

Notes

Contents

1	Introduction	4
1.1	Flow	5
2	Particle physics	7
2.1	The Standard Model	7
2.2	The weak interaction	10
2.3	Neutrino interactions	10
2.4	Neutrino oscillations	10
2.5	Cherenkov radiation	11
3	IceCube	15
3.1	Detector	15
3.2	DOMs	15
3.3	Triggers	15
3.4	Post processing	15
3.5	Levels	15
3.6	Current reconstruction methods	15
3.7	I3 file format	15
3.8	IceCube upgrade	15
4	Machine learning	16
4.1	ML basics	16
4.2	Deep learning	16
4.3	Temporal convolutional neural networks	16
4.4	Loss functions	16
4.5	Data transformation	16
5	Data	17
5.1	Monte Carlo simulation	17
5.2	OscNext	17

5.3	SQLite	17
6	Design	18
6.1	Plots and metrics	18
6.2	Opponents	18
6.3	Data structure	18
6.4	Algorithm	18
7	Results	19
8	Summary and outlook	20

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-dominated¹ and dark matter². IceCube is also capable of searching for point-like sources³, and serves in the SuperNova Early Warning System⁴.

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 PegLeg⁵ and Retro Reco^{*} have mostly been used for this purpose, but the burgeoning proliferation of the use of machine learning (ML) for scientific discovery⁶ 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

^{*}<https://github.com/IceCubeOpenSource/retro>

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

unseen 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 upgrade⁷, 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 10). At some point an interaction will create such a particle (section 2.3 on page 10), which oscillates between different flavors (section 2.4 on page 10). 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 11) which the photomultiplier tubes (section 3.2 on page 15) can detect. Today each in-ice digital optical module contains one photomultiplier tube. With the IceCube upgrade this changes (section 3.8 on page 15), 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 15), 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 in the choice

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

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 15). 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 15), and the events at all levels of cleaning are saved in an IceCube proprietary format, called I3 (section 3.7 on page 15).

With this part done, we turn towards machine learning generalities. First the basics of machine learning in general (section 4.1 on page 16), then deep learning in particular (section 4.2 on page 16). Special attention is given to temporal convolutional neural networks (section 4.3 on page 16), 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 16), 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 16).

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 17). 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 18), before the structure of the data itself used in training is explained (section 6.3 on page 18) and the algorithm laid out (section 6.4 on page 18).

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, while sterile neutrinos is a dark matter candidate and the possible CP-violating phase in the PMNS matrix can help explain baryon asymmetry. Needless to say, the more we discover about this little particle, the better we know the universe.

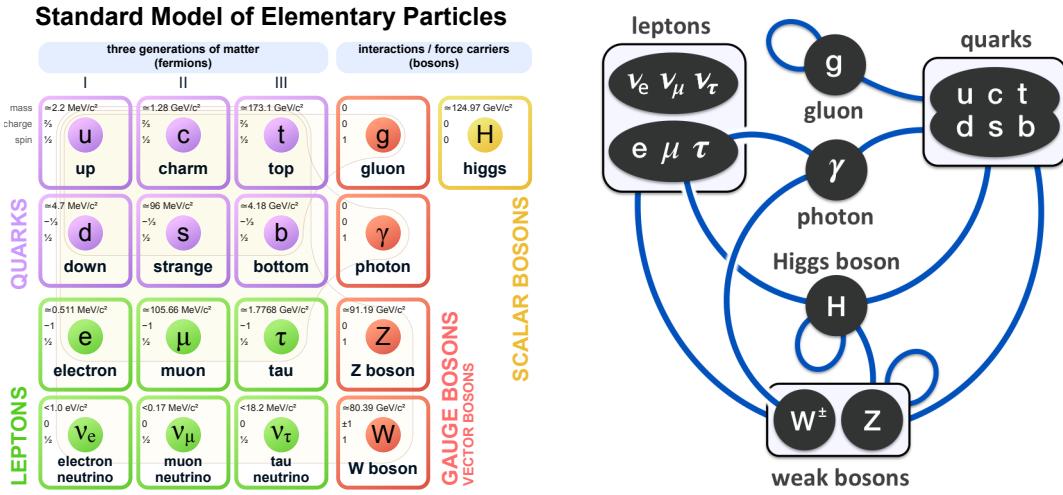
The concept of neutrino oscillations has the ability to constrain the previously mentioned CP-violating phase of the neutrino mixing matrix⁸ and so the better we are able to measure this, the better the constraint will be.

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 scientific collaboration, and tested through inter-country spending on giant particle accelerators.



