hypotheses by varying μ_s, and the corresponding CL_s values are used to determine the upper limit of the signal strength, denoted as μ_s^{UL}. The 95% confidence level (CL) upper limit value is evaluated at a point where the CL_s value falls below 0.05(5%). The upper limit scan can also measure the $\pm 1\sigma$ uncertainties of μ_s^{UL} using the method discussed in [211].

6.4.4 Expected Asimov Sensitivity

The model describing the Standard Model dimuon signal events and backgrounds can be characterized by evaluating the expected sensitivity of the search analysis before performing the fits on the IceCube data. A special dataset (comprising the event counts in the signal and control regions), called the *Asimov dataset*, is formed from the simulated events representing the observed (pseudo) data. Asimov dataset is defined such that the best-fit parameter values from the maximum likelihood estimator return all the true parameters of the injected model.

For the discovery fit, the expected sensitivity is evaluated by injecting the expected number of MC events under the nominal signal model ($\mu_s = 1$) hypothesis as the observed data and computing the significance Z_0 (referred to as median significance) with which the background-only hypothesis ($\mu_s = 0$) can be rejected. The Asimov dataset for the exclusion fit is generated using only the nominal background MC (i.e. the hypothesis with $\mu_s = 0$). This dataset is used as the observed events to compute CL_s value under the assumption of $\mu_s = 1$ hypothesis and μ_s^{UL} (95% CL)

from the upper limit scan. These results are reported as the expected sensitivity for the exclusion fit.

Since the two signal regions, SR1 and SR2, significantly overlap in the 2-D score space (Figure 6.11), they are not statistically independent. SR1 and SR2, thus, require separate fits for evaluating the expected sensitivities in the respective regions. HistFitter software package contains the built-in definitions of the fit methods discussed above and only requires constructing the likelihood function in Equation (6.11) using the MC inputs from Table 6.5 and the observed data. To build the test statistic distributions for the discovery and exclusion fits, 5000 pseudo experiments are simulated for each null and alternate hypothesis. The upper limit scan of the exclusion fits is implemented by computing the CL_s values for 40 different scan points of μ_s in the range $[0, 20]$. The fit results as the expected sensitivities of the discovery and exclusion fits for SR1 and SR2 are shown in Table 6.6. For the discovery fits, SR1 and

Signal Region	Discovery Fit			Exclusion Fit		
	p-value	med$[Z_0\|\mu_s = 1]$	TS	CLs$(\mu_s = 0)$	μ_s^{UL}(95%CL)	$[-1\sigma, +1\sigma]$
SR1	0.0214	**2.03**	1.52	0.33	**2.19**	[1.91, 5.50]
SR2	0.0361	**1.80**	1.52	0.15	**1.98**	[1.28, 4.68]

Table 6.6: Expected sensitivity assuming the Asimov dataset as the 10.67 years observed data. TS denotes the test statistic value computed for the Asimov dataset.

SR2 are expected to achieve 2.03σ and 1.80σ significance, respectively, with which the background-only hypothesis can be rejected. The 95% CL exclusion upper limit of the signal strength parameter, μ_s^{UL} is expected to be 2.19 and 1.98 for SR1 and SR2, respectively.

6.5 Results

Before the final analysis of the full 10.67 years dataset, an additional step is considered as part of the data unblinding process to check for any potential mismodelling issues. The step involves investigating a small fraction of the data (7.1% of the total data with 0.76 years of livetime) by comparing the data-MC agreement of the four final

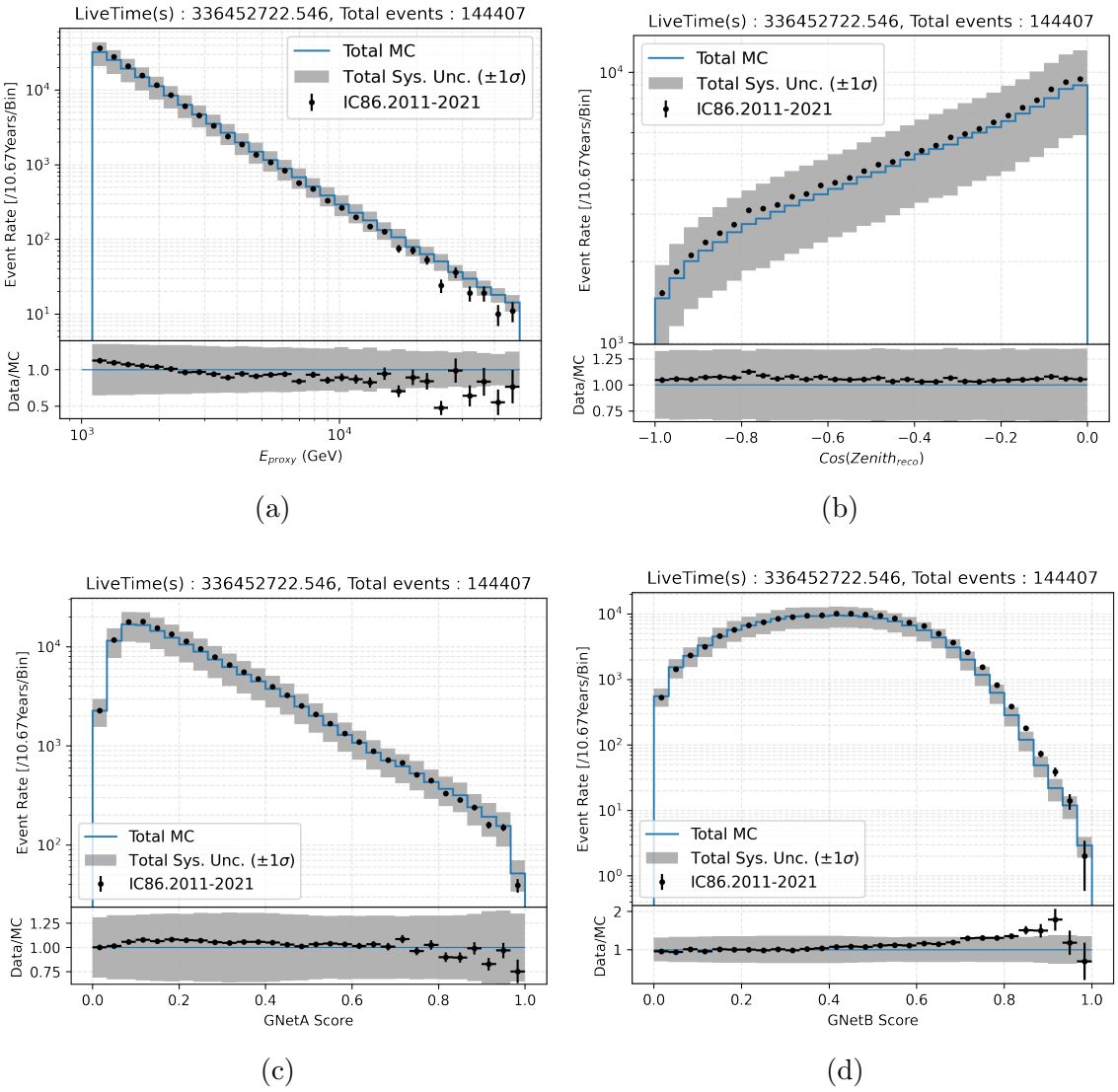


Figure 6.14: Comparison of the event rate distributions of the final observable (reconstructed energy (a), cosine zenith (b), GNetA (c), and GNetB (d)) between the total MC expectation and the observed data. The total MC expectation is calculated from the sum of all the signal and background processes shown in Figure 6.7. The shaded grey region around the nominal expectation is the $\pm 1\sigma$ prior uncertainty due to the systematic effects, i.e. all the individual systematic variations added in quadrature. The error bars of the data points denote the statistical uncertainties.

observables and by performing the background-only fit. No significant deviation in the data-MC agreement and the background-only fit results are observed.

After the preliminary checks on the small fraction of the data, the full IceCube dataset (10.67 years of livetime) is processed up to the pre-classification analysis level and is compared with the MC expectation using the event rate distributions of

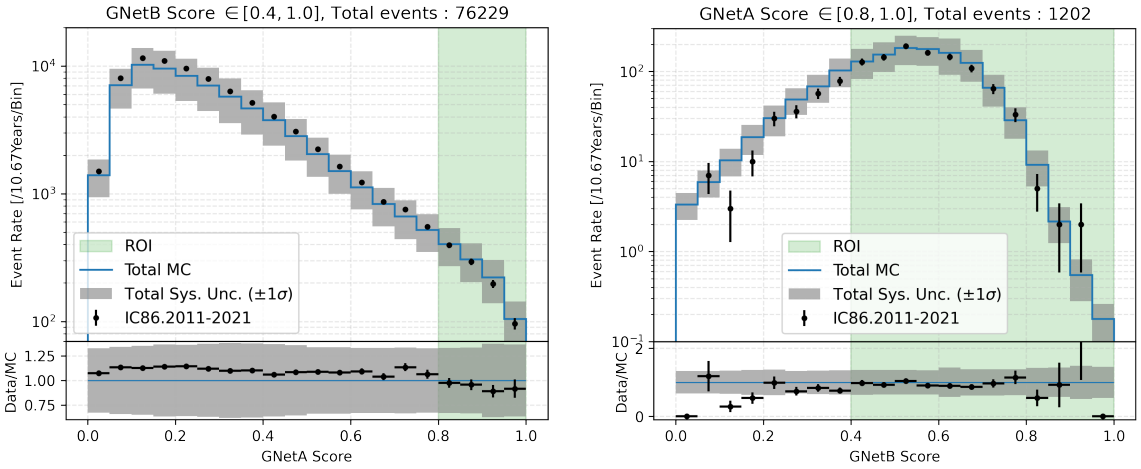


Figure 6.15: (Left) The data-MC comparison of the events with GNetB score $\in [0.4, 1.0]$ as a function of the GNetA score. The excess of events found in high GNetB score region (shown in Figure 6.14) can be attributed to the event excess in the low GNetA scores in this plot (outside the vertical green-shaded ROI region). (Right) The data-MC comparison of the events with GNetA score $\in [0.8, 1.0]$ as a function of the GNetB score. Most of the events in this plot are inside ROI and have good data-MC agreement.

the four observable: reconstructed energy, cosine zenith, GNetA, and GNetB scores. Figure 6.14 shows the comparison of the distributions and the corresponding data/MC ratio before performing any fits. All the distributions except for the GNetB score in the figure show good data-MC agreement. The global upward shift of the data in the reconstructed cosine zenith distribution is associated with the overall normalization uncertainty of the model and will get corrected in the analysis fits by the scaling factor, Φ_N.

The GNetB score distribution below 0.4 is found to have good data-MC agreement. However, the data/MC ratio with GNetB scores above 0.4 is observed to have a gradual upward shift, i.e. excess of events in the data with increasing class scores. The events from the region with data-MC disagreement (GNetB score $\in [0.4, 1.0]$) are investigated by plotting their distribution in GNetA score, as shown in Figure 6.15 (left plot). The data/MC ratio in the plot shows that the excess events have low GNetA scores, and thus, are outside the ROI. In particular, the events with the low-GNetA and high-GNetB scores belong to the top left region in the 2-D score space (as in Figure 6.11) and are outside any of the analysis regions (control, validation or

signal) used in the final fits. For further confirmation, the GNetB score distribution of the events with GNetA score ≥ 0.8 are also plotted in Figure 6.15 (right plot), and the data/MC ratio is shown to have good agreement in ROI. Therefore, the discussed data-MC disagreement in the full GNetB score distribution is considered to have no significant impact on the final dimuon search analysis.

6.5.1 Background-Only Fit Results

After comparing the data and MC in the four observable, the model validation in the validation regions is performed using the background-only fit described in Section 6.4.1. The number of observed events and the predicted background events from the fit result are shown in Table 6.7, along with the associated uncertainties. The

Fit Details	CR	VR1	VR2	VR3	VR4
Observed events	21335 ± 146	5499 ± 74	743 ± 27	3063 ± 55	1295 ± 36
Fitted Bkg events	21335 ± 146	5398 ± 265	830 ± 93	2766 ± 179	1270 ± 153

Table 6.7: The observed and predicted background events obtained from the background only fit of the CR followed by the extrapolation of the fit results to the VRs. The errors in the observed events are the statistical uncertainties. The uncertainties on the fitted background events are calculated using the error propagation method for the background-only fit (discussed in Section 6.4.1).

values from the table are then used to calculate the pulls in the validation regions using Equation (6.10), which are shown in Figure 6.16. The plot shows that the observed number of events in the three validation regions (VR1, VR2, and VR4) have small deviations ($< \pm 1\sigma$) from the MC expectations, and the pull in VR4 is slightly high with $\sim 1.5\sigma$ deviation. Since there is no statistical discrepancy between the MC and observed data, i.e. only one in four measurements is more than $\pm 1\sigma$ away, the background model is shown to be a good description of the control and validation regions and should be suitable for performing the final dimuon search analysis on the signal regions.

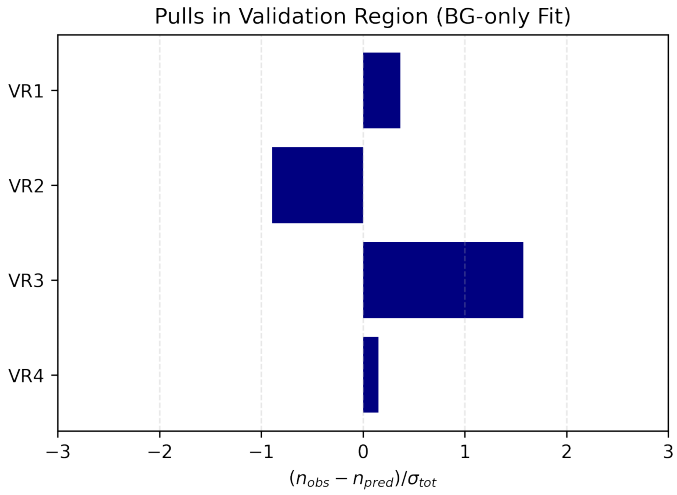


Figure 6.16: Validation pulls of the background-only fit.

Sample	CR				SR1				SR2			
	n_{exp}	$\gamma(\pm1\sigma)$	$\alpha_{corr}([+,-])$	$\alpha_{uncorr}(\pm1\sigma)$	n_{exp}	$\gamma(\pm1\sigma)$	$\alpha_{corr}([+,-])$	$\alpha_{uncorr}(\pm1\sigma)$	n_{exp}	$\gamma(\pm1\sigma)$	$\alpha_{corr}([+,-])$	$\alpha_{uncorr}(\pm1\sigma)$
Total Signal	**261**	2	[0.32, 0.31]	0.11	**1.25**	0.13	[0.34, 0.33]	0.11	**3.27**	0.21	[0.34, 0.33]	0.11
Total Background	**21343**	54	[0.33, 0.32]	0.01	**0.21**	0.13	[0.31, 0.35]	0.18	**2.13**	0.50	[0.31, 0.35]	0.18
Total Expected	**21604**	–	–	–	**1.46**	–	–	–	**5.40**	–	–	–
Total Observed	**21335**	–	–	–	**1**	–	–	–	**4**	–	–	–

Table 6.8: The nominal event expectations from signal and background processes and the nuisance parameter priors for 10.67 years. The table also shows the total expected MC events (signal + background) and observed data events in CR, SR1, and SR2.

