Notes

1 Introduction

Deep beneath the Amundsen-Scott station on the geographic South Pole, lurks a telescope in the ice. This observatory hunts for some of the most elusive particles in our universe, hoping they may answer foundational questions about nature, and glean insight into new physics that the Genevan collider—or its potential successors—may never be able to.

The IceCube Neutrino Observatory was built between 2005 and 2010, but its heritage stretches back several decades. Spanning one cubic kilometer of Antarctic ice, the telescope consists of some 60 strings, hot water drilled into the ice, with each string containing a number of photomultiplier tubes, capable of detecting photons originating from processes involving neutrinos—the aforementioned elusive particles.

Measurements of certain neutrino properties may help us understand phenomena such as why the universe is matter-dominated[1] and dark matter[2]. IceCube is also capable of searching for point-like sources[3], and serves in the SuperNova Early Warning System[4].

But one thing is the ability to observe derived photons in Antarctic ice; another is reconstruction of this light in order to infer properties of the originating particles. To date, classical methods, such as PegLeg[5] and Retro Reco[*] have mostly been used for this purpose, but the burgeoning proliferation of the use of machine learning (ML) for scientific discovery[6] inspires the development of new methods for this cause.

Whereas classical methods require a hypothesis and thus preset logic, ML "changes these [problems] from logic problems to statistical problems"[†]. That is to say, given a large enough—or rather, a representative enough—dataset, an ML algorithm may find statistical similarities in disparate neutrino events sharing similar underlying properties, giving the network an ability to predict these same properties for previously unseen

[*] https://github.com/IceCubeOpenSource/retro

[†] https://www.ben-evans.com/benedictevans/2019/9/6/face-recognition

data. As a “machine learning model is almost like a pure function at inference time”[‡], owing to the universal approximation theorems, it should be possible to construct an ML reconstruction algorithm that is both faster, preciser, more extensible, captures a wider spectrum and is more future-proof than the current methods.

The intention of this thesis is to explore that proposition, specifically for up-going muon neutrinos in the sub-teraelectronvolt range.

This problem is selected as it is the focus of the oscillations group at IceCube, studying neutrino oscillations. The current best reconstruction algorithm for this case, Retro Reco, can take minutes to reconstruct a neutrino, while an ML-based approach can be expected to reconstruct several thousand each second. An additional focus will be the IceCube upgrade[7], which will install additional improved detectors, and for which there is no current functioning reconstruction algorithm.

1.1 Flow

To weave a common thread through this document, and ensure the reader is informed about the big picture throughout, the following section presents the flow of the thesis.

The story of an interaction in IceCube from physics to analysis proceeds at a very high level as follows: The universe is created with mechanisms that only allow a certain chirality of neutrinos to interact with forces other than gravity (section 2.2 on page 9). At some point an interaction will create such a particle (section 2.3 on page 9), which oscillates between different flavors (section 2.4 on page 9). As chance allows some neutrinos to interact in or around the instrumented ice, new particles are created that in turn send out a specific form of light (section 2.5 on page 9) which the photomultiplier tubes (section 3.2 on page 14) can detect. Today each in-ice digital optical module contains one photomultiplier tube. With the IceCube upgrade this changes (section 3.8 on page 14), and new optical modules will be fitted with additional photomultiplier tubes which should help with determining the direction of neutrinos. The detector is set up with certain triggers in differing levels (section 3.3 on page 14), to ensure that not just any light is recorded and saved to disk. Even further processing is done, as the bandwidth from pole to America is quite low, and so care must be taken

[‡]Brennan Saeta, Software Engineer - TensorFlow, https://www.swiftbysundell.com/podcast/58/

in the choice of what to actually send down the wires for analysis. When the data arrives for analysis, different level of cleaning takes place to root out noise (section 3.5 on page 14). This pipeline also attempts to classify event types, using decision trees, until finally algorithms—none of them machine learning based—reconstruct neutrino properties of interest such as energy and direction (section 3.6 on page 14), and the events at all levels of cleaning are saved in an IceCube proprietary format, called I3 (section 3.7 on page 14).