(a) The fermions and bosons of The Standard Model (b) The possible interactions in the SM between particle physics. 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, and it describes three out of the four known fundamental interactions. Of these, the so-called weak one is of importance to the neutrino, while being a critical clue to why some particles acquire mass.

The SM describes two types of particles: fermions and bosons, as seen in fig. 2.1a. 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 forces are mediated by the bosons.

The fermions are divided between the quarks and the leptons, differentiated by their participation in the strong interaction, carried by the massless gluon. This boson only acts on quarks, and has the peculiar behaviour of creating new particles when quarks are pulled apart, resulting in the fact that it is only possible to observe hadrons (a composite particle made up of two or more quarks); this phenomenon, “color confinement”, is the progenitor of jets in collider experiments such as the LHC.

Both the charged leptons and the quarks interact electromagnetically and weakly, forces mediated by the photon and the W^\pm & Z^0 bosons respectively, but the leptons include a second sub-species that only interacts weakly: the neutral leptons, known as the neutrinos; see fig. 2.1b on the preceding page. Because they do not participate in anything but the weak force, neutrinos hardly ever interact, and are thus difficult (but not impossible!) to observe. An MeV neutrino has a cross section on the order of $1 \times 10^{-44} \text{ cm}^2$, and a beam of such particles must travel through a light year of lead before half of them is blocked.

The experimental particle physicist uses Feynman diagrams to describe potential interactions between particles, depending on Fermi's golden rule, the Lorentz-invariant matrix element and for fermions the Dirac equation.

Paul Dirac and others before him searched for an extension of the Schrödinger equation, which incorporated relativity. Requiring adherence to the Einstein energy-momentum relation, Dirac found the matrix equation

$$(i\gamma^\mu \partial_\mu - m) \psi = 0, \quad (2.1)$$

which requires a four-component wave function (not incidentally explaining spin and anti-particles). This is the governing equation of all fermions

Experiments generally proceed by calculating expected particle rates, and using either accelerators or passive detectors to compare reality to theory using Fermi's golden rule,

$$\Gamma_{fi} = 2\pi |T_{fi}|^2 \rho(E_i), \quad (2.2)$$

describing transition rates from an initial state $|i\rangle$ to a final state $|f\rangle$. The equation depends on the transition matrix T_{fi} from the expansion of the interaction with the perturbation Hamiltonian, ρ is the energy-dependent density of states. Equation (2.2) can be used to calculate the differential cross section

$$\frac{d\sigma}{d\Omega^*} = \frac{1}{64\pi^2 s} \frac{p_f^*}{p_i^*} |\mathcal{M}_{fi}|^2, \quad (2.3)$$

an expression of quantum mechanical probabilities for an interaction to happen, valid in the centre-of-mass frame (COM), indicated by *. s is the squared COM energy, p the particle momenta, and Ω the solid angle a particle scatters into.

The SM thus provides observables, as long as one can calculate the Lorentz invariant matrix element \mathcal{M} , which in the fermion case can be done using eq. (2.1). This process

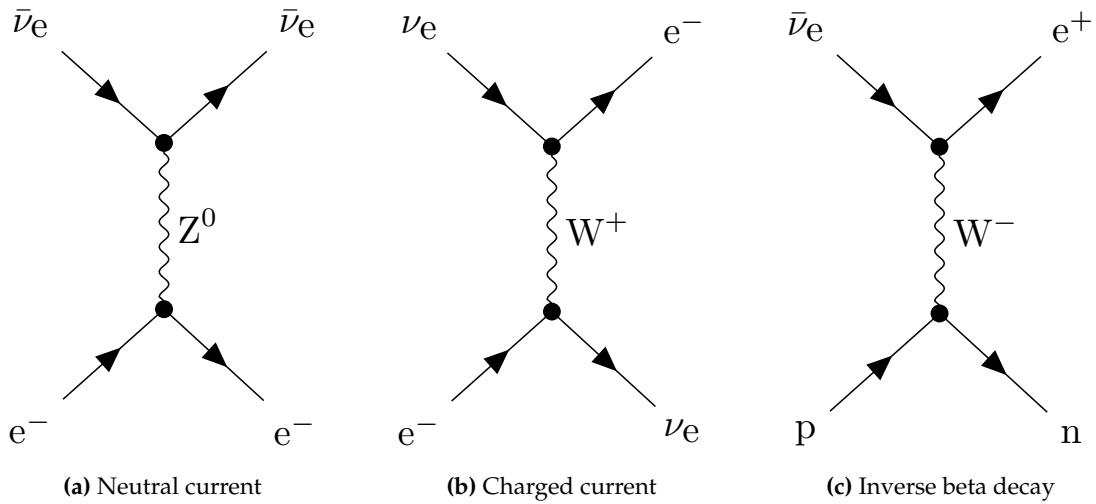


Figure 2.2: Feynman diagrams of neutral- and charged current processes

is much simplified by Feynman diagrams, which allows the theorist to draw processes and using Feynman rules stitch together a matrix element describing the interaction. Examples of Feynman diagrams can be seen in fig. 2.2.