6.5.2 Dimuon Search Results

The complete information needed to construct the likelihood model for the discovery and exclusion fits of the signal regions is summarized in Table 6.8. As shown in the table, the observed number of events in SR1 and SR2 are 1 and 4, respectively, compared to the total nominal MC expectation of 1.46 events in SR1 and 5.40 events in SR2. The methods described in Section 6.4.2 and Section 6.4.3 are used to perform the discovery and exclusion fits on the observed data. The measurements from the fit results are summarized in Table 6.9. The observed significance of the discovery fits (Z_0) is less than the expected sensitivity (shown in Table 6.6) due to the small under fluctuation of the data compared to the nominal MC in the signal regions. For the one observed event in SR1, the background-only hypothesis is rejected at 1.31σ. The measured CL_s value under the assumption of dimuons from the Standard Model

Signal Region	Discovery Fit			Exclusion Fit		
	p-value	Z_0	TS	CLs	$\mu_s^{UL}(95\%\text{CL})$	$[-1\sigma, +1\sigma]$
SR1	0.0956	**1.31**	0.74	0.43	**3.99**	[1.76, 5.32]
SR2	0.1246	**1.15**	0.58	0.37	**2.38**	[1.13, 5.44]

Table 6.9: The final analysis results from the discovery and exclusion fits. TS refers to the test statistic value of the observed data. The last column denotes the $\pm 1\sigma$ uncertainty of μ_s^{UL} obtained from the upper limit scans.

($\mu_s = 1$) and the exclusion upper limit, μ_s^{UL} at 95% CL are reported to be 0.43 and 3.99, respectively, for SR1. The discovery fit based on the four observed events in SR2 results in a rejection of the background-only hypothesis at 1.15σ. The results related to the exclusion fits in SR2 are reported to be 0.37 and 2.38 for CL_s and μ_s^{UL}, respectively. The test statistic distributions of the null and alternate hypotheses and the test statistic value of the observed data for the discovery fits are shown in Figure 6.17.

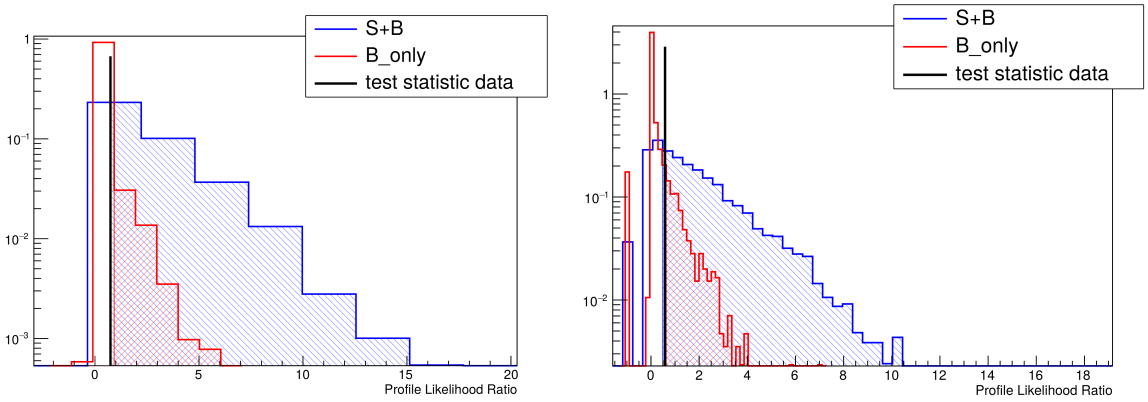


Figure 6.17: The test statistic distributions of SR1 (left) and SR2 (right) for the discovery fit, generated from 5000 pseudo experiments for each hypothesis model. The distribution for the null hypothesis (B_only) is constructed under the model assumption with $\mu_s = 0$. The alternate hypothesis (S+B) refers to the nominal signal model ($\mu_s = 1$). The black vertical line shows the test statistic value of the observed data, $q_{0,\text{obs}}$. The areas shown in hatches denote the part of the distributions used for corresponding p-value calculation.

6.5.3 Posterior Checks

In addition to the dimuon search fits, the properties of the observed events in the signal regions are analyzed as a post-unblinding step. The details of these observed events are provided in Table 6.10. The reconstructed energy, cosine zenith, GNetA

Events	ID	E_{reco} (TeV)	$\cos(\text{zenith}_{reco})$	GNetA Score	GNetB Score
Event 1 (SR1, SR2)	136132/46343498	1.22	-0.30	0.96	0.90
Event 2 (SR2)	126778/25509256	1.31	-0.11	0.99	0.69
Event 3 (SR2)	126724/44041024	1.34	-0.96	0.99	0.75
Event 4 (SR2)	118186/37203675	5.71	-0.11	0.99	0.68

Table 6.10: The details of the observed events in the signal regions. Event 1 is observed in both SR1 and SR2.

and GNetB score distributions of the signal and background MC in SR1 (Figure 6.18) and SR2 (Figure 6.19) are also shown with an overlay of the observed event properties. The plots show that the properties of most of the observed events are either at or near the bins with the highest MC expectations. One event (Event 4) in SR2 has high reconstructed energy of 5.7 TeV, where the MC expectations from both the signal and background events are minimal. The visualization of each observed event in the signal regions is shown in Appendix C.

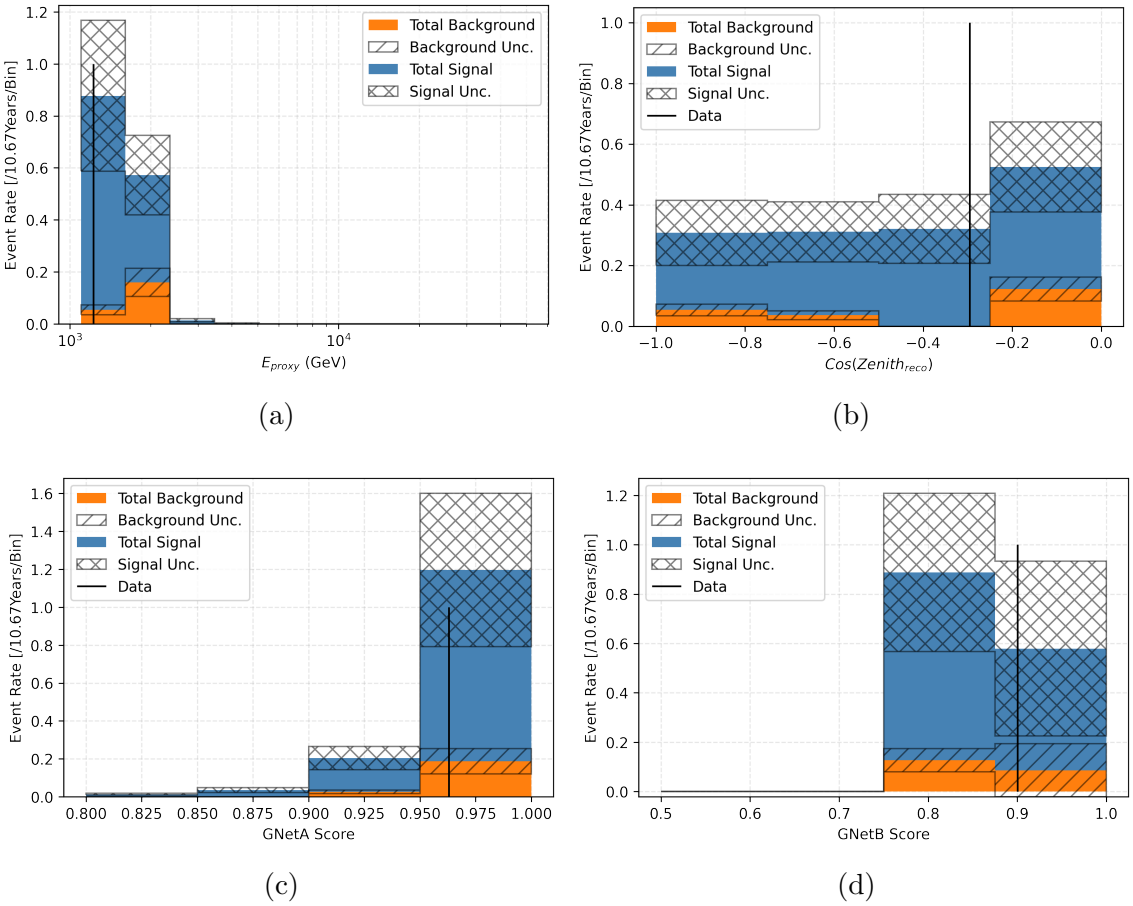


Figure 6.18: Stacked distributions of the expected signal (blue) and background (orange) events as a function of the reconstructed energy (a), cosine zenith (b), GNetA (c), and GNetB (d) scores in SR1. The properties of the observed event in SR1 are shown as the vertical black line in the plots. The prior $\pm 1\sigma$ systematic uncertainties are shown using the hatches around the signal and background expectations.

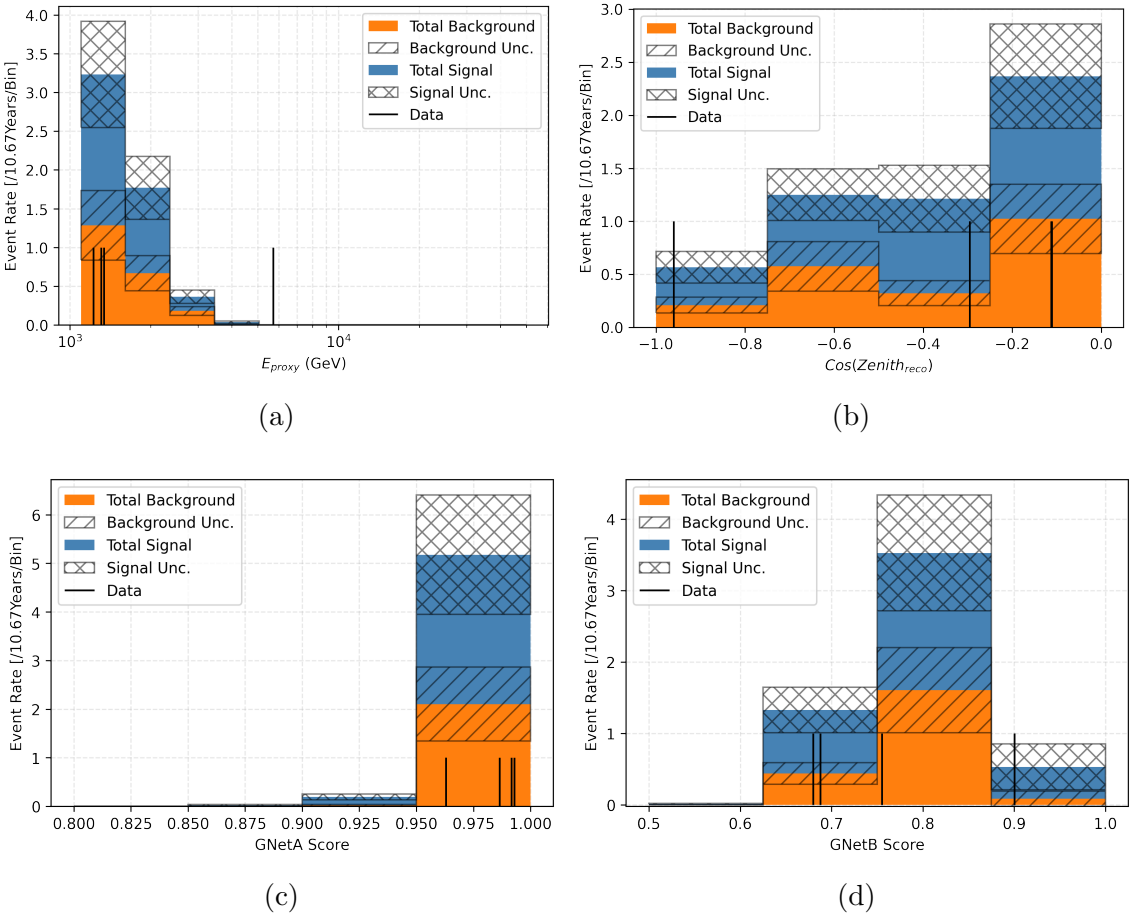


Figure 6.19: Stacked distributions of the expected signal (blue) and background (orange) events as a function of the reconstructed energy (a), cosine zenith (b), GNetA (c), and GNetB (d) scores in SR2. The properties of the 4 observed events in SR2 are shown as the vertical black line in the plots. The prior $\pm 1\sigma$ systematic uncertainties are shown using the hatches around the signal and background expectations.

Chapter 7

Conclusion and Outlook

The search for dimuons presented in this work reports the first non-zero observation of such events in IceCube. The observation of one (in SR1) and four (in SR2) events has a discovery significance (Z_0) of 1.31σ and 1.15σ, respectively. The search analysis based on the observed events also reports the exclusion upper limit of the signal strength parameter (μ_s^{UL}) at 95% CL to be 3.99 in SR1 and 2.38 in SR2. The results are entirely consistent with the Standard Model expectations. While the statistical significance fails to reach the 3σ threshold needed for claiming unambiguous evidence for dimuon events, this work is the first IceCube analysis that has shown the hints for any positive evidence for dimuon events in the detector, i.e. it is more likely to have been observed dimuon events than no dimuon scenario. Moreover, the pursuit of this work leads to the possibility of significant improvements in the search analysis, which will enable the prospect of probing QCD and new physics with higher statistics of observed dimuons. Although implementing such improvements is beyond the scope of this work, a few potential improvements are outlined below to assist the future efforts for dimuon searches.

The pre-classification analysis cuts (shown in Figure 6.12) in the event selection process are currently developed based on the reconstructed energy and direction of the tracks. Additional selection criteria can be constructed to further increase the selected tracks' quality in the context of dimuon classification. The events with a significant fraction of their tracks passing through the dust layer (at a depth between -50 m and -250 m in IceCube coordinate) contains an increased amount of scattered lights in the DOM hits, resulting in the smearing of the potential dimuon hit signature.

Such an effect makes the dimuons almost impossible to identify, and thus removing all tracks passing through the dust layer should not impact the signal efficiency but will remove many potential background events. These events can be removed by imposing cuts on the 2-D space of the reconstructed cosine zenith and the Z-coordinate of the track's centre for the charge distribution of the event (denoted as *COG-Z*). The nearly horizontal events ($\cos(zenith_{reco}) \sim 0$) with COG-Z $\in [-50$ m, -250 m] indicate a large portion of the track segments passing through the dust layer and, thus, can be removed by defining a cut region in the 2-D space. Since the DOMs in the DeepCore strings have a denser configuration than the IceCube strings, the dimuon events with hits in the DeepCore strings are expected to contain more information on the double-track signature. An improved separation between the signal and background distributions in GNetA and GNetB scores can be achieved by requiring the events to pass through the DeepCore volume. However, such a criterion will reduce the event statistics significantly and needs further investigation to examine its benefit.

In addition to the event selection process, the statistical analysis of the search can also be improved. The current method of counting analysis performs the fitting procedure with only two bins (control and signal region). This method can be replaced with a binned profile-likelihood fit of the entire 2-D distribution of the GNetA and GNetB scores. The higher number of bins in the improved fit allows for a better constraint on the free parameters (signal strength and nuisance parameters of the model), resulting in a search result with more statistical power.

Another significant analysis improvement can come from reconstructing the events with a double-track hypothesis. A likelihood model with double-track parameters such as opening angle, track-, and intersection-plane (shown in Figure 5.12) can be used to fit each event in a sample with high purity of dimuons (e.g. events in ROI). Such a likelihood model can be constructed to fit a subset of the hit strings with the highest number of DOM hits. An example MC dimuon event with reconstructed properties similar to the single event observed in both SR1 and SR2 (Event 1 from Table 6.10) is shown in Figure 7.1. The plots in the figure show the DOM hits and the light curves from reconstructed single track (MuEx) and two true muons ($\mu_{1,2}$) for the seven brightest strings of the event. The objective of the double-track likelihood model would be to reconstruct the two muon light curves close to their

true values. Combining the fit results from the double-track reconstruction and the classification scores can achieve much higher separation power in identifying dimuon events. Implementing these improvements can lead to a potentially high number of observed dimuons in IceCube and also provide information about the event kinematics related to the two muons.