With this part done, we turn towards machine learning generalities. First the basics of machine learning in general (section 4.1 on page 15), then deep learning in particular (section 4.2 on page 15). Special attention is given to temporal convolutional neural networks (section 4.3 on page 15), as this is the architecture used. Because loss functions are such an important part of any deep learning project, a section is devoted to this concept as well (section 4.4 on page 15), and for the most part, before data can be meaningfully trained in a neural network, some form of data transformation is required (section 4.5 on page 15).

The last part of the thesis concerns the work by the author. The data format provided by IceCube is not machine learning friendly, so a new solution fixing those shortcomings was built (section 5.3 on page 16). So that the reader is on the same page as the author, plots and metrics used for comparison are explained. The opponent algorithm is detailed (section 6.2 on page 17), before the structure of the data itself used in training is explained (section 6.3 on page 17) and the algorithm laid out (section 6.4 on page 17).

2 Particle physics

Neutrinos are some of the universe's most interesting and mysterious particles, and offer one of the better paths to researching some still unsolved problems in physics. The Large Hadron Collider nailed down the final missing piece of the model, the Higgs boson, but questions still remain; and some of them concern neutrinos.

The masses of the neutrinos themselves is a mystery, but sterile neutrinos may also be a dark matter candidate and the possible CP-violating phase in the PMNS matrix may help explain baryon asymmetry. Needless to say, the more we discover about this little particle, the better we know the universe.

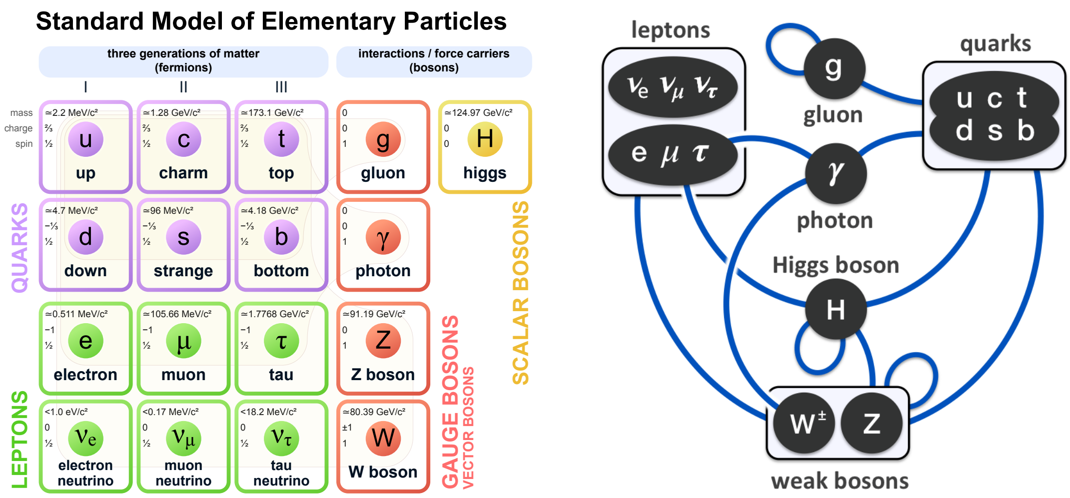
The concept of neutrino oscillations has the ability to constrain the previously mentioned CP-violating phase of the neutrino mixing matrix[the_t2k_collaboration_constraint_2020] and so the better we are able to measure this, the better the constraint will be.

This chapter will seek to outline the physics behind the interactions that are captured in the IceCube detector, including the concept of Cherenkov radiation, which is of critical importance to the process.

First, we shortly need to touch on what The Standard Model of particle physics is.

2.1 The Standard Model

The Standard Model (SM) of particle physics is the foundation of modern particle physics research. It was patched together during the 20th century by a significant amount of scientific collaboration, and validated through country-level spending.



(a) The fermions and bosons of The Standard Model of particle physics. Image from (ref)

(b) The possible interactions in the SM between fermions and bosons. Image from (ref)

Figure 2.1: Three simple graphs

The theory itself is a gauge quantum field theory, the mathematics of which is not particularly relevant to this thesis. It describes three out of the four known fundamental interactions, of which the so-called weak one, as we shall see, is of importance to the neutrino.