2.2 The weak interaction

2.3 Neutrino interactions

2.4 Neutrino oscillations

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

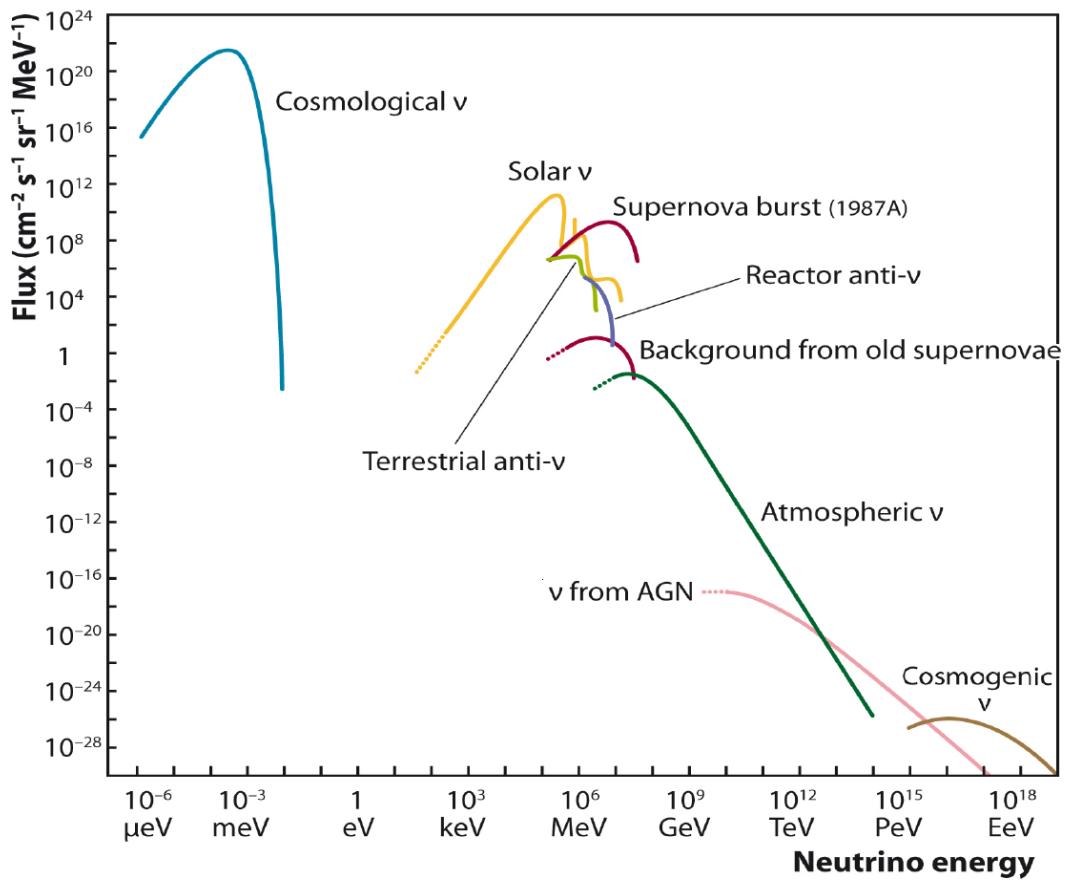


Figure 2.3: Neutrino spectrum. Image from [spiering_towards_2012](#).

