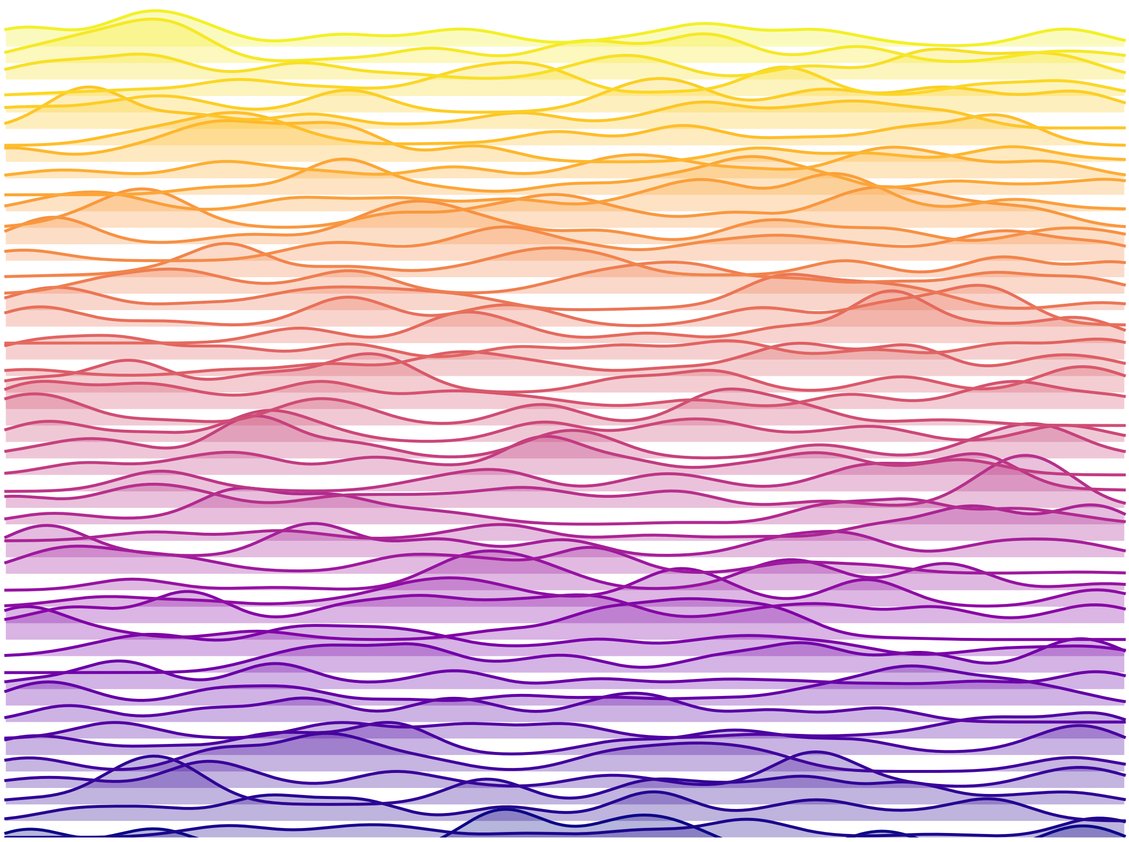
ELEMENTAL COMPOSITION OF COSMIC RAYS

ANALYSIS OF ICECUBE DATA USING GRAPH NEURAL NETWORKS

PARAS KOUNDAL

Karlsruher Institut für Technologie



ELEMENTAL COMPOSITION OF COSMIC RAYS: ANALYSIS OF ICECUBE DATA USING GRAPH NEURAL NETWORKS

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN (Dr. rer. nat.)
von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte
DISSERTATION
von

Paras Koundal (Masters of Science, with Major in Physics)
aus Himachal Pradesh, India

Referent: Prof. Dr. Ralph Engel
Korreferent: Prof. Dr. Ulrich Husemann
Betreuer: Dr. Andreas Haungs
Tag der mündlichen Prüfung: July 14th, 2023

Paras Koundal - *Elemental Composition of Cosmic Rays: Analysis of IceCube data using Graph Neural Networks*. Defended On : July 14th, 2023.
The author extends gratitude to André Miede for developing the `classicthesis` template used for this work.