The SM describes two types of particles: fermions and bosons. The distinguishing feature between the two sets is the statistics they obey; whereas the fermions follow Fermi-Dirac statistics—as a consequence of their half-integer spin—the bosons adhere to Bose-Einstein statistics. Matter in the universe consists of fermions, while the fundamental force mediation is carried out by the bosons.

But what is the Standard Model trying to describe? Well, every macroscopic event in our lives (and, indeed, the universe) consists of interactions: the tweeting of a bird, the drizzle falling on cold autumn day, even the keyboard clacking sounds of a graduate student furiously typing up his thesis; all interactions. But these are macroscopic, made from relations at a lower level still. The bird, spurred on by the invisible and unending power of evolution, forces air through membranes to produce flattering sounds that attract mates. Thermodynamics and gravity create conditions in which small water drops fall towards the ground, and the graduate student is compelled by coffee and his supervisor to produce some result or other. But this still only explains the phenomenon on a macroscopic level. It is possible to keep peeling away abstraction layers, and hopefully arrive at some lowest level that can describe the fundamental forces that conspire to create the universe and life as we know it. This is the job of the Standard Model, the most succesful mathematical formulation of fundamentality we have.

The progression from Galilei to Higgs is a remarkable journey. It never ceases to amaze the author that Dirac hit upon his equation (eq. (2.1)) and discovered that the structure was that of Clifford algebra, discovered 50 years before as a purely mathematical concept. The use of group theory, seemingly far removed from everyday life, permeates the Standard Model, and in fact the applicability of abstract mathematics in particle physics is a bewildering thing, contemplated by Wigner[8]. Of course this thesis is not concerned with this philosophical aspect of physics, but the preceding sentences serve as an introduction to the Dirac equation, which will be used as a mathematical introduction to the concepts that need explaining in this chapter.

The Dirac equation reads as

$$\left(\mathrm{i}\gamma^\mu \partial_\mu - m\right)\psi = 0, \tag{2.1}$$

with $\mu \in \{0, 1, 2, 3\}$ and where γ^μ represents the gamma matrices and ∂ is the partial differential operator. The gamma matrices are defined by their anti-commutation relation

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu} I_4, \tag{2.2}$$

where η is the Minkowski metric with signature $(+ - - -)$ and I_4 is the 4 dimensional identity matrix. Equation (2.2) generates the previously mentioned Clifford algebra, and is on its own enough to define the system.

Equation (2.1) came about out of needing to reconcile special relativity and quantum mechanics; in short, the wave equation needed to be Lorentz invariant and conform to Einstein's energy-momentum relation, $E^2 = (pc)^2 + \left(mc^2\right)^2$.

Maybe expound a bit on Dirac's journey and a little more on the equation

Todo: Fermi's golden rule and invariant matrix element

The Dirac equation leads to the form of the matrix elements in the different possible interactions, for example the electromagnetic scattering of an electron and tau (seen

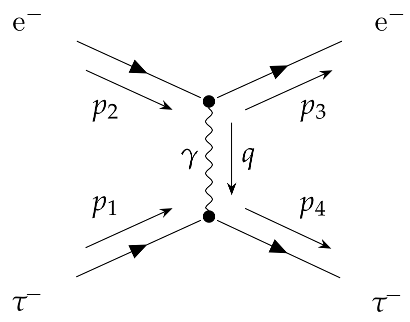


Figure 2.2: Feynman diagram showing the t-channel scattering of an electron and a tau.