In addition, the collaboration is working on an ongoing effort to deploy state-of-the-art light sensors in a dense configuration within the DeepCore volume, referred to as IceCube Upgrade [212], which can further enhance the dimuon event reconstruction and identification due to improved resolution. Therefore, the proposed future improvements of this work will facilitate IceCube's ability to discover TeV-scale neutrino dimuon events. Such an outcome will allow IceCube to join the global effort of next-generation ν-induced dimuon search programs. Future experiments like DUNE [213] and SHiP [214, 215] will lead the search for dimuon events in the low energy sector ($1 - 100$ GeV). In the high energy sector, IceCube will be accompanied by other future neutrino telescopes like P-ONE [216, 217] and KM3Net [218] and the proposed Forward Physics Facility at the LHC, which is expected to detect neutrinos with energies up to 5 TeV [219, 220].

In summary, detecting dimuon production in neutrino interactions is essential in studying weak interactions, QCD physics, and searching for new physics. Event kinematics of the dimuons from the charm production in ν CC DIS interactions offer the PDF measurements of the target nucleon's quark content (primarily strange quark). Studying the charm production is also important for ν_τ flavour identification as the high energy charm hadron can travel a short distance before decay (similar to $\tau^\pm$), making it a crucial background for future neutrino flavour analyses. Observing dimuons from neutrino trident production also has significant implications due to its ability to probe W-boson production and search for various BSM scenarios.

In conclusion, the methodology developed in this work has established the baseline performance in the search for dimuon events in IceCube and has found the first IceCube event that is more likely to be a dimuon signal than a single muon background. This baseline provides a great starting point for the next generation of analyses which will demonstrate clear evidence for dimuon events in IceCube and lead the high energy frontiers of neutrino dimuon studies.

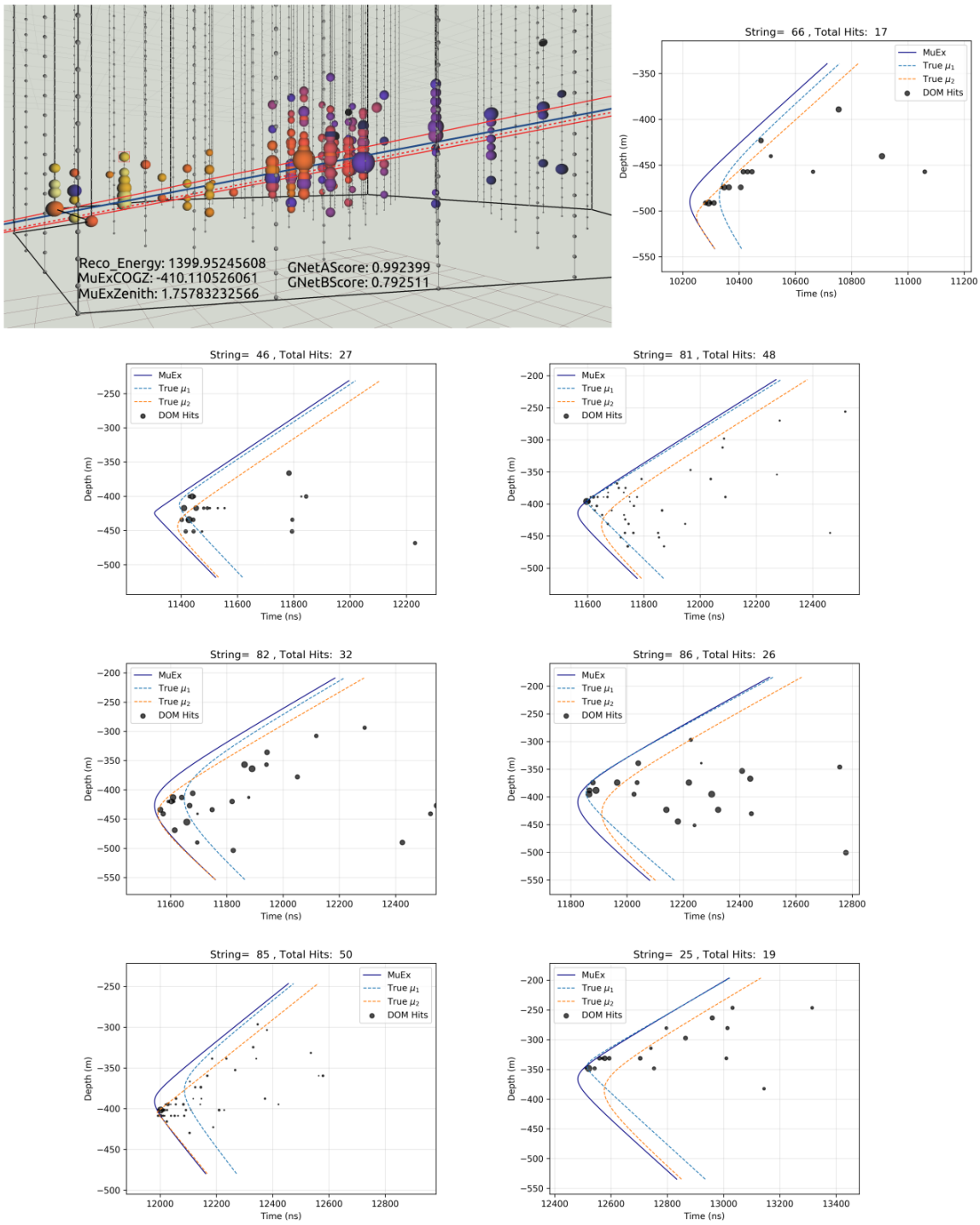


Figure 7.1: An example MC dimuon event with observable (reconstructed energy, cosine zenith, GNetA and GNetB scores) similar to Event 1 observed in SR1 and SR2. The event display (top left) shows the true muons (red solid lines) and reconstructed single-track (blue solid line) with the DOM hits. The plots for the seven brightest strings show the DOM hits at different depths as a function of time. The size of the points corresponds to the charges of the hits. The proposed double-track likelihood model will attempt to fit the true muon light curves shown in these plots.

Bibliography

[1] Mohammad Sajjad Athar et al. "Status and perspectives of neutrino physics". In: *Prog. Part. Nucl. Phys.* 124 (2022), p. 103947. DOI: 10.1016/j.ppnp.2022.103947. arXiv: 2111.07586 [hep-ph].

[2] Naoko Kurahashi, Kohta Murase, and Marcos Santander. "High-Energy Extragalactic Neutrino Astrophysics". In: *Annual Review of Nuclear and Particle Science* 72.1 (2022), pp. 365–387. DOI: 10.1146/annurev-nucl-011122-061547. eprint: https://doi.org/10.1146/annurev-nucl-011122-061547. URL: https://doi.org/10.1146/annurev-nucl-011122-061547.

[3] Martina Gerbino and Massimiliano Lattanzi. "Status of Neutrino Properties and Future Prospects—Cosmological and Astrophysical Constraints". In: *Frontiers in Physics* 5 (2018). ISSN: 2296-424X. DOI: 10.3389/fphy.2017.00070. URL: https://www.frontiersin.org/articles/10.3389/fphy.2017.00070.

[4] Q. R. Ahmad et al. "Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^{8}B solar neutrinos at the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 87 (2001), p. 071301. DOI: 10.1103/PhysRevLett.87.071301. arXiv: nucl-ex/0106015.

[5] Y. Fukuda et al. "Evidence for oscillation of atmospheric neutrinos". In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. arXiv: hep-ex/9807003.

[6] M. Aker et al. "Direct neutrino-mass measurement with sub-electronvolt sensitivity". In: *Nature Phys.* 18.2 (2022), pp. 160–166. DOI: 10.1038/s41567-021-01463-1. arXiv: 2105.08533 [hep-ex].

[7] R. Abbasi et al. "Low energy event reconstruction in IceCube DeepCore". In: *Eur. Phys. J. C* 82.9 (2022), p. 807. DOI: 10.1140/epjc/s10052-022-10721-2. arXiv: 2203.02303 [hep-ex].

[8] M. G. Aartsen et al. "Measurement of Atmospheric Neutrino Oscillations at 6–56 GeV with IceCube DeepCore". In: *Phys. Rev. Lett.* 120.7 (2018), p. 071801. DOI: 10.1103/PhysRevLett.120.071801. arXiv: 1707.07081 [hep-ex].

[9] M. G. Aartsen et al. "Measurement of Atmospheric Tau Neutrino Appearance with IceCube DeepCore". In: *Phys. Rev. D* 99.3 (2019), p. 032007. DOI: 10.1103/PhysRevD.99.032007. arXiv: 1901.05366 [hep-ex].

[10] M. G. Aartsen et al. "Development of an analysis to probe the neutrino mass ordering with atmospheric neutrinos using three years of IceCube DeepCore data". In: *Eur. Phys. J. C* 80.1 (2020), p. 9. DOI: 10.1140/epjc/s10052-019-7555-0. arXiv: 1902.07771 [hep-ex].

[11] M. G. Aartsen et al. "Neutrino oscillation studies with IceCube-DeepCore". In: *Nucl. Phys. B* 908 (2016), pp. 161–177. DOI: 10.1016/j.nuclphysb.2016.03.028.

[12] M. G. Aartsen et al. "The IceCube Neutrino Observatory - Contributions to ICRC 2015 Part V: Neutrino Oscillations and Supernova Searches". In: Oct. 2015. arXiv: 1510.05227 [astro-ph.HE].

[13] M. G. Aartsen et al. "Determining neutrino oscillation parameters from atmospheric muon neutrino disappearance with three years of IceCube DeepCore data". In: *Phys. Rev. D* 91.7 (2015), p. 072004. DOI: 10.1103/PhysRevD.91.072004. arXiv: 1410.7227 [hep-ex].

[14] M. G. Aartsen et al. "Measurement of Atmospheric Neutrino Oscillations with IceCube". In: *Phys. Rev. Lett.* 111.8 (2013), p. 081801. DOI: 10.1103/PhysRevLett.111.081801. arXiv: 1305.3909 [hep-ex].

[15] Francis Halzen. "High-energy neutrino astrophysics". In: *Nature Phys.* 13.3 (2016), pp. 232–238. DOI: 10.1038/nphys3816.

[16] M. G. Aartsen et al. "Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data". In: *Phys. Rev. Lett.* 113 (2014), p. 101101. DOI: 10.1103/PhysRevLett.113.101101. arXiv: 1405.5303 [astro-ph.HE].

[17] M. G. Aartsen et al. "Detection of a particle shower at the Glashow resonance with IceCube". In: *Nature* 591.7849 (2021). [Erratum: Nature 592, E11 (2021)], pp. 220–224. DOI: 10.1038/s41586-021-03256-1. arXiv: 2110.15051 [hep-ex].

[18] M. G. Aartsen et al. "Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert". In: *Science* 361.6398 (2018), pp. 147–151. DOI: 10.1126/science.aat2890. arXiv: 1807.08794 [astro-ph.HE].

[19] R. Abbasi et al. "Evidence for neutrino emission from the nearby active galaxy NGC 1068". In: *Science* 378.6619 (2022), pp. 538–543. DOI: 10.1126/science.abg3395. arXiv: 2211.09972 [astro-ph.HE].

[20] M. G. Aartsen et al. "Measurement of the multi-TeV neutrino cross section with IceCube using Earth absorption". In: *Nature* 551 (2017), pp. 596–600. DOI: 10.1038/nature24459. arXiv: 1711.08119 [hep-ex].

[21] M. G. Aartsen et al. "Measurements using the inelasticity distribution of multi-TeV neutrino interactions in IceCube". In: *Phys. Rev. D* 99.3 (2019), p. 032004. DOI: 10.1103/PhysRevD.99.032004. arXiv: 1808.07629 [hep-ex].

[22] R. Abbasi et al. "Measurement of the high-energy all-flavor neutrino-nucleon cross section with IceCube". In: (Nov. 2020). DOI: 10.1103/PhysRevD.104.022001. arXiv: 2011.03560 [hep-ex].

[23] A. Albert et al. "Combined search for neutrinos from dark matter self-annihilation in the Galactic Center with ANTARES and IceCube". In: *Phys. Rev. D* 102.8 (2020), p. 082002. DOI: 10.1103/PhysRevD.102.082002. arXiv: 2003.06614 [astro-ph.HE].

[24] M. G. Aartsen et al. "Search for neutrinos from decaying dark matter with IceCube". In: *Eur. Phys. J. C* 78.10 (2018), p. 831. DOI: 10.1140/epjc/s10052-018-6273-3. arXiv: 1804.03848 [astro-ph.HE].

[25] M. G. Aartsen et al. "Search for annihilating dark matter in the Sun with 3 years of IceCube data". In: *Eur. Phys. J. C* 77.3 (2017). [Erratum: Eur.Phys.J.C 79, 214 (2019)], p. 146. DOI: 10.1140/epjc/s10052-017-4689-9. arXiv: 1612.05949 [astro-ph.HE].

[26] Tao Han and Dan Hooper. "The Particle physics reach of high-energy neutrino astronomy". In: *New J. Phys.* 6 (2004), p. 150. DOI: 10.1088/1367-2630/6/1/150. arXiv: hep-ph/0408348.

[27] Valerio Bertone, Rhorry Gauld, and Juan Rojo. "Neutrino Telescopes as QCD Microscopes". In: *JHEP* 01 (2019), p. 217. DOI: 10.1007/JHEP01(2019)217. arXiv: 1808.02034 [hep-ph].

[28] Edmond L. Berger et al. "Charm-Quark Production in Deep-Inelastic Neutrino Scattering at Next-to-Next-to-Leading Order in QCD". In: *Phys. Rev. Lett.* 116.21 (2016), p. 212002. DOI: 10.1103/PhysRevLett.116.212002. arXiv: 1601.05430 [hep-ph].

[29] Shin'ichiro Ando et al. "Probing new physics with long-lived charged particles produced by atmospheric and astrophysical neutrinos". In: *JCAP* 04 (2008), p. 029. DOI: 10.1088/1475-7516/2008/04/029. arXiv: 0711.2908 [hep-ph].

[30] Paolo Lipari. "CP violation effects and high-energy neutrinos". In: *Phys. Rev. D* 64 (2001), p. 033002. DOI: 10.1103/PhysRevD.64.033002. arXiv: hep-ph/0102046.

[31] P. Vilain et al. "Leading order QCD analysis of neutrino induced dimuon events". In: *Eur. Phys. J. C* 11 (1999), pp. 19–34. DOI: 10.1007/s100520050611.

[32] S. A. Rabinowitz et al. "Measurement of the strange sea distribution using neutrino charm production". In: *Phys. Rev. Lett.* 70 (1993), pp. 134–137. DOI: 10.1103/PhysRevLett.70.134.

[33] M. Goncharov et al. "Precise Measurement of Dimuon Production Cross-Sections in ν_μ Fe and $\bar{\nu}_\mu$ Fe Deep Inelastic Scattering at the Tevatron." In: *Phys. Rev. D* 64 (2001), p. 112006. DOI: 10.1103/PhysRevD.64.112006. arXiv: hep-ex/0102049.

[34] A. Kayis-Topaksu et al. "Measurement of charm production in neutrino charged-current interactions". In: *New J. Phys.* 13 (2011), p. 093002. DOI: 10.1088/1367-2630/13/9/093002. arXiv: 1107.0613 [hep-ex].