Dedicated to,
My family and friends

In memory of,
Those Gone but Never Forgotten

ABSTRACT

Elemental Composition of Cosmic Rays: Analysis of IceCube data using Graph Neural Networks

Cosmic rays (CRs) are high-energy ionized nuclei that emanate from astrophysical sources and regularly bombard Earth's atmosphere. The IceCube Neutrino Observatory (ICNO), a cubic-kilometer observatory embedded in South-Pole Antarctic ice, is very well suited to investigate CRs in the energy regime where the transition from Galactic to the extra-galactic origin of CRs occurs. This work delves into the analysis of the elemental composition at ICNO using the footprint of extensive air showers (EAS) detected at IceTop (the surface detector component of IceCube) in conjunction with the in-ice component. Using dedicated Monte Carlo simulations, the work leverages energy deposits by TeV-muons (within EASs) in IceCube, to devise various CR composition-sensitive physics observables. In addition to providing composition-sensitivity, the observables also provide the possibility for testing internal consistencies in phenomenological models which describe hadronic interactions of CRs with atmospheric-nuclei.

An extended part of this work focuses on developing a Graph Neural Network (GNN)-based approach, informed by the physics of EASs. This approach utilizes the observed footprint of EAS at both IceTop and IceCube to estimate logarithmic mass for each individual EAS. As a first-of-its-kind endeavor, the GNN architecture incorporates multiple EAS-physics-inspired inductive-biases to obtain the mass estimate in an efficient manner. Moreover, in order to provide enhanced stability to the network the work also adapts ideas from other leading fields in Deep Learning. In addition to the physics-informed application of the GNN, the network also benefits from the composition-sensitive observables developed in this work, along with other shower observables, to provide enhanced composition sensitivity. A gradient-boosted decision tree-based approach is used for the energy estimate. The different Machine Learning (ML) methods are concurrently utilized to estimate the elemental composition of CRs. The methods developed on simulations are tested and validated with a burnsample of real data. Finally the work presents the elemental composition of real-data (using 10% of 2012-data, spread throughout the year) observed at ICNO, obtaining fractional spectra for four primary groups (proton, Helium, Oxygen, and Iron). The fractions obtained are validated using distributions of EAS-observables in real data.

This work makes fractional estimates of primary types which indicate a heavier composition than the prior work at ICNO. This has strong physics implications as it might indicate the presence of the highest-energy galactic CRs in the transition region between Galactic to the extragalactic origin of CRs. A future update of this work, incorporating greater statistics and including quantified detector systematics, has the potential to enhance the physics significance of the results obtained from this work. This work also serves as a very useful test case for using GNNs in the planned upgrade of ICNO, namely IceCube-Gen2. In addition, to that this work also establishes GNNs as a very useful tool for exploring unique problems in astroparticle physics.

ZUSAMMENFASSUNG

Elementzusammensetzung der kosmischen Strahlung: Analyse von IceCube-Daten mit Graph Neural Networks

Kosmische Strahlung (CRs) sind hochenergetische ionisierte Kerne, die von astrophysikalischen Quellen ausgehend regelmäßig die Erdatmosphäre bombardieren. Das IceCube-Neutrino-Observatorium (ICNO), ein Kubikkilometer großer Detektor, der in das antarktische Südpol-Eis eingebettet ist, eignet sich sehr gut für die Untersuchung von CRs im Energiebereich, in dem der Übergang von galaktischem zu extra-galaktischem Ursprung von CRs stattfindet. Diese Arbeit befasst sich mit der Analyse der Elementzusammensetzung bei ICNO unter Verwendung des Fußabdrucks ausgedehnter Luftschauer (EAS), die mit IceTop (der Oberflächendetektor-Komponente von IceCube) in Verbindung mit der In-Eis-Komponente nachgewiesen wurden. Unter Verwendung spezieller Monte-Carlo-Simulationen nutzt die Arbeit Energieverluste durch TeV-Muonen (innerhalb von EASs) in IceCube, um verschiedene CR-empfindliche physikalische Observable zu rekonstruieren. Neben der Kompositionssensitivität bieten die Observablen auch die Möglichkeit, interne Konsistenzen in phänomenologischen Modellen zu testen, die hadronische Wechselwirkungen von CRs mit atmosphärischen Kernen beschreiben.