in fig. 2.2),

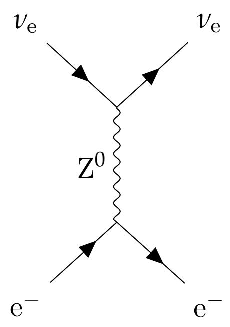
$$\mathcal{M} = -e^2 \left[\bar{u}_e(p_3)\gamma^\mu u_e(p_1)\right] \frac{g_{\mu\nu}}{q^2} \left[\bar{u}_\tau(p_4)\gamma^\nu u_\tau(p_2)\right], \tag{2.3}$$

where $u_\ell, \ell \in \{e, \tau\}$ represents the spinor of a particle, $\bar{u}_\ell, \ell \in \{e, \tau\}$ the anti-spinor and the momenta $\{q, p_1, p_2, p_3, p_4\}$ correspond to fig. 2.2. In this case, an electron and a tau lepton scatter by exchanging a photon, with momentum $q = p_1 - p_3$, which changes the momentum of both the electron and the tau. The Standard Model describes these fundamental interactions of the universe, embodied in Feynman diagrams such as fig. 2.2. Equation (2.3) shows the prominence of the gamma matrices in the fundamental interactions, which is of some importance in the weak interaction, to be discussed in section 2.2 on the following page.

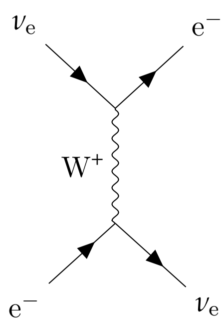
Figure 2.1a on page 6 shows the standard representation of The Standard Model.

A little more about Feynman diagrams

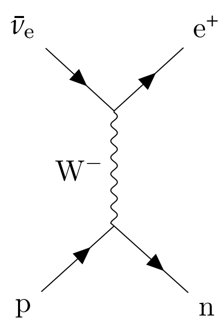
The different forces and associated Feynman diagrams



(a) Neutral current



(b) Charged current



(c) Inverse beta decay

Figure 2.3: Feynman diagrams of neutral- and charged current processes

2.2 The weak interaction

2.3 Neutrino interactions

2.4 Neutrino oscillations

$$\begin{bmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{bmatrix} = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix}$$