[35] O. Samoylov et al. "A Precision Measurement of Charm Dimuon Production in Neutrino Interactions from the NOMAD Experiment". In: *Nucl. Phys. B* 876 (2013), pp. 339–375. DOI: 10.1016/j.nuclphysb.2013.08.021. arXiv: 1308.4750 [hep-ex].

[36] D. Geiregat et al. "First observation of neutrino trident production". In: *Phys. Lett. B* 245 (1990), pp. 271–275. DOI: 10.1016/0370-2693(90)90146-W.

[37] S. R. Mishra et al. "Neutrino tridents and W Z interference". In: *Phys. Rev. Lett.* 66 (1991), pp. 3117–3120. DOI: 10.1103/PhysRevLett.66.3117.

[38] T. Adams et al. "Neutrino trident production from NuTeV". In: *29th International Conference on High-Energy Physics*. July 1998, pp. 631–634. arXiv: hep-ex/9811012.

[39] Shao-Feng Ge, Manfred Lindner, and Werner Rodejohann. "Atmospheric Trident Production for Probing New Physics". In: *Phys. Lett. B* 772 (2017), pp. 164–168. DOI: 10.1016/j.physletb.2017.06.020. arXiv: 1702.02617 [hep-ph].

[40] Bei Zhou and John F. Beacom. "W-boson and trident production in TeV–PeV neutrino observatories". In: *Phys. Rev. D* 101.3 (2020), p. 036010. DOI: 10.1103/PhysRevD.101.036010. arXiv: 1910.10720 [hep-ph].

[41] Stephan Meighen-Berger et al. "New constraints on supersymmetry using neutrino telescopes". In: *Phys. Lett. B* 811 (2020), p. 135929. DOI: 10.1016/j.physletb.2020.135929. arXiv: 2005.07523 [hep-ph].

[42] Vernon Barger et al. "Neutrino signals in IceCube from weak production of top and charm quarks". In: *Phys. Rev. D* 95.9 (2017), p. 093002. DOI: 10.1103/PhysRevD.95.093002. arXiv: 1611.00773 [hep-ph].

[43] ChuanLe Sun et al. "Search for charm-quark production via dimuons in neutrino telescopes*". In: *Chin. Phys. C* 47.2 (2023), p. 023109. DOI: 10.1088/1674-1137/aca465. arXiv: 2201.03580 [hep-ph].

[44] C. L. Cowan et al. "Detection of the free neutrino: A Confirmation". In: *Science* 124 (1956), pp. 103–104. DOI: 10.1126/science.124.3212.103.

[45] W. Pauli. "Dear radioactive ladies and gentlemen". In: *Phys. Today* 31N9 (1978), p. 27.

[46] E. Fermi. "An attempt of a theory of beta radiation. 1." In: *Z. Phys.* 88 (1934), pp. 161–177. DOI: 10.1007/BF01351864.

[47] Luigi Di Lella and Carlo Rubbia. "The Discovery of the W and Z Particles". In: *Adv. Ser. Direct. High Energy Phys.* 23 (2015), pp. 137–163. DOI: 10.1142/9789814644150_0006.

[48] C. Patrignani et al. "Review of Particle Physics". In: *Chin. Phys. C* 40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.

[49] Nicola Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531.

[50] Makoto Kobayashi and Toshihide Maskawa. "CP Violation in the Renormalizable Theory of Weak Interaction". In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.

[51] B. T. Cleveland et al. "Measurement of the solar electron neutrino flux with the Homestake chlorine detector". In: *Astrophys. J.* 496 (1998), pp. 505–526. DOI: 10.1086/305343.

[52] B. Pontecorvo. "Mesonium and anti-mesonium". In: *Sov. Phys. JETP* 6 (1957), p. 429.

[53] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. "Remarks on the unified model of elementary particles". In: *Prog. Theor. Phys.* 28 (1962), pp. 870–880. DOI: 10.1143/PTP.28.870.

[54] L. Ludhova et al. "Solar neutrino physics with Borexino I". In: *47th Rencontres de Moriond on Electroweak Interactions and Unified Theories*. May 2012, p. 341. arXiv: 1205.2989 [hep-ex].

[55] K. Abe et al. "Solar neutrino results in Super-Kamiokande-III". In: *Phys. Rev. D* 83 (2011), p. 052010. DOI: 10.1103/PhysRevD.83.052010. arXiv: 1010.0118 [hep-ex].

[56] Volodymyr Takhistov. "Review of Atmospheric Neutrino Results from Super-Kamiokande". In: *PoS* ICHEP2020 (2021), p. 181. DOI: 10.22323/1.390.0181. arXiv: 2012.06864 [hep-ex].

[57] P. Adamson et al. "Measurement of Neutrino and Antineutrino Oscillations Using Beam and Atmospheric Data in MINOS". In: *Phys. Rev. Lett.* 110.25 (2013), p. 251801. DOI: 10.1103/PhysRevLett.110.251801. arXiv: 1304.6335 [hep-ex].

[58] Lukas Berns. "Recent Results from T2K". In: *55th Rencontres de Moriond on Electroweak Interactions and Unified Theories*. May 2021. arXiv: 2105.06732 [hep-ex].

[59] Erika Catano-Mur. "Recent results from NOvA". In: *56th Rencontres de Moriond on Electroweak Interactions and Unified Theories*. June 2022. arXiv: 2206.03542 [hep-ex].

[60] F. P. An et al. "Precision Measurement of Reactor Antineutrino Oscillation at Kilometer-Scale Baselines by Daya Bay". In: *Phys. Rev. Lett.* 130.16 (2023), p. 161802. DOI: 10.1103/PhysRevLett.130.161802. arXiv: 2211.14988 [hep-ex].

[61] Aldiyar Oralbaev. "Latest Results From Double Chooz Collaboration". In: *19th Lomonosov Conference on Elementary Particle Physics*. 2021. DOI: 10.1142/9789811233913_0007.

[62] Hisakazu Minakata and Hiroshi Nunokawa. "How to measure CP violation in neutrino oscillation experiments?" In: *Phys. Lett. B* 413 (1997), pp. 369–377. DOI: 10.1016/S0370-2693(97)01108-8. arXiv: hep-ph/9706281.

[63] NuFit3.1. *Three Neutrino Oscillation Fit.* URL: http://www.nu-fit.org/.

[64] Ivan Esteban et al. "Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity". In: *JHEP* 01 (2017), p. 087. DOI: 10.1007/JHEP01(2017)087. arXiv: 1611.01514 [hep-ph].

[65] Philipp Eller. *Oscillogram.* URL: https://philippeller.github.io/osc.html.

[66] P. F. De Salas et al. "Neutrino Mass Ordering from Oscillations and Beyond: 2018 Status and Future Prospects". In: *Front. Astron. Space Sci.* 5 (2018), p. 36. DOI: 10.3389/fspas.2018.00036. arXiv: 1806.11051 [hep-ph].

[67] Stefano Gariazzo et al. "Neutrino mass and mass ordering: no conclusive evidence for normal ordering". In: *JCAP* 10 (2022), p. 010. DOI: 10.1088/1475-7516/2022/10/010. arXiv: 2205.02195 [hep-ph].

[68] S. Bilenky. "Neutrinos: Majorana or Dirac?" In: (Aug. 2020). arXiv: 2008.02110 [hep-ph].

[69] Hiroshi Nunokawa, Stephen J. Parke, and Jose W. F. Valle. "CP Violation and Neutrino Oscillations". In: *Prog. Part. Nucl. Phys.* 60 (2008), pp. 338–402. DOI: 10.1016/j.ppnp.2007.10.001. arXiv: 0710.0554 [hep-ph].

[70] Bonnie T. Fleming et al. "Recent structure function results from CCFR". In: *8th International Workshop on Deep Inelastic Scattering and QCD (DIS 2000).* July 2000, pp. 103–104. arXiv: hep-ex/0011095.

[71] Debdatta Bhattacharya. "Neutrino and antineutrino inclusive charged-current cross section measurement with the MINOS near detector". PhD thesis. Pittsburgh U., 2009. DOI: 10.2172/952653.

[72] M. Tzanov et al. "Precise measurement of neutrino and anti-neutrino differential cross sections". In: *Phys. Rev. D* 74 (2006), p. 012008. DOI: 10.1103/PhysRevD.74.012008. arXiv: hep-ex/0509010.

[73] G. Onengut et al. "Measurement of nucleon structure functions in neutrino scattering". In: *Phys. Lett. B* 632 (2006), pp. 65–75. DOI: 10.1016/j.physletb.2005.10.062.

[74] Q. Wu et al. "A Precise measurement of the muon neutrino-nucleon inclusive charged current cross-section off an isoscalar target in the energy range 2.5 < E(nu) < 40-GeV by NOMAD". In: *Phys. Lett. B* 660 (2008), pp. 19–25. DOI: 10.1016/j.physletb.2007.12.027. arXiv: 0711.1183 [hep-ex].

[75] J. Mousseau et al. "Measurement of Partonic Nuclear Effects in Deep-Inelastic Neutrino Scattering using MINERvA". In: *Phys. Rev. D* 93.7 (2016), p. 071101. DOI: 10.1103/PhysRevD.93.071101. arXiv: 1601.06313 [hep-ex].

[76] R. Devenish and A. Cooper-Sarkar. *Deep inelastic scattering.* 2004. DOI: 10.1093/acprof:oso/9780198506713.001.0001.

[77] V. N. Gribov and L. N. Lipatov. "Deep inelastic e p scattering in perturbation theory". In: *Sov. J. Nucl. Phys.* 15 (1972), pp. 438–450.

[78] Guido Altarelli and G. Parisi. "Asymptotic Freedom in Parton Language". In: *Nucl. Phys. B* 126 (1977), pp. 298–318. DOI: 10.1016/0550-3213(77)90384-4.

[79] Yuri L. Dokshitzer. "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics." In: *Sov. Phys. JETP* 46 (1977), pp. 641–653.

[80] Tie-Jiun Hou et al. "New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC". In: *Phys. Rev. D* 103.1 (2021), p. 014013. DOI: 10.1103/PhysRevD.103.014013. arXiv: 1912.10053 [hep-ph].

[81] Amanda Cooper-Sarkar. "Proton Structure from HERA to LHC". In: *40th International Symposium on Multiparticle Dynamics*. Dec. 2010. arXiv: 1012.1438 [hep-ph].

[82] Richard D. Ball et al. "The PDF4LHC21 combination of global PDF fits for the LHC Run III". In: *J. Phys. G* 49.8 (2022), p. 080501. DOI: 10.1088/1361-6471/ac7216. arXiv: 2203.05506 [hep-ph].

[83] A. D. Martin et al. "Parton distributions for the LHC". In: *Eur. Phys. J. C* 63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].

[84] Amanda Cooper-Sarkar, Philipp Mertsch, and Subir Sarkar. "The high energy neutrino cross-section in the Standard Model and its uncertainty". In: *JHEP* 08 (2011), p. 042. DOI: 10.1007/JHEP08(2011)042. arXiv: 1106.3723 [hep-ph].

[85] Choy-Heng Lai. "Charm Contribution to Neutrino Induced Production of Opposite Sign Dimuons". In: *Phys. Rev. D* 18 (1978), p. 1422. DOI: 10.1103/PhysRevD.18.1422.

[86] M. Klasen et al. "Global analyses of nuclear PDFs with heavy-quark and neutrino data". In: *PoS* ICHEP2022 (2022), p. 443. DOI: 10.22323/1.414.0443. arXiv: 2210.10284 [hep-ph].

[87] Sayipjamal Dulat et al. "New parton distribution functions from a global analysis of quantum chromodynamics". In: *Phys. Rev. D* 93.3 (2016), p. 033006. DOI: 10.1103/PhysRevD.93.033006. arXiv: 1506.07443 [hep-ph].

[88] Raj Gandhi et al. "Ultrahigh-energy neutrino interactions". In: *Astropart. Phys.* 5 (1996), pp. 81–110. DOI: 10.1016/0927-6505(96)00008-4. arXiv: hep-ph/9512364.

[89] Carlos A. Argüelles et al. "High-energy behavior of photon, neutrino, and proton cross sections". In: *Phys. Rev. D* 92.7 (2015), p. 074040. DOI: 10.1103/PhysRevD.92.074040. arXiv: 1504.06639 [hep-ph].

[90] Bei Zhou and John F. Beacom. "Dimuons in neutrino telescopes: New predictions and first search in IceCube". In: *Phys. Rev. D* 105.9 (2022), p. 093005. DOI: 10.1103/PhysRevD.105.093005. arXiv: 2110.02974 [hep-ph].

[91] Sheldon L. Glashow. "Resonant Scattering of Antineutrinos". In: *Phys. Rev.* 118 (1960), pp. 316–317. DOI: 10.1103/PhysRev.118.316.

[92] W. Czyz, G. C. Sheppey, and J. D. Walecka. "Neutrino production of lepton pairs through the point four-fermion interaction". In: *Nuovo Cim.* 34 (1964), pp. 404–435. DOI: 10.1007/BF02734586.

[93] J. Lovseth and M. Radomiski. "Kinematical distributions of neutrino-produced lepton triplets". In: *Phys. Rev. D* 3 (1971), pp. 2686–2706. DOI: 10.1103/PhysRevD.3.2686.

[94] K. Fujikawa. "The self-coupling of weak lepton currents in high-energy neutrino and muon reactions". In: *Annals Phys.* 68 (1971), pp. 102–162. DOI: 10.1016/0003-4916(71)90244-2.

[95] R. W. Brown et al. "Intermediate boson. iii. virtual-boson effects in neutrino trident production". In: *Phys. Rev. D* 6 (1972), pp. 3273–3292. DOI: 10.1103/PhysRevD.6.3273.

[96] R. Belusevic and J. Smith. "W - Z Interference in Neutrino - Nucleus Scattering". In: *Phys. Rev. D* 37 (1988), p. 2419. DOI: 10.1103/PhysRevD.37.2419.

[97] D. Seckel. "Neutrino photon reactions in astrophysics and cosmology". In: *Phys. Rev. Lett.* 80 (1998), pp. 900–903. DOI: 10.1103/PhysRevLett.80.900. arXiv: hep-ph/9709290.

[98] I. Alikhanov. "Hidden Glashow resonance in neutrino–nucleus collisions". In: *Phys. Lett. B* 756 (2016), pp. 247–253. DOI: 10.1016/j.physletb.2016.03.009. arXiv: 1503.08817 [hep-ph].

[99] Wolfgang Altmannshofer et al. "Neutrino Tridents at DUNE". In: *Phys. Rev. D* 100.11 (2019), p. 115029. DOI: 10.1103/PhysRevD.100.115029. arXiv: 1902.06765 [hep-ph].

[100] Wolfgang Altmannshofer et al. "Neutrino Trident Production: A Powerful Probe of New Physics with Neutrino Beams". In: *Phys. Rev. Lett.* 113 (2014), p. 091801. DOI: 10.1103/PhysRevLett.113.091801. arXiv: 1406.2332 [hep-ph].

[101] Gabriel Magill and Ryan Plestid. "Neutrino Trident Production at the Intensity Frontier". In: *Phys. Rev. D* 95.7 (2017), p. 073004. DOI: 10.1103/PhysRevD.95.073004. arXiv: 1612.05642 [hep-ph].

[102] Peter Ballett et al. "Neutrino trident scattering at near detectors". In: *Journal of High Energy Physics* 2019.1 (Jan. 2019). DOI: 10.1007/jhep01(2019)119. URL: https://doi.org/10.1007%2Fjhep01%282019%29119.

[103] Bei Zhou and John F. Beacom. "Neutrino-nucleus cross sections for W-boson and trident production". In: *Phys. Rev. D* 101.3 (2020), p. 036011. DOI: 10.1103/PhysRevD.101.036011. arXiv: 1910.08090 [hep-ph].