Ein erweiterter Teil dieser Arbeit konzentriert sich auf die Entwicklung eines Graph Neural Network (GNN)-basierten Ansatzes, der auf der Physik von EASs beruht. Dieser Ansatz nutzt den beobachteten Fußabdruck von EAS sowohl bei IceTop als auch in IceCube, um die logarithmische Masse für jedes einzelne EAS zu schätzen. Die GNN-Architektur ist das erste Projekt dieser Art, das mehrere von der Physik inspirierte „inductive Bias“ einbezieht, um die Masse auf effiziente Weise zu schätzen. Um die Stabilität des Netzwerks zu erhöhen, werden außerdem Ideen aus anderen führenden Bereichen des Deep Learning übernommen. Neben der physikalisch informierten Anwendung des GNN profitiert das Netzwerk auch von den kompositions-sensitiven Observablen, die in dieser Arbeit entwickelt wurden, zusammen mit anderen rekonstruierten Luftschauergrößen, um eine verbesserte Kompositionssensitivität zu erreichen. Für die Energieschätzung wird ein auf einem gradientenverstärkten Entscheidungsbaum basierender Ansatz verwendet. Die verschiedenen Methoden des Maschinellen Lernens (ML) werden gleichzeitig eingesetzt, um die Elementzusammensetzung von CRs zu schätzen. Die in Simulationen entwickelten Methoden werden mit einer Stichprobe realer Daten getestet und validiert. Schließlich wird die Elementzusammensetzung von realen Daten (unter Verwendung von 10% der 2012-Daten, die über das Jahr verteilt sind) vorgestellt, um Anteilsspektren für vier Primärgruppen (Proton, Helium, Sauerstoff und Eisen) zu erhalten. Die erhaltenen Anteile werden anhand von Verteilungen von EAS-Beobachtungsdaten in realen Daten validiert.

Aus dieser Arbeit resultieren Schätzungen der Anteile der primären Massengruppen, die auf eine schwerere Zusammensetzung als in den früheren Arbeiten zu ICNO hinweisen. Dies hat starke physikalische Implikationen, da es auf das Vorhandensein von galaktischen CRs mit höchster Energie im Übergangsbereich zwischen galaktischem und extragalaktischem CRs hinweisen könnte. Eine künftige Aktualisierung dieser Arbeit unter Einbeziehung größerer statistischer Daten und einer besser verstandenen Systematik des Detektors könnte die physikalische Bedeutung der Ergebnisse dieser Arbeit erhöhen. Diese Arbeit dient auch als ein sehr nützlicher Testfall für die Verwendung von GNNs in der geplanten Erweiterung von ICNO, IceCube-Gen2. Darüber hinaus etabliert diese Arbeit auch GNNs als ein sehr interessantes Werkzeug für die Erforschung weiterer Fragestellungen in der Astroteilchenphysik.

...I can no other answer make but thanks,
And thanks; and ever thanks...

— Twelfth Night (Act III, Scene 3) : William Shakespeare [1]

ACKNOWLEDGMENTS

This research endeavor of mine was no solo journey. There are many to whom I would like to express my profound gratitude.

Foremost I would like to thank my supervisor Prof. Dr. Ralph Engel, who provided me with the opportunity to work in the scientifically rich environment at the Institute of Astroparticle Physics, KIT and with the IceCube Collaboration. I would also like to thank my co-supervisor Prof. Dr. Ulrich Husemann for giving me useful scientific advice. A very special thanks go to my research-supervisor Dr. Andreas Haungs who has not only guided me every step of the way in my research journey but has also acted as the lighthouse of wisdom on every aspect of life. The triad has helped shape my approach to scientific as well as personal matters.