2.5 Cherenkov radiation

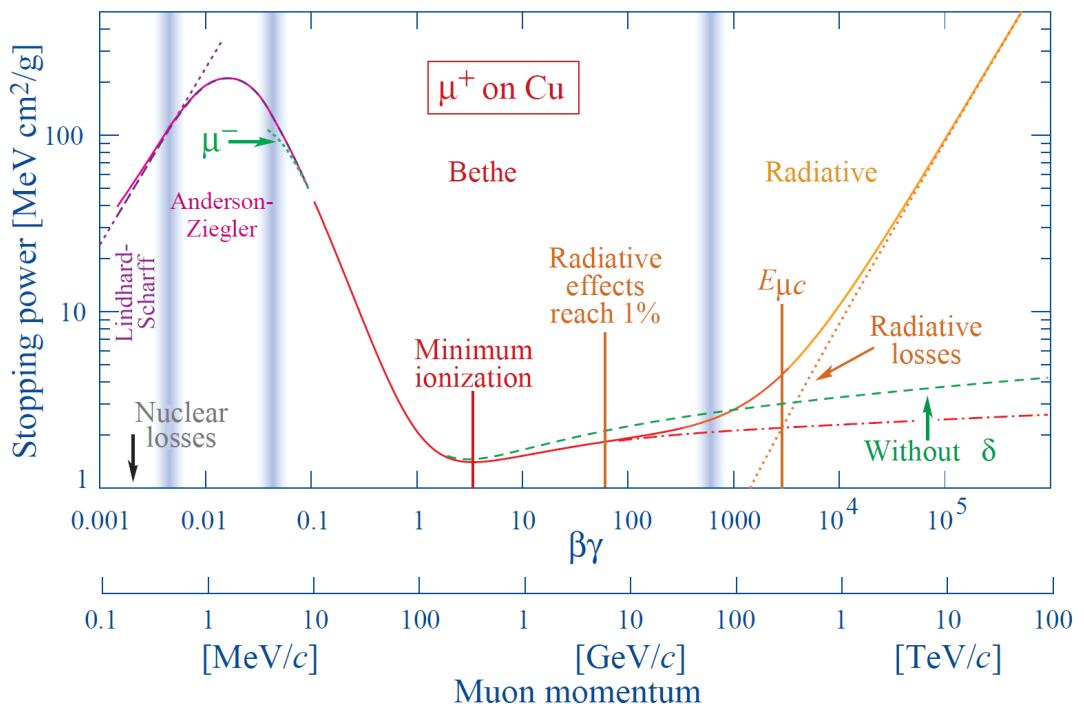


Figure 2.4: Stopping power

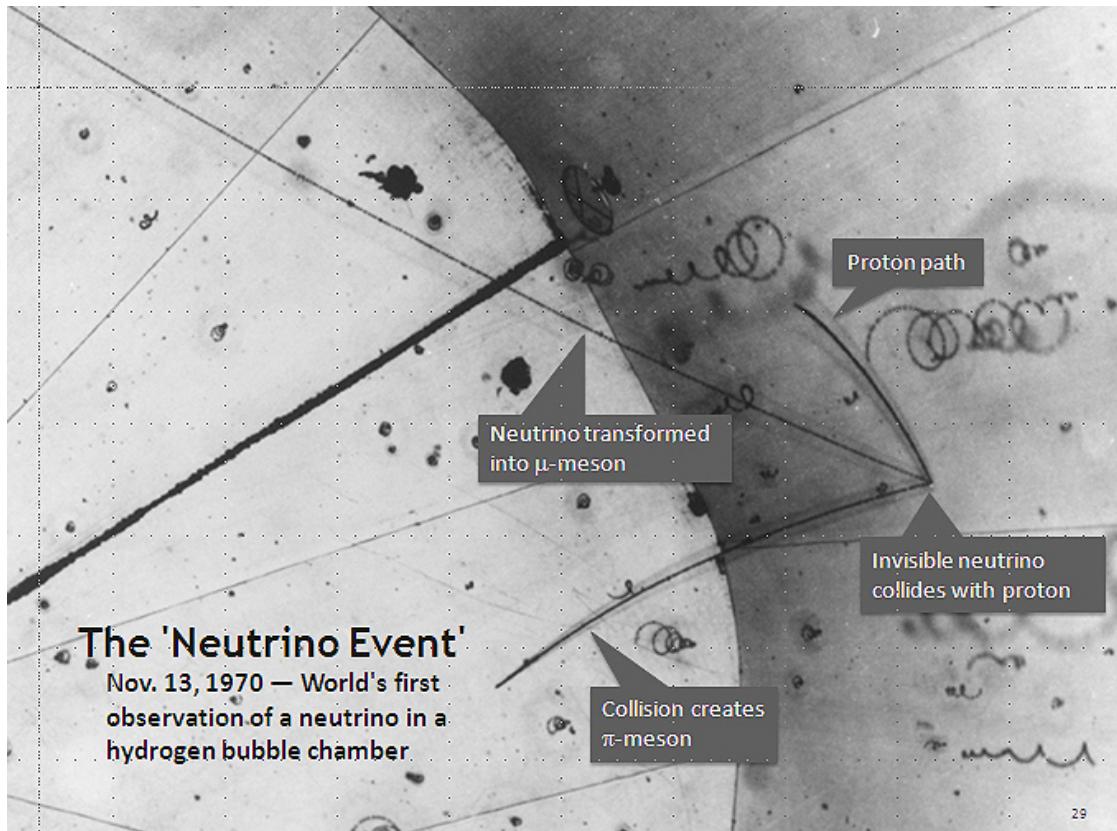


Figure 2.5: First neutrino seen in a bubble chamber. Image from [argonne_national_laboratory_first_1970](#).