[104] G. Fricke et al. "Nuclear Ground State Charge Radii from Electromagnetic Interactions". In: *Atom. Data Nucl. Data Tabl.* 60 (1995), pp. 177–285. DOI: 10.1006/adnd.1995.1007.

[105] U. D. Jentschura and V. G. Serbo. "Nuclear form factor, validity of the equivalent photon approximation and Coulomb corrections to muon pair production in photon-nucleus and nucleus-nucleus collisions". In: *Eur. Phys. J. C* 64 (2009), pp. 309–317. DOI: 10.1140/epjc/s10052-009-1147-3. arXiv: 0908.3853 [hep-ph].

[106] Gabriel Magill and Ryan Plestid. "Probing new charged scalars with neutrino trident production". In: *Phys. Rev. D* 97.5 (2018), p. 055003. DOI: 10.1103/PhysRevD.97.055003. arXiv: 1710.08431 [hep-ph].

[107] Wolfgang Altmannshofer et al. "Quark flavor transitions in $L_\mu - L_\tau$ models". In: *Phys. Rev. D* 89 (2014), p. 095033. DOI: 10.1103/PhysRevD.89.095033. arXiv: 1403.1269 [hep-ph].

[108] Julian Heeck and Werner Rodejohann. "Gauged $L_\mu - L_\tau$ Symmetry at the Electroweak Scale". In: *Phys. Rev. D* 84 (2011), p. 075007. DOI: 10.1103/PhysRevD.84.075007. arXiv: 1107.5238 [hep-ph].

[109] H. N. Brown et al. "Precise measurement of the positive muon anomalous magnetic moment". In: *Phys. Rev. Lett.* 86 (2001), pp. 2227–2231. DOI: 10.1103/PhysRevLett.86.2227. arXiv: hep-ex/0102017.

[110] R Aaij et al. "Measurement of Form-Factor-Independent Observables in the Decay $B^0 \to K^{*0}\mu^+\mu^-$". In: *Phys. Rev. Lett.* 111 (2013), p. 191801. DOI: 10.1103/PhysRevLett.111.191801. arXiv: 1308.1707 [hep-ex].

[111] Ivone Albuquerque, Gustavo Burdman, and Z. Chacko. "Neutrino telescopes as a direct probe of supersymmetry breaking". In: *Phys. Rev. Lett.* 92 (2004), p. 221802. DOI: 10.1103/PhysRevLett.92.221802. arXiv: hep-ph/0312197.

[112] Ivone F. M. Albuquerque, Gustavo Burdman, and Z. Chacko. "Direct detection of supersymmetric particles in neutrino telescopes". In: *Phys. Rev. D* 75 (2007), p. 035006. DOI: 10.1103/PhysRevD.75.035006. arXiv: hep-ph/0605120.

[113] P. S. Bhupal Dev. "Signatures of Supersymmetry in Neutrino Telescopes". In: *Probing Particle Physics with Neutrino Telescopes*. 2020, pp. 317–352. DOI: 10.1142/9789813275027_0010. arXiv: 1906.02147 [hep-ph].

[114] Sandro Kopper. "Search for Neutrino Induced Double Tracks as an Exotic Physics Signature in IceCube". In: *PoS* ICRC2015 (2016), p. 1104. DOI: 10.22323/1.236.1104.

[115] Jan-Henrik Schmidt-Dencker et al. "Stau Search in IceCube". In: *PoS* ICRC2021 (2021), p. 1117. DOI: 10.22323/1.395.1117.

[116] Michael F. L'Annunziata. "7 - Cherenkov Radiation". In: *Radioactivity*. Ed. by Michael F. L'Annunziata. Amsterdam: Elsevier Science B.V., 2007, pp. 465–495. ISBN: 978-0-444-52715-8. DOI: https://doi.org/10.1016/B978-044452715-8.50010-4. URL: https://www.sciencedirect.com/science/article/pii/B9780444527158500104.

[117] I. M. Frank and I. E. Tamm. "Coherent visible radiation of fast electrons passing through matter". In: *Compt. Rend. Acad. Sci. URSS* 14.3 (1937), pp. 109–114. DOI: 10.3367/UFNr.0093.196710o.0388.

[118] C. Patrignani et al. "Review of Particle Physics". In: *Chin. Phys. C* 40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001. URL: https://pdg.lbl.gov/2022/AtomicNuclearProperties/HTML/water_ice.html.

[119] Donald E. Groom, Nikolai V. Mokhov, and Sergei I. Striganov. "Muon stopping power and range tables 10-MeV to 100-TeV". In: *Atom. Data Nucl. Data Tabl.* 78 (2001), pp. 183–356. DOI: 10.1006/adnd.2001.0861.

[120] Dmitry Chirkin and Wolfgang Rhode. "Muon Monte Carlo: A High-precision tool for muon propagation through matter". In: (July 2004). arXiv: hep-ph/0407075.

[121] E. Andres et al. "The AMANDA neutrino telescope: Principle of operation and first results". In: *Astropart. Phys.* 13 (2000), pp. 1–20. DOI: 10.1016/S0927-6505(99)00092-4. arXiv: astro-ph/9906203.

[122] T. Benson et al. "IceCube Enhanced Hot Water Drill functional description". In: *Annals of Glaciology* 55.68 (2014), pp. 105–114. DOI: 10.3189/2014AoG68A032.

[123] M. G. Aartsen et al. "The IceCube Neutrino Observatory: Instrumentation and Online Systems". In: *JINST* 12.03 (2017), P03012. DOI: 10.1088/1748-0221/12/03/P03012. arXiv: 1612.05093 [astro-ph.IM].

[124] Matsusada Precision. *Photonmultipler Tube*. URL: https://www.matsusada.com/application/ps/photomultiplier_tubes/.

[125] R. Abbasi et al. "Calibration and Characterization of the IceCube Photomultiplier Tube". In: *Nucl. Instrum. Meth. A* 618 (2010), pp. 139–152. DOI: 10.1016/j.nima.2010.03.102. arXiv: 1002.2442 [astro-ph.IM].

[126] R. Abbasi et al. "The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping". In: *Nucl. Instrum. Meth. A* 601 (2009), pp. 294–316. DOI: 10.1016/j.nima.2009.01.001. arXiv: 0810.4930 [physics.ins-det].

[127] Hamamatsu Corp. *R7081-02 Photomultiplier Tube Data*. URL: https://www.hamamatsu.com/eu/en/product/optical-sensors/pmt/pmt_tube-alone/head-on-type/R7081.html.

[128] Photonis. *Photomultiplier tube basics*. URL: https://psec.uchicago.edu/library/photomultipliers/Photonis_PMT_basics.pdf.

[129] M. G. Aartsen et al. "Energy Reconstruction Methods in the IceCube Neutrino Telescope". In: *JINST* 9 (2014), P03009. DOI: 10.1088/1748-0221/9/03/P03009. arXiv: 1311.4767 [physics.ins-det].

[130] M. G. Aartsen et al. "In-situ calibration of the single-photoelectron charge response of the IceCube photomultiplier tubes". In: *JINST* 15.06 (2020), P06032. DOI: 10.1088/1748-0221/15/06/P06032. arXiv: 2002.00997 [physics.ins-det].

[131] Nicholas Kulacz. "In Situ Measurement of the IceCube DOM Efficiency Factor Using Atmospheric Minimum Ionizing Muons". MA thesis. University of Alberta, 2019. URL: https://era.library.ualberta.ca/items/23cdc715-25d9-4f19-910b-c25e55e8a6b9.

[132] M. G. Aartsen et al. "Measurement of South Pole ice transparency with the IceCube LED calibration system". In: *Nucl. Instrum. Meth. A* 711 (2013), pp. 73–89. DOI: 10.1016/j.nima.2013.01.054. arXiv: 1301.5361 [astro-ph.IM].

[133] P. B. Price, K. Woschnagg, and D. Chirkin. "Age vs depth of glacial ice at South Pole". In: *Geophysical Research Letters* 27.14 (2000), pp. 2129–2132. DOI: https://doi.org/10.1029/2000GL011351. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000GL011351.

[134] Spencer Nicholas Gaelan Axani. "Sterile Neutrino Searches at the IceCube Neutrino Observatory". PhD thesis. MIT, 2019. arXiv: 2003.02796 [hep-ex].

[135] P. Buford Price et al. "Temperature profile for glacial ice at the South Pole: Implications for life in a nearby subglacial lake". In: *Proceedings of the National Academy of Sciences* 99.12 (2002), pp. 7844–7847. DOI: 10.1073/pnas.082238999. URL: https://www.pnas.org/doi/abs/10.1073/pnas.082238999.

[136] Dmitry Chirkin. "Evidence of optical anisotropy of the South Pole ice". In: *33rd International Cosmic Ray Conference*. 2013, p. 0580.

[137] M. G. Aartsen et al. "Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector". In: *Science* 342 (2013), p. 1242856. DOI: 10.1126/science.1242856. arXiv: 1311.5238 [astro-ph.HE].

[138] Logan Wille and Donglian Xu. "Astrophysical Tau Neutrino Identification with IceCube Waveforms". In: *PoS* ICRC2019 (2020), p. 1036. DOI: 10.22323/1.358.1036. arXiv: 1909.05162 [astro-ph.HE].

[139] R. Abbasi et al. "Detection of astrophysical tau neutrino candidates in IceCube". In: *Eur. Phys. J. C* 82.11 (2022), p. 1031. DOI: 10.1140/epjc/s10052-022-10795-y. arXiv: 2011.03561 [hep-ex].

[140] Cern. *Cosmic Rays: particles from outer space*. URL: https://home.cern/science/physics/cosmic-rays-particles-outer-space.

[141] Thomas K. Gaisser, Todor Stanev, and Serap Tilav. "Cosmic Ray Energy Spectrum from Measurements of Air Showers". In: *Front. Phys. (Beijing)* 8 (2013), pp. 748–758. DOI: 10.1007/s11467-013-0319-7. arXiv: 1303.3565 [astro-ph.HE].

[142] Felix Riehn et al. "The hadronic interaction model SIBYLL 2.3c and Feynman scaling". In: *PoS* ICRC2017 (2018), p. 301. DOI: 10.22323/1.301.0301. arXiv: 1709.07227 [hep-ph].

[143] Thomas K. Gaisser, Ralph Engel, and Elisa Resconi. *Cosmic Rays and Particle Physics*. 2nd ed. Cambridge University Press, 2016. DOI: 10.1017/CBO9781139192194.

[144] Atri Bhattacharya et al. "Prompt atmospheric neutrino fluxes: perturbative QCD models and nuclear effects". In: *JHEP* 11 (2016), p. 167. DOI: 10.1007/JHEP11(2016)167. arXiv: 1607.00193 [hep-ph].

[145] IceCube Collaboration. *nuFlux*. URL: https://docs.icecube.aq/nuflux/v2.0.0/fluxes.html.

[146] R. Abbasi et al. "Improved Characterization of the Astrophysical Muon–neutrino Flux with 9.5 Years of IceCube Data". In: *Astrophys. J.* 928.1 (2022), p. 50. DOI: 10.3847/1538-4357/ac4d29. arXiv: 2111.10299 [astro-ph.HE].

[147] D. Soldin. "Atmospheric Muons Measured with IceCube". In: *EPJ Web Conf.* 208 (2019). Ed. by B. Pattison et al., p. 08007. DOI: 10.1051/epjconf/201920808007. arXiv: 1811.03651 [astro-ph.HE].

[148] IceCube Collaboration. *IceCube Coordinate System*. URL: https://docs.icecube.aq/icetray/main/projects/dataclasses/coordinates.html.

[149] R. Abbasi et al. "LeptonInjector and LeptonWeighter: A neutrino event generator and weighter for neutrino observatories". In: *Comput. Phys. Commun.* 266 (2021), p. 108018. DOI: 10.1016/j.cpc.2021.108018. arXiv: 2012.10449 [physics.comp-ph].

[150] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[151] Stefan Höche and Stefan Prestel. "The midpoint between dipole and parton showers". In: *Eur. Phys. J. C* 75.9 (2015), p. 461. DOI: 10.1140/epjc/s10052-015-3684-2. arXiv: 1506.05057 [hep-ph].

[152] Stefan Prestel. *PYTHIA + DIRE worksheet*. URL: https://dire.gitlab.io/Documentation/worksheet1500.pdf.

[153] *Why MB/UE physics prefer LO PDFs*. URL: https://indico.cern.ch/event/266994/contributions/601393/attachments/476981/659966/mepdf.pdf.

[154] *impy - (hadronic) interaction models in python*. URL: https://github.com/impy-project/impy.

[155] Felix Riehn et al. "Hadronic interaction model Sibyll 2.3d and extensive air showers". In: *Phys. Rev. D* 102.6 (2020), p. 063002. DOI: 10.1103/PhysRevD.102.063002. arXiv: 1912.03300 [hep-ph].

[156] Gary Binder. "Measurements of the Flavor Composition and Inelasticity Distribution of High-EnergyNeutrino Interactions in IceCube". PhD thesis. University of California, Berkeley, 2017. URL: https://escholarship.org/uc/item/2bg3x5x1.

[157] D. Heck et al. "CORSIKA: A Monte Carlo code to simulate extensive air showers". In: (Feb. 1998).

[158] Stefan Roesler, Ralph Engel, and Johannes Ranft. "The Monte Carlo event generator DPMJET-III". In: *International Conference on Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications (MC 2000)*. Dec. 2000, pp. 1033–1038. DOI: 10.1007/978-3-642-18211-2_166. arXiv: hep-ph/0012252.

[159] Carlos A. Argüelles, Jordi Salvado, and Christopher N. Weaver. "nuSQuIDS: A toolbox for neutrino propagation". In: *Comput. Phys. Commun.* 277 (2022), p. 108346. DOI: 10.1016/j.cpc.2022.108346. arXiv: 2112.13804 [hep-ph].

[160] A. M. Dziewonski and D. L. Anderson. "Preliminary reference earth model". In: *Phys. Earth Planet. Interiors* 25 (1981), pp. 297–356. DOI: 10.1016/0031-9201(81)90046-7.

[161] Georges Aad et al. "Determination of the parton distribution functions of the proton using diverse ATLAS data from pp collisions at $\sqrt{s} = 7$, 8 and 13 TeV". In: *Eur. Phys. J. C* 82.5 (2022), p. 438. DOI: 10.1140/epjc/s10052-022-10217-z. arXiv: 2112.11266 [hep-ex].

[162] Alexander Belyaev, Neil D. Christensen, and Alexander Pukhov. "CalcHEP 3.4 for collider physics within and beyond the Standard Model". In: *Comput. Phys. Commun.* 184 (2013), pp. 1729–1769. DOI: 10.1016/j.cpc.2013.01.014. arXiv: 1207.6082 [hep-ph].

[163] Andy Buckley et al. "LHAPDF6: parton density access in the LHC precision era". In: *Eur. Phys. J. C* 75 (2015), p. 132. DOI: 10.1140/epjc/s10052-015-3318-8. arXiv: 1412.7420 [hep-ph].

[164] Carl Schmidt et al. "CT14QED parton distribution functions from isolated photon production in deep inelastic scattering". In: *Phys. Rev. D* 93.11 (2016), p. 114015. DOI: 10.1103/PhysRevD.93.114015. arXiv: 1509.02905 [hep-ph].

[165] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[166] Sourav Sarkar. *Trident Dimuon Generator*. URL: https://github.com/ssarkarbht/trident-generator.