I believe that scientific discussions with Dr. Donghwa Kang, Dr. Markus Roth, Dr. Tanguy Pierog, Prof. Dr. Frank G. Schöder, Dr. Dennis Soldin and members of cosmic-ray community in IceCube Collaboration have helped me in evolving as a more confident cosmic-ray researcher, than would have been possible otherwise. It goes without saying that a very special thanks to Sabine Bucher, Doris Wochele and Frank Polgart for providing a very smooth operation of the administrative and technical work environment is fully deserved.

I would like to express my heartfelt gratitude to Julian Saffer, Pranav Sampathkumar, Steffen Hahn, Federico Bontempo, and Fiona Ellwanger who came together with me to create an environment where we could learn from the experiences each other in Machine Learning and helped each other become a better practitioner of the skill.

My passionate thanks also go to my very dear colleagues Tista Mukherjee, Shefali, and Mark Weyrauch for their unwavering support, continuous encouragement, and shared wisdom. I am also very thankful to Sara Martinelli, Megha Venugopal, Wenjie Hou, Vladimir Lenok, Viktoria Tokareva, Olena Tkachenko, Emily Pereira Martins Edyvania, Nikos Karastathis, Max Büsken and everyone else for building a conducive work environment. A token of appreciation goes to Agnieszka Leszczyńska, Thomas Huber, Hrvoje Dujmovic, and Aswathi Balagopal who helped guide me through the intricacies of the research and bureaucratic life at the advent of my journey. I am also thankful to my flatmates Anmol Shukla and Prerak Dhawan who always made me feel like home.

No words can express the gratitude I feel towards my family who provided me with the foundation on which I am standing today. I am indebted to their support, love, and patience.

I would also like to thank essential care workers in every part of the world. They have been the warriors for all of us.

CONTENTS

1 INTRODUCTION 1
2 COSMIC RAY PHYSICS 5
2.1 The Landscape of Cosmic Rays . . . 5
2.2 Extensive Air-Showers . . . 10
2.2.1 Muons . . . 13
3 ICECUBE OBSERVATORY 19
3.1 IceCube : The Cubic-Km Eye on the Cosmos . . . 19
3.1.1 Building IceCube - The Rationale . . . 19
3.1.2 Coordinate System . . . 21
3.1.3 Digital Optical Module - DOM . . . 22
3.1.4 IceCube Laboratory . . . 24
3.1.5 IceTop . . . 25
3.1.6 In-Ice Array . . . 26
3.1.7 Data Acquisition . . . 29
3.1.8 Calibration . . . 32
3.1.9 Environmental Characteristics @ South Pole . . . 33
3.2 IceCube-Gen2: The window to the extreme Universe . . . 34
4 SCIENCE @ ICECUBE 39
4.1 Neutrino Astronomy . . . 39
4.2 Cosmic-ray Physics . . . 42
4.3 Multimessenger Astronomy . . . 45
4.4 Other Science . . . 46
4.4.1 Beyond Standard Model (BSM) Physics . . . 46
4.4.2 Dark Matter . . . 46
4.4.3 Quantum Gravity . . . 46
4.4.4 Glaciology . . . 47
4.4.5 More . . . 47
5 SIMULATION AND RECONSTRUCTION 49
5.1 EAS Simulations . . . 49
5.1.1 CORSIKA . . . 49
5.1.2 Hadronic Interaction Models . . . 50
5.1.3 Thinning . . . 53
5.2 Detector Simulation . . . 54
5.2.1 Dethinning . . . 54
5.2.2 Resampling . . . 55
5.2.3 Detector Response . . . 57
5.3 Reconstructions . . . 57
5.3.1 IceTop Reconstructions . . . 58
5.3.2 IceCube Reconstructions . . . 61
5.3.3 Quality Cuts . . . 63
5.3.4 Performance . . . 65
6 GRAPH NEURAL NETWORK - A PRIMER 69