2.5 Cherenkov radiation

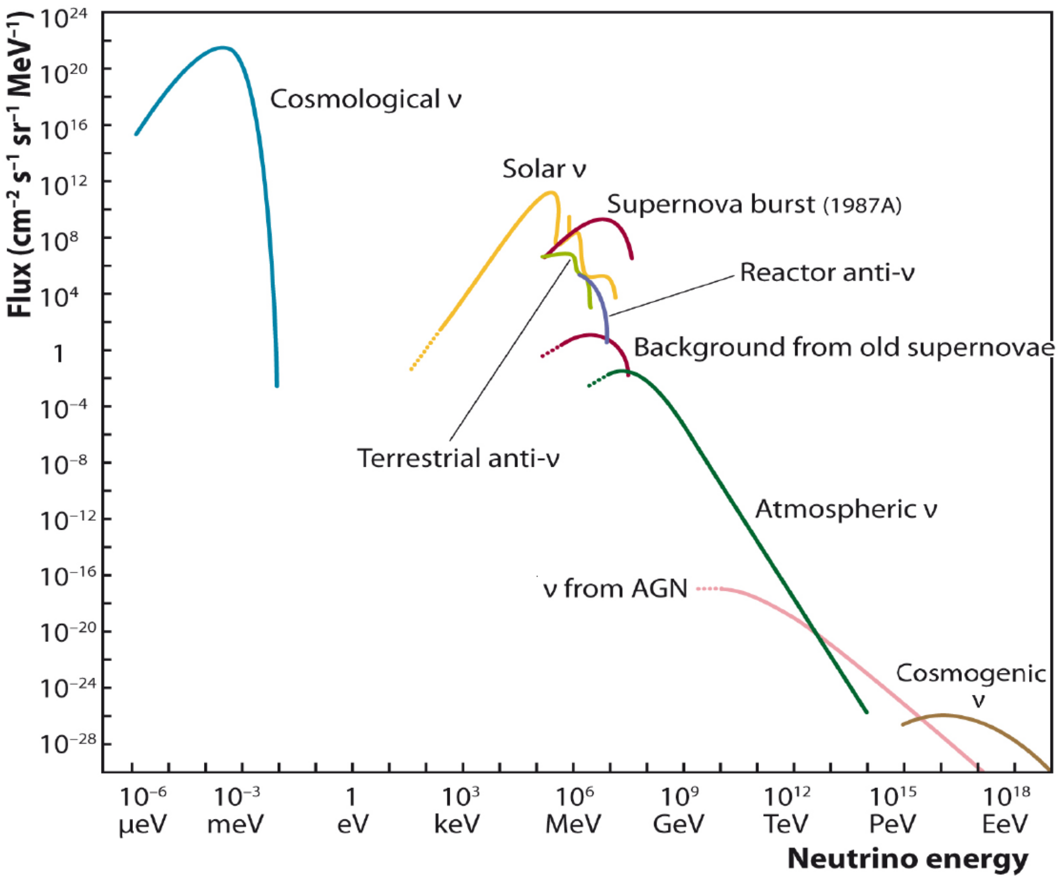


Figure 2.4: Neutrino spectrum. Image from**spiering_towards_2012**.