[167] M. G. Aartsen et al. "eV-Scale Sterile Neutrino Search Using Eight Years of Atmospheric Muon Neutrino Data from the IceCube Neutrino Observatory". In: *Phys. Rev. Lett.* 125.14 (2020), p. 141801. DOI: 10.1103/PhysRevLett.125.141801. arXiv: 2005.12942 [hep-ex].

[168] C. A. Arguelles. *Lepton Weighter*. URL: https://github.com/icecube/LeptonWeighter.

[169] J. H. Koehne et al. "PROPOSAL: A tool for propagation of charged leptons". In: *Comput. Phys. Commun.* 184 (2013), pp. 2070–2090. DOI: 10.1016/j.cpc.2013.04.001.

[170] Mario Dunsch et al. "Recent Improvements for the Lepton Propagator PROPOSAL". In: *Comput. Phys. Commun.* 242 (2019), pp. 132–144. DOI: 10.1016/j.cpc.2019.03.021. arXiv: 1809.07740 [hep-ph].

[171] Sebastian Panknin. "Search for neutrino-induced cascade events in the IceCube detector". PhD thesis. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät I, 2011. DOI: http://dx.doi.org/10.18452/16397.

[172] Claudio Kopper. *CLSim*. URL: https://github.com/claudiok/clsim.

[173] Christopher Weaver. "Evidence for Astrophysical Muon Neutrinos from the Northern Sky". PhD thesis. University of Wisconsin–Madison, 2015. URL: https://docushare.icecube.wisc.edu/dsweb/Get/Document-73829/weaver_thesis_2015.pdf.

[174] M. G. Aartsen et al. "Improvement in Fast Particle Track Reconstruction with Robust Statistics". In: *Nucl. Instrum. Meth. A* 736 (2014), pp. 143–149. DOI: 10.1016/j.nima.2013.10.074. arXiv: 1308.5501 [astro-ph.IM].

[175] Chris Weaver. *Topological Splitter*. URL: https://docs.icecube.aq/icetray/main/projects/TopologicalSplitter/index.html.

[176] Jakob van Santen. *MuEx Energy Reconstruction*. URL: https://docs.icecube.aq/icetray/main/projects/mue/muex.html.

[177] Mike Daily et al. "Self-Driving Cars". In: *Computer* 50.12 (2017), pp. 18–23. DOI: 10.1109/MC.2017.4451204.

[178] Zhong-Qiu Zhao et al. "Object Detection with Deep Learning: A Review". In: *CoRR* abs/1807.05511 (2018). arXiv: 1807.05511. URL: http://arxiv.org/abs/1807.05511.

[179] Daniel W. Otter, Julian R. Medina, and Jugal K. Kalita. "A Survey of the Usages of Deep Learning in Natural Language Processing". In: *CoRR* abs/1807.10854 (2018). arXiv: 1807.10854. URL: http://arxiv.org/abs/1807.10854.

[180] Fernanda Psihas et al. "A Review on Machine Learning for Neutrino Experiments". In: *Int. J. Mod. Phys. A* 35.33 (2020), p. 2043005. DOI: 10.1142/S0217751X20430058. arXiv: 2008.01242 [physics.comp-ph].

[181] Kim Albertsson et al. "Machine Learning in High Energy Physics Community White Paper". In: *J. Phys. Conf. Ser.* 1085.2 (2018), p. 022008. DOI: 10.1088/1742-6596/1085/2/022008. arXiv: 1807.02876 [physics.comp-ph].

[182] Pushpalatha C. Bhat. "Multivariate Analysis Methods in Particle Physics". In: *Ann. Rev. Nucl. Part. Sci.* 61 (2011), pp. 281–309. DOI: 10.1146/annurev.nucl.012809.104427.

[183] Nicholas Choma et al. "Graph Neural Networks for IceCube Signal Classification". In: *CoRR* abs/1809.06166 (2018). arXiv: 1809.06166. URL: http://arxiv.org/abs/1809.06166.

[184] Mirco Huennefeld et al. "Combining Maximum-Likelihood with Deep Learning for Event Reconstruction in IceCube". In: *PoS* ICRC2021 (2021), p. 1065. DOI: 10.22323/1.395.1065. arXiv: 2107.12110 [astro-ph.HE].

[185] R. Abbasi et al. "A Convolutional Neural Network based Cascade Reconstruction for the IceCube Neutrino Observatory". In: *JINST* 16 (2021), P07041. DOI: 10.1088/1748-0221/16/07/P07041. arXiv: 2101.11589 [hep-ex].

[186] R. Abbasi et al. "Graph Neural Networks for low-energy event classification & reconstruction in IceCube". In: *JINST* 17.11 (2022), P11003. DOI: 10.1088/1748-0221/17/11/P11003. arXiv: 2209.03042 [hep-ex].

[187] Vijay Prakash Dwivedi et al. "Benchmarking Graph Neural Networks". In: *CoRR* abs/2003.00982 (2020). arXiv: 2003.00982. URL: https://arxiv.org/abs/2003.00982.

[188] Thomas N. Kipf and Max Welling. "Semi-Supervised Classification with Graph Convolutional Networks". In: *CoRR* abs/1609.02907 (2016). arXiv: 1609.02907. URL: http://arxiv.org/abs/1609.02907.

[189] Petar Veličković et al. "Graph Attention Networks". In: (2017). DOI: 10.48550/ARXIV.1710.10903. URL: https://arxiv.org/abs/1710.10903.

[190] Zonghan Wu et al. "A Comprehensive Survey on Graph Neural Networks". In: *CoRR* abs/1901.00596 (2019). arXiv: 1901.00596. URL: http://arxiv.org/abs/1901.00596.

[191] Rex Ying et al. *Hierarchical Graph Representation Learning with Differentiable Pooling*. 2018. DOI: 10.48550/ARXIV.1806.08804. URL: https://arxiv.org/abs/1806.08804.

[192] Martin Wolf. *Common Variables*. URL: https://docs.icecube.aq/icetray/main/projects/CommonVariables/track_characteristics/index.html#commonvariables-track-characteristics.

[193] *PyTorch Geometric*. URL: https://pytorch-geometric.readthedocs.io/en/latest/.

[194] *PyTorch*. URL: https://pytorch.org/.

[195] *DenseGCNConv*. URL: https://pytorch-geometric.readthedocs.io/en/latest/_modules/torch_geometric/nn/dense/dense_gcn_conv.html#DenseGCNConv.

[196] *DenseDiffPool*. URL: https://pytorch-geometric.readthedocs.io/en/latest/_modules/torch_geometric/nn/dense/diff_pool.html#dense_diff_pool.

[197] Sergey Ioffe and Christian Szegedy. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift". In: *CoRR* abs/1502.03167 (2015). arXiv: 1502.03167. URL: http://arxiv.org/abs/1502.03167.

[198] Nitish Srivastava et al. "Dropout: a simple way to prevent neural networks from overfitting." In: *Journal of Machine Learning Research* 15.1 (2014), pp. 1929–1958. URL: http://www.cs.toronto.edu/~rsalakhu/papers/srivastava14a.pdf.

[199] Tianxiang Zhao, Xiang Zhang, and Suhang Wang. "GraphSMOTE: Imbalanced Node Classification on Graphs with Graph Neural Networks". In: *CoRR* abs/2103.08826 (2021). arXiv: 2103.08826. URL: https://arxiv.org/abs/2103.08826.

[200] Shuhao Shi et al. "Boosting-GNN: Boosting Algorithm for Graph Networks on Imbalanced Node Classification". In: *Frontiers in Neurorobotics* 15 (2021). ISSN: 1662-5218. DOI: 10.3389/fnbot.2021.775688. URL: https://www.frontiersin.org/articles/10.3389/fnbot.2021.775688.

[201] Hongwei Yong et al. "Momentum Batch Normalization for Deep Learning with Small Batch Size". In: *Computer Vision – ECCV 2020*. Ed. by Andrea Vedaldi et al. Cham: Springer International Publishing, 2020, pp. 224–240.

[202] Augustin-Louis Cauchy. *Gradient Descent*. URL: https://en.wikipedia.org/wiki/Gradient_descent.

[203] Williams R. Rumelhart D. Hinton G. "Learning representations by back-propagating errors". In: *Nature* 323.1 (1986), pp. 533–536. URL: https://www.nature.com/articles/323533a0.

[204] Diederik P. Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization". In: *arXiv e-prints*, arXiv:1412.6980 (Dec. 2014), arXiv:1412.6980. arXiv: 1412.6980 [cs.LG].

[205] PyTorch. *Reduce on Plateau*. URL: https://pytorch.org/docs/stable/generated/torch.optim.lr_scheduler.ReduceLROnPlateau.html.

[206] M. G. Aartsen et al. "Efficient propagation of systematic uncertainties from calibration to analysis with the SnowStorm method in IceCube". In: *JCAP* 10 (2019), p. 048. DOI: 10.1088/1475-7516/2019/10/048. arXiv: 1909.01530 [hep-ex].

[207] Martin Rongen. "Measuring the optical properties of IceCube drill holes". In: *EPJ Web Conf.* 116 (2016). Ed. by A. Capone et al., p. 06011. DOI: 10.1051/epjconf/201611606011.

[208] Anatoli Fedynitch, Julia Becker Tjus, and Paolo Desiati. "Influence of hadronic interaction models and the cosmic ray spectrum on the high energy atmospheric muon and neutrino flux". In: *Physical Review D* 86.11 (Dec. 2012). DOI: 10.1103/physrevd.86.114024. URL: https://doi.org/10.1103%2Fphysrevd.86.114024.

[209] Paolo Montini. "Cosmic ray physics with ARGO-YBJ". In: *Nucl. Part. Phys. Proc.* 279-281 (2016). Ed. by Gabriella Cataldi, Ivan De Mitri, and Daniele Martello, pp. 7–14. DOI: 10.1016/j.nuclphysbps.2016.10.003. arXiv: 1608.01251 [astro-ph.HE].

[210] M. Baak et al. "HistFitter software framework for statistical data analysis". In: *Eur. Phys. J. C* 75 (2015), p. 153. DOI: 10.1140/epjc/s10052-015-3327-7. arXiv: 1410.1280 [hep-ex].

[211] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[212] Aya Ishihara. "The IceCube Upgrade - Design and Science Goals". In: *PoS* ICRC2019 (2021), p. 1031. DOI: 10.22323/1.358.1031. arXiv: 1908.09441 [astro-ph.HE].

[213] Andrea Falcone. "Deep underground neutrino experiment: DUNE". In: *Nucl. Instrum. Meth. A* 1041 (2022), p. 167217. DOI: 10.1016/j.nima.2022.167217.

[214] Marilisa De Serio. "Neutrino Physics with the SHiP experiment at CERN". In: *PoS* ICHEP2020 (2021), p. 123. DOI: 10.22323/1.390.0123.

[215] Sergey Shirobokov. "Search for New Physics with the SHiP experiment at CERN". In: *PoS* ICHEP2020 (2021), p. 282. DOI: 10.22323/1.390.0282.

[216] Elisa Resconi. "The Pacific Ocean Neutrino Experiment". In: *PoS* ICRC2021 (2022), p. 024. DOI: 10.22323/1.395.0024. arXiv: 2111.13133 [astro-ph.IM].

[217] Matteo Agostini et al. "The Pacific Ocean Neutrino Experiment". In: *Nature Astron.* 4.10 (2020), pp. 913–915. DOI: 10.1038/s41550-020-1182-4. arXiv: 2005.09493 [astro-ph.HE].

[218] Matteo Sanguineti. "Status and physics results of the KM3NeT experiment". In: *Nuovo Cim. C* 46.1 (2023), p. 4. DOI: 10.1393/ncc/i2023-23004-3.

[219] Felix Kling and Laurence J. Nevay. "Forward neutrino fluxes at the LHC". In: *Phys. Rev. D* 104.11 (2021), p. 113008. DOI: 10.1103/PhysRevD.104.113008. arXiv: 2105.08270 [hep-ph].

[220] Henso Abreu et al. "Detecting and Studying High-Energy Collider Neutrinos with FASER at the LHC". In: *Eur. Phys. J. C* 80.1 (2020), p. 61. DOI: 10.1140/epjc/s10052-020-7631-5. arXiv: 1908.02310 [hep-ex].

Appendix A: Features for the Classification Model

A.1 Feature Correlations

The selected DOM- and event-level input features for the dimuon classifiers are defined based on the potential dimuon event signatures in the detector as described in Section 5.2.5. The performance of ML algorithms relies on finding correlations among these features that can discriminate between signal and background events. Therefore, an investigation of the correlations among the features is an important step to avoid the use of any redundant features in the model, i.e. the features that are fully correlated (or anti-correlated) for both classes (single- and double-track) will not improve the model performance. For this purpose, the Spearman correlation matrix is computed for all combinations of two features (in the DOM- and event-level model input space separately) for the classifier. Spearman correlation is a statistical measure to quantify the degree to which two variables can be described using a monotonic function. Given a paired dataset of two features comprising n samples, denoted as $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$, the correlation coefficient between feature x and y is defined as,

$$C_{xy} = \frac{\sum_{i=1}^{n} \left(R(x_i) - R(\bar{x})\right) \left(R(y_i) - R(\bar{y})\right)}{\sqrt{\sum_{i=1}^{n} \left(R(x_i) - R(\bar{x})\right)^2} \sqrt{\sum_{i=1}^{n} \left(R(y_i) - R(\bar{y})\right)^2}} \quad , \tag{A.1}$$

where $R(x_i)$ and $R(y_i)$ are the relative positions (referred to as 'ranks') of the variables x_i and y_i in the ordered set of $\{x_1, x_2, ..., x_n\}$ and $\{y_1, y_2, ..., y_n\}$, respectively. $R(\bar{x})$ and $R(\bar{y})$ denote the ranks of the average values $\bar{x}$ and $\bar{y}$ for the two features, respectively. The values of the Spearman correlation C_{xy} range from -1 to +1. If the two variables x and y are associated with a perfectly monotonic function (decreasing or increasing), the Spearman correlation coefficient is measured to be either -1 (for a

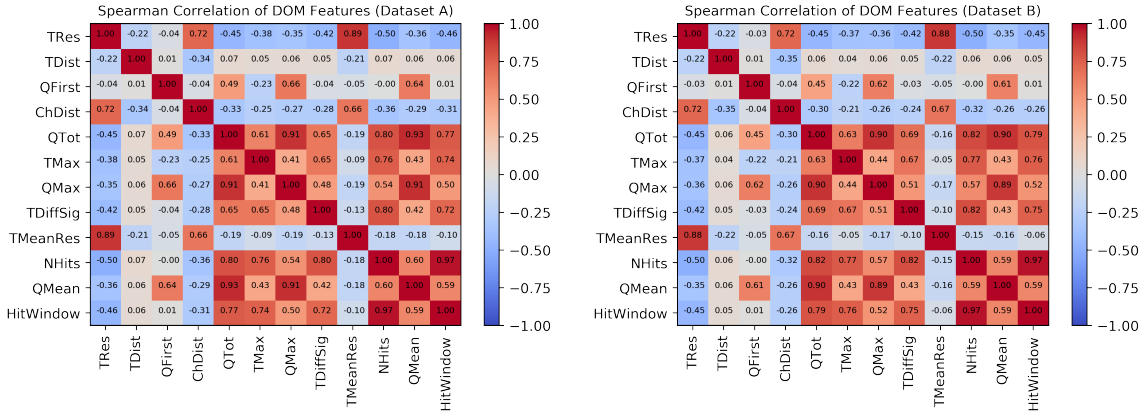


Figure A.1: Spearman correlation matrix among the DOM-level features in the training dataset. The left plot shows the correlation in DatasetA. The right plot shows the correlation matrix for DatasetB.