6.1 Graphs - A Tour d'Horizon . . . 70
6.1.1 Graph Representation . . . 71
6.1.2 Graph Attributes . . . 72
6.2 Graph-Based Learning . . . 73
6.3 Graph Neural Network . . . 74
6.3.1 Message Passing . . . 74
6.3.2 Limits of Graph Neural Networks . . . 75
6.4 Bridging Ideas with CNNs . . . 80
6.4.1 DenseNet . . . 80
6.4.2 InceptionNet . . . 81
6.5 Graph Neural Networks at IceCube Observatory . . . 81
7 CHARACTERIZING COMPOSITION: USING TEV MUONS 85
7.1 Differential Energy Loss : $\frac{dE}{dX_{1500m}}$ . . . 86
7.2 Total Stochastic Energy . . . 90
7.3 Ratio Parameter . . . 93
7.4 Mean Radii and Mean Charge . . . 95
7.4.1 Mean Radii . . . 97
7.4.2 Mean Charge . . . 99
7.5 Composition Consistency Among Observables . . . 101
7.5.1 $\frac{dE}{dX_{1500m}}$ vs Mean Radii . . . 103
7.5.2 Mean Charge vs Mean Radii . . . 105
8 ANALYSIS 109
8.1 Mass-Estimate using Graph Neural Network . . . 109
8.1.1 Nodes: The Fine Print . . . 111
8.1.2 The Beauty of Meaningful-Connections . . . 112
8.1.3 Homophily . . . 120
8.1.4 Point-Pair Convolution . . . 123
8.1.5 Point Transformer . . . 126
8.1.6 Dynamic-Edge Convolution . . . 128
8.1.7 Bridging Ideas: DenseNet and InceptionNet . . . 128
8.1.8 Global Pooling . . . 129
8.1.9 Adding Physics Features . . . 129
8.1.10 Confluence of Representations . . . 129
8.1.11 Loss Function and other details . . . 130
8.1.12 Mass Prediction . . . 130
8.2 Energy-Estimate using Gradient Boosted Decision Trees . . . 131
8.3 Mass Prediction on Measured Data . . . 133
8.4 Template Fitting : Method and Testing . . . 134
9 RESULTS 139
9.1 Predicting Composition . . . 139
9.2 Composition : Consistency Check . . . 145
10 CONCLUSIONS 151

I APPENDIX 153
A PHYSICS 155
A.1 CORSIKA Atmosphere . . . 155
A.2 Muon Range @ IceCube . . . 155

A.3 Figure of Merit : FOM 156
A.4 Normalized Overlap Area: NOA 156
A.5 Mean Radii and Mean Charge 157
A.6 Testing Template Fitting 157
A.7 Template Fits on Real Data : Using GNN Predictions 157
A.8 Template Fits on Real Data : Using Air-Shower Observables 157
A.9 Compatibility Check: Fraction Estimates 175
B DETECTOR 181
B.1 MILLIPEDE 181
B.2 DDDDR - Data Derived Differential Deposition Reconstruction 182
C MACHINE LEARNING 185
C.1 Weisfeiler-Lehman (WL) Test 185
C.2 CNN @ IceCube 185
C.3 Homophily 185
C.4 PointNet 187
C.5 Transformer and Attention 187
D SCIENCE COMMUNICATION 189