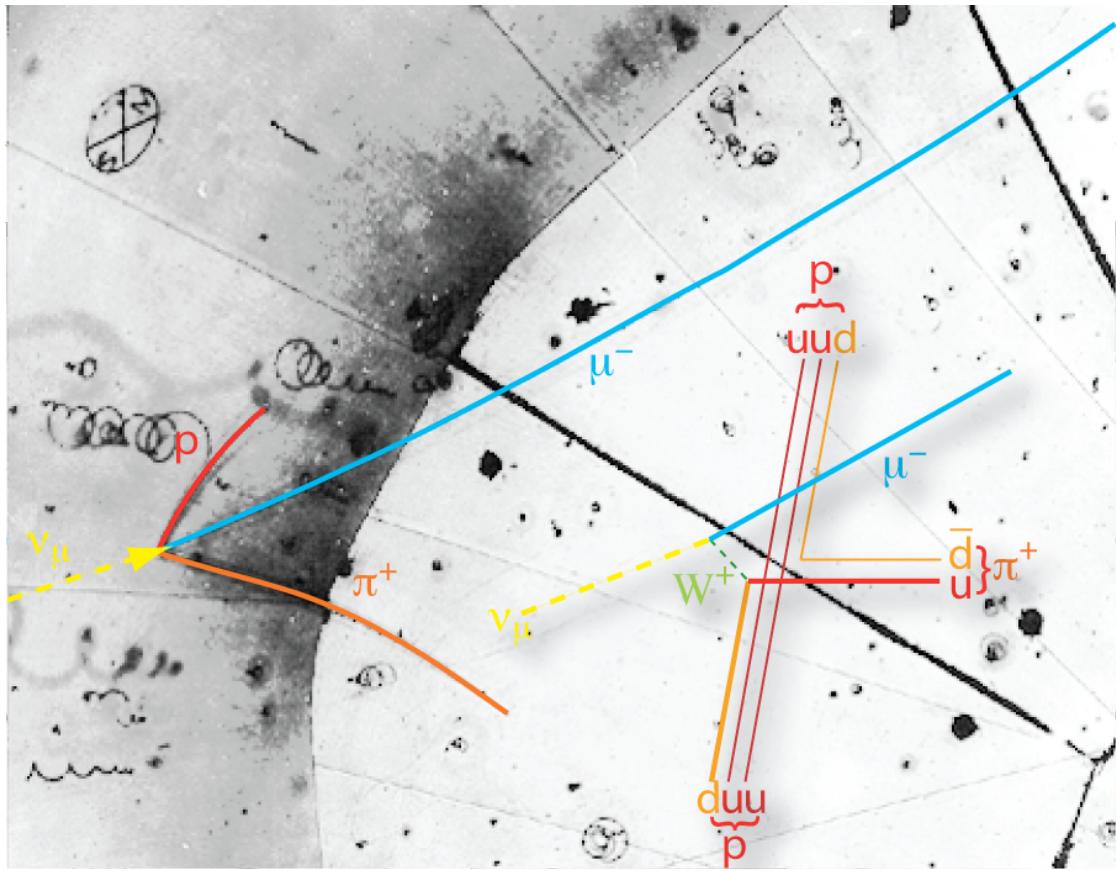


Figure 2.6: Bubble chamber neutrino event with annotation. Image from [argonne_national_laboratory_neutrino_2012](#).

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 4).
- [2] S Baur, “Dark matter searches with the icecube upgrade,” arXiv preprint arXiv:1908.08236 (2019) (cited on page 4).
- [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 4).
- [4] L Köpke, “Supernova neutrino detection with icecube,” arXiv preprint arXiv:1106.6225 (2011) (cited on page 4).
- [5] M Leuermann, M Krämer, and CHV Wiebusch, “Testing the neutrino mass ordering with icecube deepcore,” s3.cern.ch (2018) (cited on page 4).
- [6] M Raghu and E Schmidt, “A survey of deep learning for scientific discovery,” arXiv preprint arXiv:2003.11755 (2020) (cited on page 4).
- [7] A Ishihara, “The icecube upgrade–design and science goals,” arXiv preprint arXiv:1908.09441 (2019) (cited on page 5).
- [8] K Abe, R Akutsu, A Ali, C Alt, and C Andreopoulos..., “Constraint on the matter-antimatter symmetry-violating phase in neutrino oscillations,” arXiv preprint arXiv ... (2019) (cited on page 7).