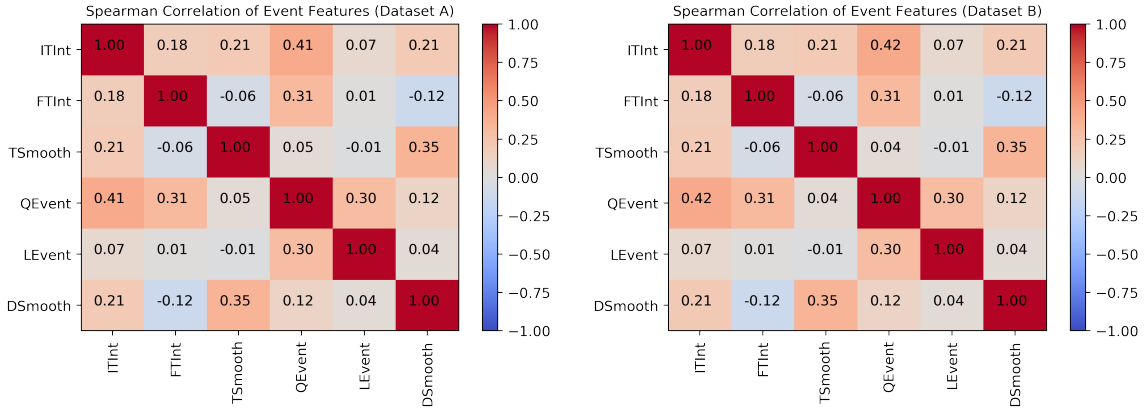
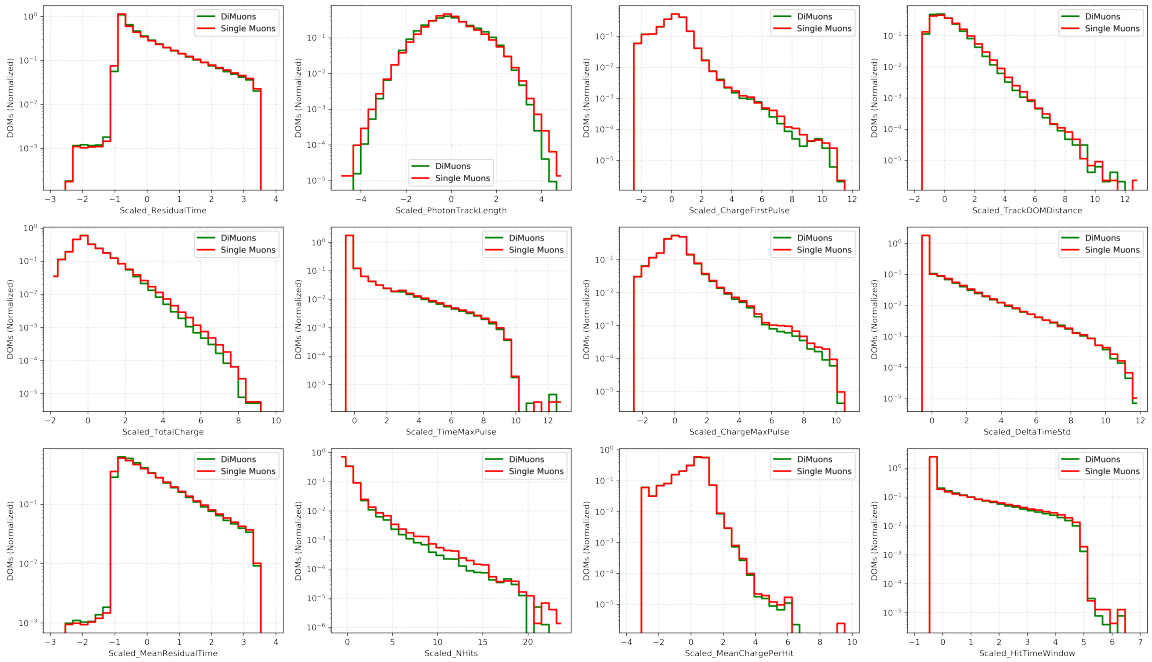


Figure A.2: Spearman correlation among the event-level features in the training dataset. The left plot shows the correlation in DatasetA. The right plot shows the correlation matrix for DatasetB.

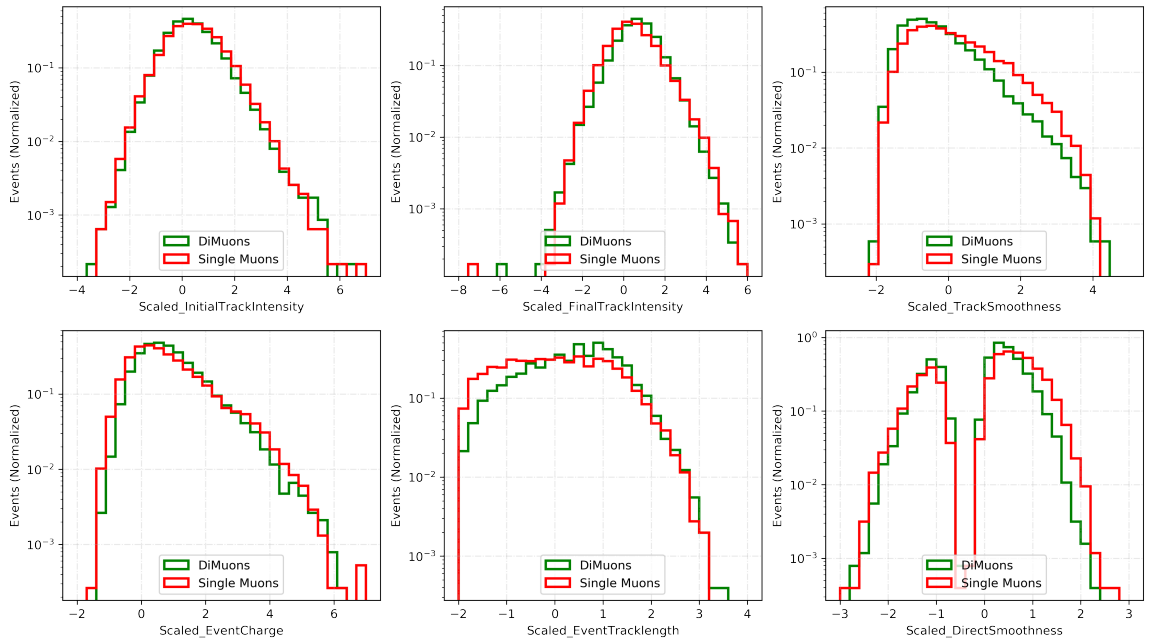
monotonically decreasing function) or +1 (for a monotonically increasing function). The correlation matrix for the input features in the training datasets (DatasetA and DatasetB) are shown in Figure A.1 for DOM-level features and Figure A.2 for event-level features.

A.2 Feature Distribution in Training Datasets

In addition to the feature correlations discussed in the previous section, the distributions of individual features are compared between the signal dimuon and background single muon events and are shown in Figure A.3 for DatasetA and in Figure A.4 for DatasetB. The plots show the raw distributions (i.e. without event weights) of the scaled features (scaling method is discussed in Section 5.2.5) with very little separation power on individual features.

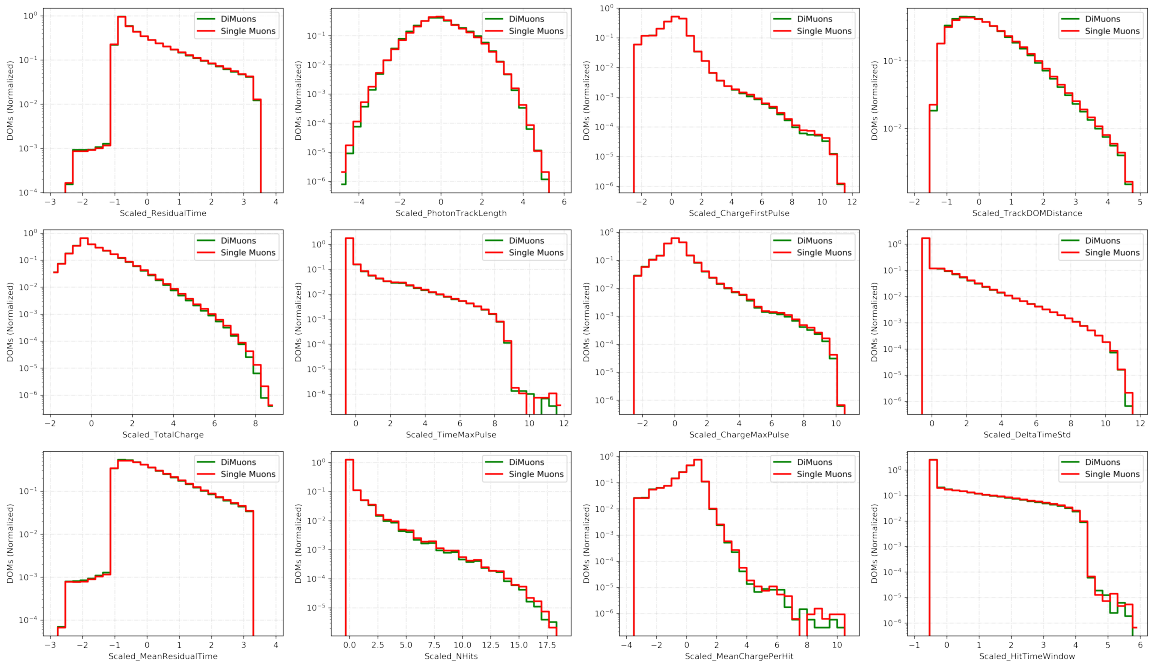


(a)

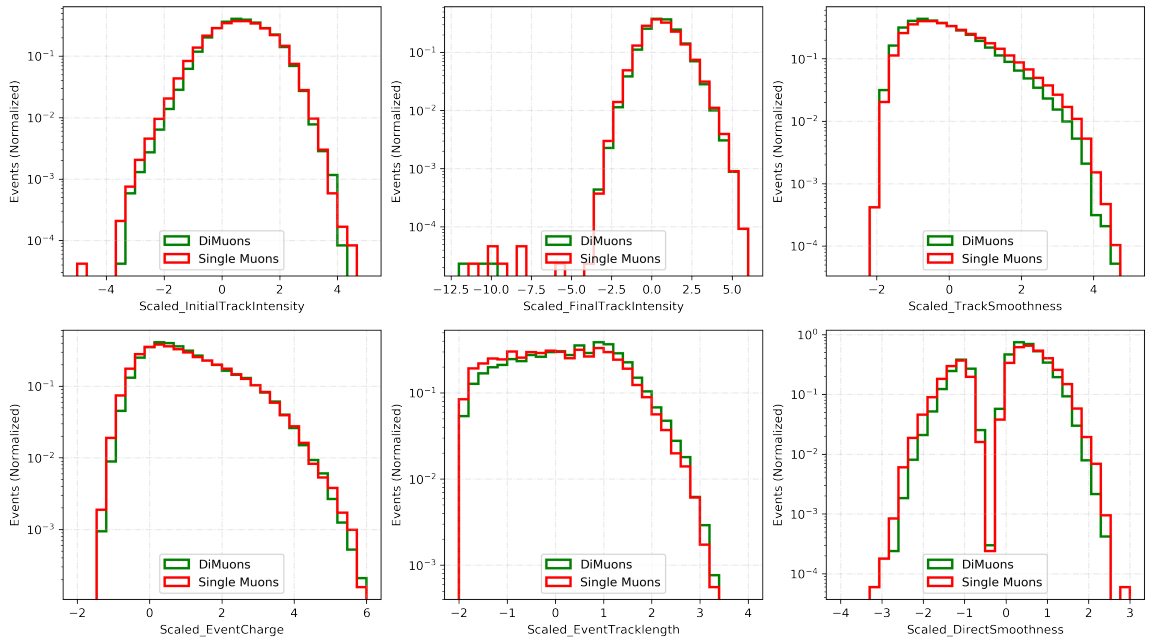


(b)

Figure A.3: Scaled feature distributions of DatasetA for DOM-level (a) and event-level (b) features.



(a)



(b)

Figure A.4: Scaled feature distributions of DatasetB for DOM-level (a) and event-level (b) features.

Appendix B: Geometry Resampling

Class A dimuons are only $\sim 3\%$ of all dimuon events triggered in the detector (see Section 5.1). It requires a huge computation power to generate an MC dataset with large enough class A dimuon events required for the model training dataset. In this appendix, we describe the technique used to resample the interaction geometry of the already simulated events to create more class A events while keeping the original random sampling distributions as intact as possible without requiring much computation power.

Let us take a track-like particle with starting position $\overrightarrow{P}_0 = (x_0, y_0, z_0)$ and direction unit vector $\overrightarrow{v}$ ($\|\overrightarrow{v}\| = 1$). The goal of the resampling is to move the position (interaction vertex) $\overrightarrow{P}_0$ to a new position $\overrightarrow{Q}_0$ along the direction vector $\overrightarrow{v}$ so as to meet the criteria of Class A events (minimum track segment $\geq$ 200 m, maximum track separation $\geq$ 25 m in the detector volume). If the track has the entry and exit points in the detector volume (assuming a perfect cylinder of radius = 500 m and height = 1000 m) as $\overrightarrow{X}_i$ and $\overrightarrow{X}_f$, respectively, they can be expressed using the parametric form of a line as,

$$\begin{aligned} \overrightarrow{X}_i &= \overrightarrow{P}_0 + t_i \cdot \overrightarrow{v} \\ \overrightarrow{X}_f &= \overrightarrow{P}_0 + t_f \cdot \overrightarrow{v} \quad , \end{aligned} \tag{B.1}$$

where t_i, t_f are the corresponding length parameters of the line equation. For a dimuon event, we choose the trailing muon, i.e. the muon with lower energy as the track particle discussed above. With respect to the starting position $\overrightarrow{P}_0$, let us denote t_{25} as the length parameter where the separation between the tracks is 25 m and t_{end} as the length parameter indicating the end point of the trailing muon. The track and the interaction position with all the parameters discussed above are shown in

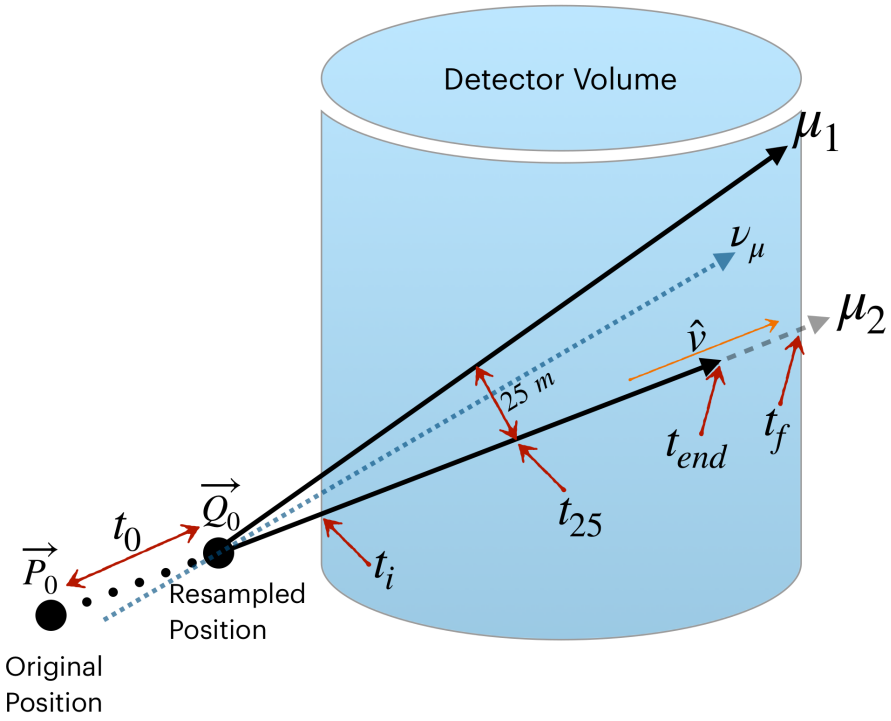


Figure B.1: An illustration of the parameters used in the geometry resampling technique.