BIBLIOGRAPHY 191

LIST OF FIGURES

Figure 2.1 The energy spectrum of cosmic rays 7
Figure 2.2 Total and individual primaries flux-parametrization by GSF Model. 9
Figure 2.3 Illustration of Extensive Air-Shower. 10
Figure 2.4 CORSIKA simulations of extensive air showers initiated by gamma, proton, and Iron primaries. 11
Figure 2.5 Correlation between abundance of muons and electrons in EAS, at sea-level. 12
Figure 2.6 Muon multiplicity expectation from Elbert formula. 15
Figure 2.7 Muon energy-loss in ice. 16
Figure 2.8 Cherenkov radiation by subatomic particle 17
Figure 3.1 A Schematic of IceCube Observatory 20
Figure 3.2 Anticipated astrophysical neutrino fluxes at ICNO 21
Figure 3.3 Schematic of DOMs used at IceCube Observatory 22
Figure 3.4 DOM on an in-ice cable 24
Figure 3.5 IceCube Laboratory 24
Figure 3.6 Top view of IceTop array 25
Figure 3.7 Height variation among IceTop (IT) tanks, without snow coverage 26
Figure 3.8 Snow-accumulation on IT-tanks over multiple years 27
Figure 3.9 Dust concentration variation as a function depth 28
Figure 3.10 Neutrino event topologies at IceCube. 30
Figure 3.11 Cabling to check local-coincidence at an IT station 31
Figure 3.12 Charge spectrum at IT and fits to EM and muonic component . . 33
Figure 3.13 Prototype of mDOM and D-Egg 35
Figure 3.14 Planned design of IceCube-Gen2 Observatory 35
Figure 3.15 Planned design of the Surface-Array component of IceCube-Gen2 Observatory 37
Figure 3.16 Elevated scintillator-panel and radio-antenna at South-Pole . . . 37
Figure 3.17 Illustration of hybrid radio station 38
Figure 4.1 Bert, Ernie and Big Bird 39
Figure 4.2 Hydrangea 40
Figure 4.3 Containment region corresponding to observation of blazar TXS 0506+056 at IceCube Observatory. 41
Figure 4.4 Milky Way galactic plane in photons and neutrinos 41
Figure 4.5 Cosmic-ray (CR) intensity map at using 6 year of observed data . 43
Figure 4.6 Full-sky CR relative-intensity map using HAWC and IceCube data 43
Figure 4.7 Observation of blazar TXS 0506+056 using multiple messengers. 45
Figure 5.1 Excess average muon-multiplicity predicted by hadronic-models w.r.t. SIBYLL 2.1 53
Figure 5.2 Excess average muon-multiplicity predicted by hadronic-models w.r.t. SIBYLL 2.1 53
Figure 5.3 Resampling radii for IT as a function of MC-primary energy . . . 55

Figure 5.4 EAS response at IT, and the reconstructed LDF 62
Figure 5.5 IT core containment cut 64
Figure 5.6 $\mathrm{Log}_{10}S_{125\mathrm{m}}/\mathrm{VEM}$ as a function of true MC-energy 66
Figure 5.7 $\mathrm{Log}_{10}S_{125\mathrm{m}}/\mathrm{VEM}$ distribution of MC-simulations which pass quality cuts 67
Figure 5.8 Zenith and Azimuth distribution of MC-simulations which pass quality cuts 67
Figure 6.1 Different delineations for identical underlying connectivity. . . . 70
Figure 6.2 An example graph with node-attributes and weighted edges. . . 72
Figure 6.3 Message Passing in a GNN 74
Figure 6.4 Receptive field among CNN layers 77
Figure 6.5 Receptive field for an IC event mapped as a graph, among GNN layers 78
Figure 6.6 Measure of over-smoothing for *CiteSeer*-dataset 79
Figure 6.7 A block used in Densely Connected Convolutional Network . . . 81
Figure 6.8 A module used in InceptionNet, with dimension reduction . . . 82
Figure 7.1 MILLIPEDE reconstructed energy-deposits as a function of slant depth. 87
Figure 7.2 Average energy loss at a slant-depth of 1500 m as a function of in-ice muon-multiplicity. 87
Figure 7.3 Composition sensitivity of $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ and Data-MC overlap. 88
Figure 7.4 Composition sensitivity of TSE and Data-MC overlap. 92
Figure 7.5 Composition sensitivity of Ratio Parameter and Data-MC overlap. 94
Figure 7.6 Representation of in-ice charge deposit and maximum-radii cut. 96
Figure 7.7 Composition sensitivity of Mean Radii (Max. DOM Radii = 200 m) and Data-MC overlap. 98
Figure 7.8 Composition sensitivity of Mean Charge (Max. DOM Radii = 200 m) and Data-MC overlap. 100
Figure 7.9 Confidence-region for Mean Radii as a function of $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ and Data-MC overlap for different hadronic interaction models. . . . 104
Figure 7.10 Confidence-region for Mean Radii as a function of Mean Charge and Data-MC overlap for different hadronic interaction models. 106
Figure 8.1 GNN-based architecture for CR-Composition estimate at IceCube Observatory 110
Figure 8.2 The role of inductive biases in ML models. 113
Figure 8.3 Kernel-size dependence on the convolution output. 117
Figure 8.4 Number of graph-edges as a function of graph-cardinality and $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$. 119
Figure 8.5 Illustration of an example EAS mapped as a graph. 121
Figure 8.6 Homophily for Total Charge and time of maximum charge-deposit as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$, with maximum node connections = 10. 124
Figure 8.7 Homophily for Charge weighted x and z coordinate as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$, with maximum node connections = 10. 125