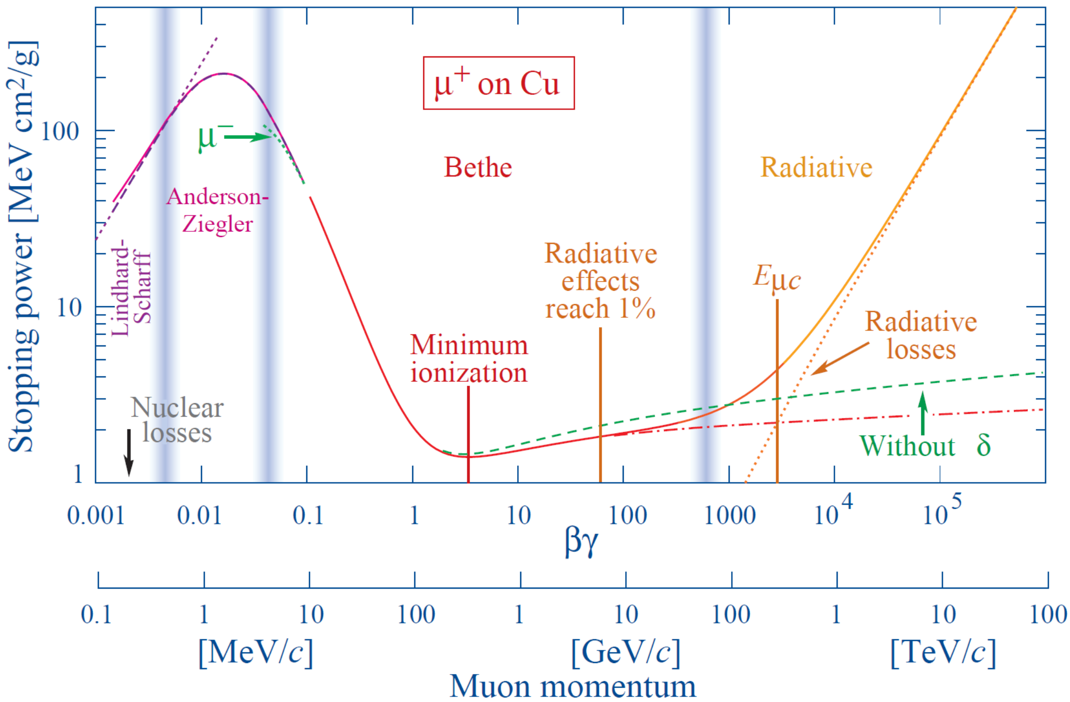


Figure 2.5: Stopping power

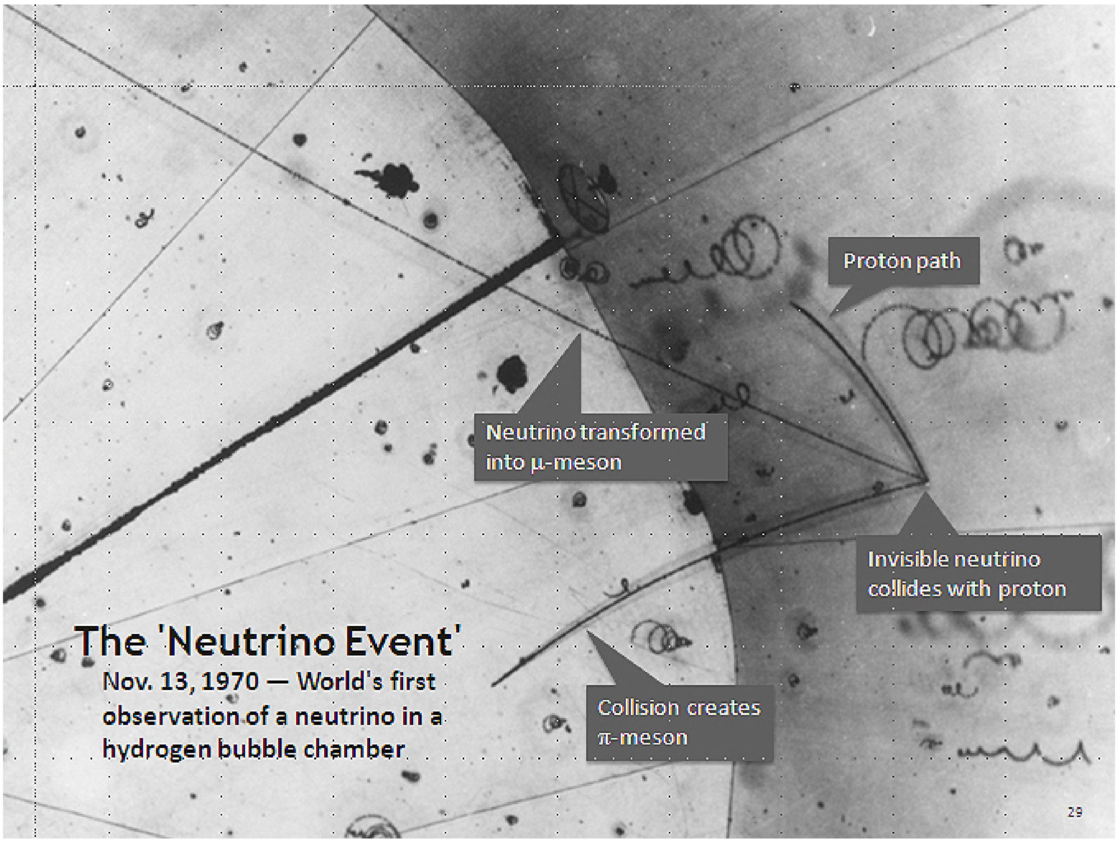


Figure 2.6: First neutrino seen in a bubble chamber. Image from **argonne_national_laboratory_first_1970**.