Figure B.1. In order to achieve Class A criteria (for the events that did not originally meet the criteria), the following condition applies to the event,

$$t_{end} \geq t_{25} \geq 200 \text{ m} \quad . \tag{B.2}$$

Therefore, the bounds on the new interaction vertex position, $[\overrightarrow{Q}_0^{min}, \overrightarrow{Q}_0^{max}]$ give the bounds on the entry and exit points (constant positions) for the event to be Class A as,

$$\begin{aligned} \overrightarrow{X}_i &= \overrightarrow{Q}_0^{min} + (t_{end} - 200) \cdot \overrightarrow{v} \\ \overrightarrow{X}_f &= \overrightarrow{Q}_0^{max} + t_{25} \cdot \overrightarrow{v} \quad . \end{aligned} \tag{B.3}$$

Combining Equation (B.1) and Equation (B.3), we get

$$\begin{aligned} \overrightarrow{Q}_0^{min} &= \overrightarrow{P}_0 + (200 + t_i - t_{end}) \cdot \overrightarrow{v} \\ \overrightarrow{Q}_0^{max} &= \overrightarrow{P}_0 + (t_f - t_{25}) \cdot \overrightarrow{v} \quad . \end{aligned} \tag{B.4}$$

If we express the new interaction vertex $\overrightarrow{Q}_0$ in terms of the original vertex position

$\overrightarrow{P}_0$ as,

$$\overrightarrow{Q}_0 = \overrightarrow{P}_0 + t_0 \cdot \overrightarrow{v} \quad , \tag{B.5}$$

a sampling of the length parameter t_0 to translate the vertex position has the bound,

$$\overrightarrow{Q}_0^{min} \leq \overrightarrow{Q}_0 \leq \overrightarrow{Q}_0^{max}$$
$$(200 + t_i - t_{end}) \leq t_0 \leq (t_f - t_{25}) \quad . \tag{B.6}$$

Therefore, we can randomly sample t_0 from the bounds to get a new vertex position where the new resampled event can be injected along the same direction to transform the event into a guaranteed Class A event.

One caveat of the technique is that the formulation holds for $(X_f - X_i) \geq 200$ m, i.e. the maximum possible line segment within the detector volume is ≥ 200 m. Some tracks that enter and exit the detector near the corners of the cylindrical volume, i.e. tracks with high closest approach distance value, might not have more than 200 m track segment available for applying the resampling method. However, we can move those events to a new closest approach distance by sampling uniformly from a smaller injection radius disk around the centre of the detector and then applying the resampling method.

The following steps summarize the resampling algorithm used to generate the dimuon in the training sample with high statistics Class A events:

- Check if the minimum energy muon has a track length ≥ 200 m.
- If not, the event kinematics cannot produce class A criteria. Discard the event and skip the next steps.
- Check if the event already satisfies Class A criteria.
- If not, check if the line segment of the minimum energy muon inside the detector volume is ≥ 200 m.
- If not, sample a new closest approach distance from the smaller injection radius.
- Calculate the length parameters $t_i,\ t_f,\ t_{25},\ t_{end}$.
- Sample t_0 uniformly from the bound in Equation (B.6).

- Compute and overwrite the new interaction vertex, $\overrightarrow{Q}_0 = \overrightarrow{P}_0 + t_0 \cdot \overrightarrow{v}$.

Application of the resampling method results in approximately 100 times more class A dimuons in the MC set compared to its original simulation, and the normalized event distributions as a function of the maximum track separation and minimum track segment compared between the original and resampled simulation set are shown in Figure B.2. Therefore, the resampled MC dataset provides a high statistics class A dimuon events for the training of the dimuon classifier without simulating a computationally expensive event generation process.

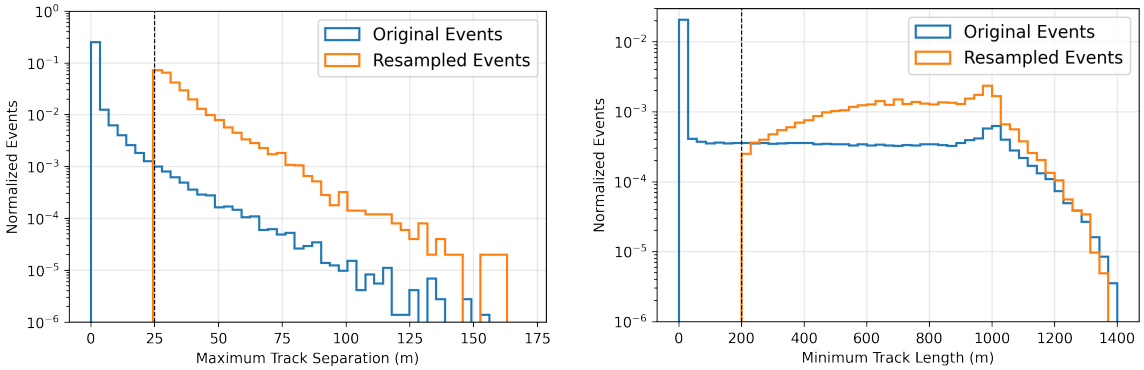


Figure B.2: Comparison of the original simulation and the translated events after applying the geometry resampling, as a function of the maximum track separation (left) and minimum track segment (right). The vertical dashed line shows the class A dimuon cut criteria.

Appendix C: Visualization of the Observed Events

The observed events in SR1 and SR2 are visualized in Figure C.1, Figure C.2, Figure C.3, and Figure C.4. The reconstructed single track (MuEx) is shown in the green solid line with an illustrated blue Cherenkov cone along the track direction. The DOM hits from the cleaned pulses (TTPulses) are shown in coloured blobs. The colour gradient in the DOM hits represents the time information with blue being the earliest hits and red being the latest hits. The size of the blobs represents the total charge of the detected light in each DOM. Additional event information is also shown in the plots in text.

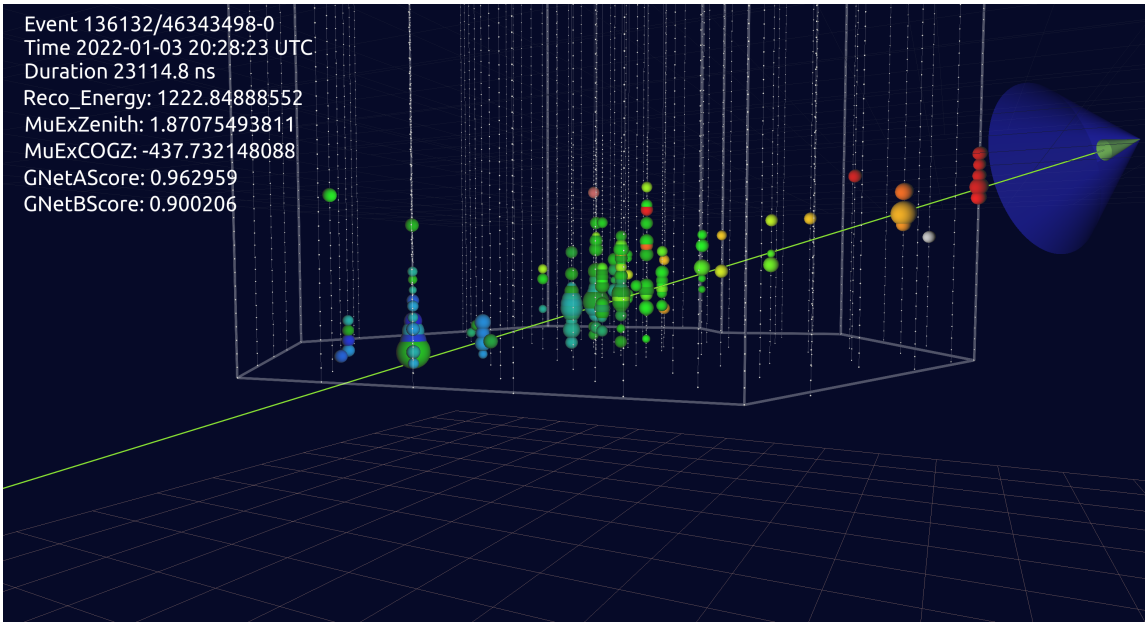


Figure C.1: Event 1 observed in both SR1 and SR2.

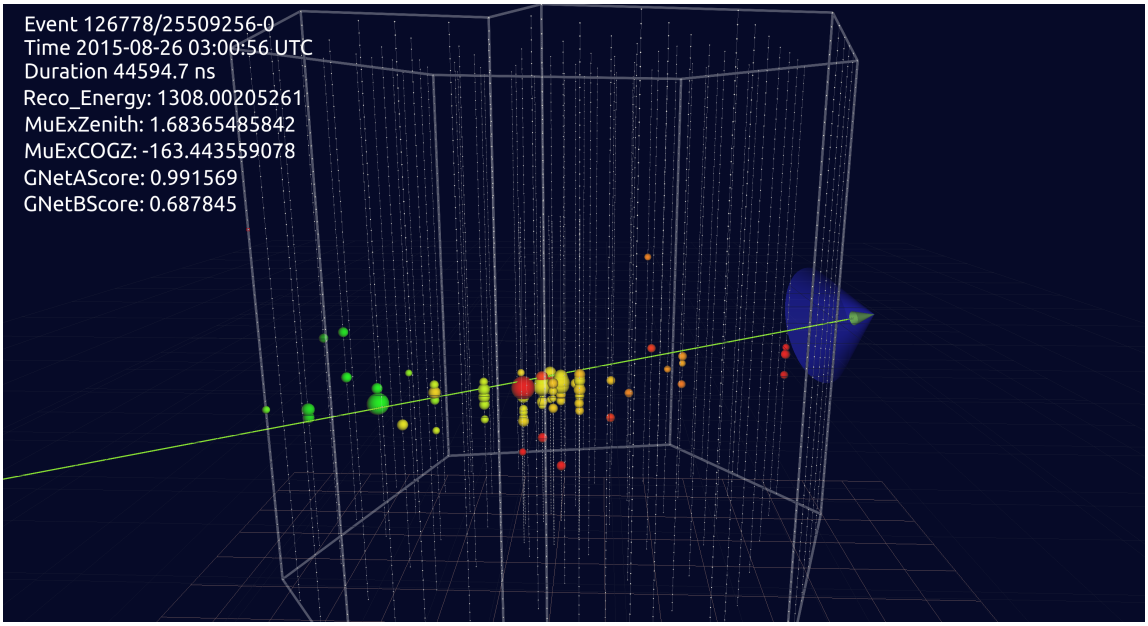


Figure C.2: Event 2 observed in SR2.

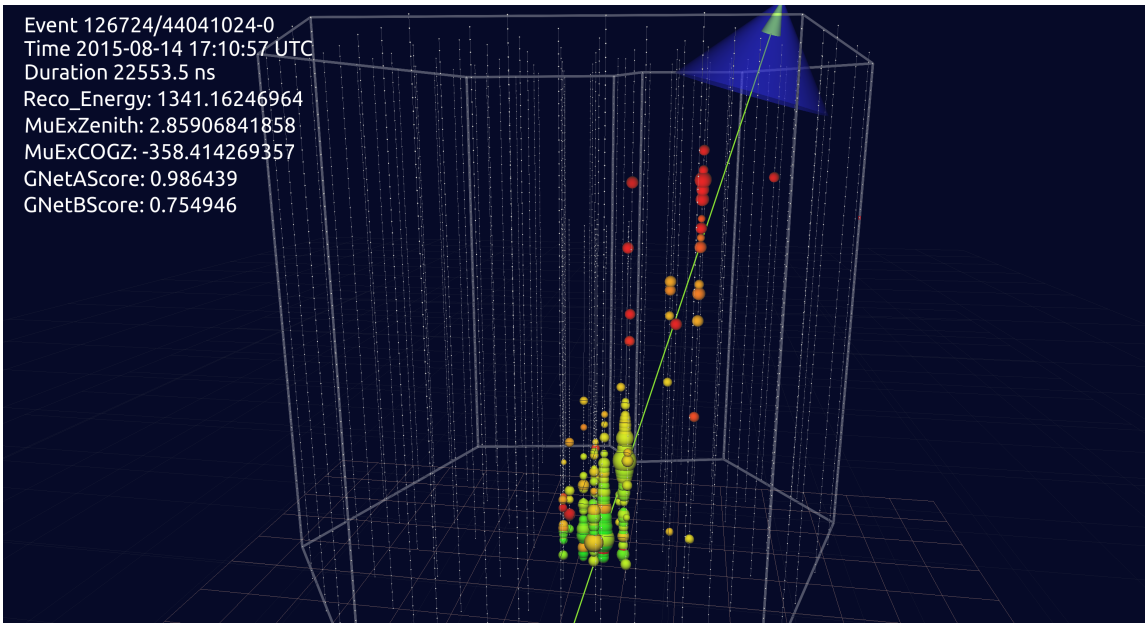


Figure C.3: Event 3 observed in SR2.

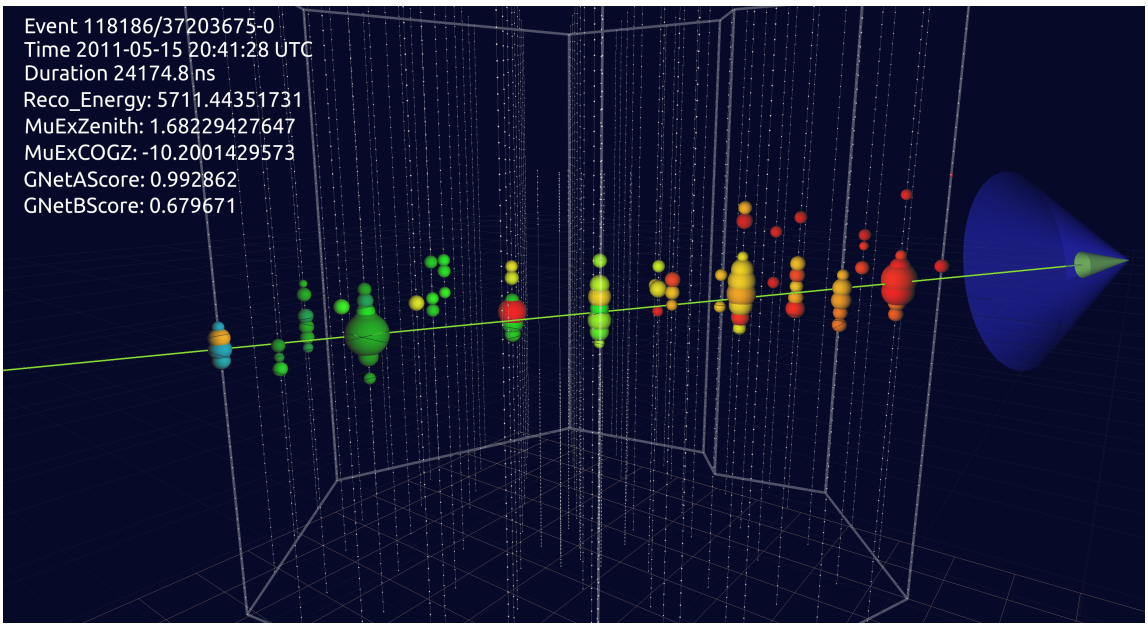


Figure C.4: Event 4 observed in SR2.