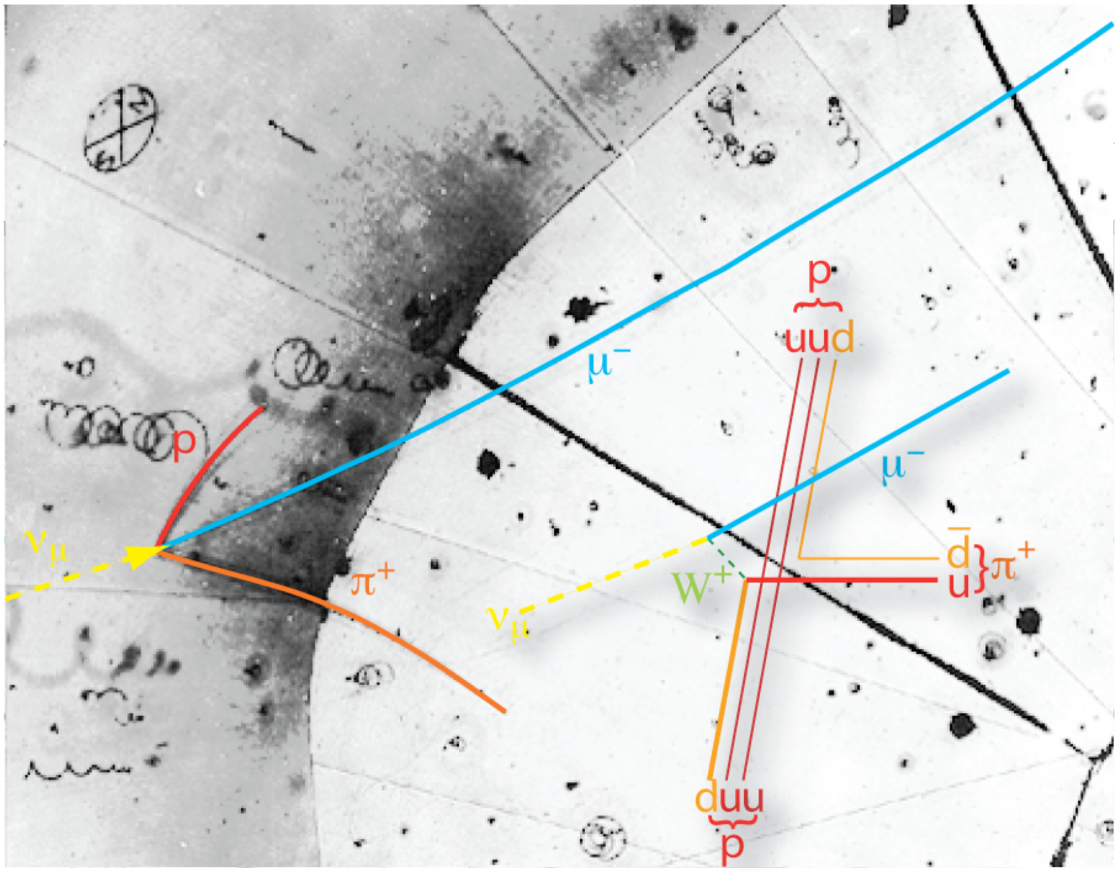


Figure 2.7: Bubble chamber neutrino event with annotation. Image from[argonne_national_laboratory_neutrino_2012].

3 IceCube

3.1 Detector

3.2 DOMs

3.3 Triggers

3.4 Post processing

3.5 Levels

3.6 Current reconstruction methods

3.7 I3 file format

3.8 IceCube upgrade

4 Machine learning

4.1 ML basics

4.2 Deep learning

4.3 Temporal convolutional neural networks

4.4 Loss functions

4.5 Data transformation

5 Data

5.1 Monte Carlo simulation

5.2 OscNext

5.3 SQLite

6 Design

6.1 Plots and metrics

6.2 Opponents

6.3 Data structure

6.4 Algorithm

7 Results

8 Summary and outlook

Bibliography

[1] H Nunokawa, S Parke, and JWF Valle, "Cp violation and neutrino oscillations," Progress in Particle and Nuclear … (2008) (cited on page 2).

[2] S Baur, "Dark matter searches with the icecube upgrade," arXiv preprint arXiv:1908.08236 (2019) (cited on page 2).

[3] MG Aartsen, M Ackermann, J Adams, and JA Aguilar…, "Time-integrated neutrino source searches with 10 years of icecube data," Physical review … (2020) (cited on page 2).

[4] L Köpke, "Supernova neutrino detection with icecube," arXiv preprint arXiv:1106.6225 (2011) (cited on page 2).

[5] M Leuermann, M Krämer, and CHV Wiebusch, "Testing the neutrino mass ordering with icecube deepcore," s3.cern.ch (2018) (cited on page 2).

[6] M Raghu and E Schmidt, "A survey of deep learning for scientific discovery," arXiv preprint arXiv:2003.11755 (2020) (cited on page 2).

[7] A Ishihara, "The icecube upgrade–design and science goals," arXiv preprint arXiv:1908.09441 (2019) (cited on page 3).

[8] Eugene P. Wigner, "The unreasonable effectiveness of mathematics in the natural sciences. richard courant lecture in mathematical sciences delivered at new york university, may 11, 1959," Communications on Pure and Applied Mathematics **13**, 1–14 (1960) (cited on page 7).