Figure 8.8 Response of GNN-based architecture to various primary-types on test-dataset, as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$. 132
Figure 8.9 Bias and resolution of energy-prediction 133
Figure 8.10 Feature importance for energy-prediction 134
Figure 8.11 Response of GNN-based architecture to MC and burnsample, as a function of reconstructed energy 135
Figure 8.12 Reconstructed logarithmic mean-mass for a test of different underlying composition fractions. 137
Figure 9.1 Raw energy-spectrum using burnsample-data 140
Figure 9.2 Reconstructed primary type fractions in burnsample, using ln(A) predictions from GNN-based architecture. 141
Figure 9.3 Mean logarithmic mass expectation from this work (using GNN predictions) and the comparison with previous IceCube-Collaboration (ICoII) publication. 142
Figure 9.4 Comparison of $\langle \ln(A) \rangle$ expectation from this work(using GNN predictions) and the previous ICoII publication with results of $\langle \ln(A) \rangle$ from KASCADE 145
Figure 9.5 Outline of steps carried out for compatibility-check of primary-type fractions obtained 146
Figure 9.6 Linear-scale: Compatibility checks for composition-fractions obtained from GNN for $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ and comparison with fractions from other flux-models 147
Figure 9.7 Linear-scale: Compatibility check for composition-fractions obtained from GNN for Mean-Radii (with maximum DOM-radii cut of 200 m) and comparison with fractions from other flux-models 148
Figure A.1 Normalized Overlap Area for two distributions 156
Figure A.2 Composition sensitivity of Mean Radii (Max. DOM Radii = 400 m) and Data-MC overlap. 158
Figure A.3 Composition sensitivity of Mean Radii (Max. DOM Radii = 600 m) and Data-MC overlap. 159
Figure A.4 Composition sensitivity of Mean Charge (Max. DOM Radii = 600 m) and Data-MC overlap. 160
Figure A.5 Composition sensitivity of Mean Charge (Max. DOM Radii = 600 m) and Data-MC overlap. 161
Figure A.6 Underlying KDE-template fits for the case of pure Proton. 162
Figure A.7 Underlying KDE-template fits for the case of pure Helium. 163
Figure A.8 Underlying KDE-template fits for the case of pure Oxygen. . . . 164
Figure A.9 Underlying KDE-template fits for the case of pure Iron. 165
Figure A.10 Underlying KDE-template fits for the case of sinusoidal variation between Proton and Iron. 166
Figure A.11 KDE-template fits for burnsample, using ln(A) predictions from GNN and SIBYLL 2.1 as the hadronic interaction model. 167
Figure A.12 Reconstructed primary type fractions in burnsample (using ln(A) predictions from GNN-based architecture) and comparison with last relevant ICoII publication. 168

Figure A.13 Comparison of $\langle \ln(A) \rangle$ expectation from this work (using GNN predictions) and the previous ICoII publication for the choice of different hadronic interaction models . . . 169
Figure A.14 KDE-template fits for burnsample, using ln(A) predictions from GNN and SIBYLL 2.3 as the hadronic interaction model. . . . 170
Figure A.15 KDE-template fits for burnsample, using ln(A) predictions from GNN and EPOS-LHC as the hadronic interaction model. . . . 171
Figure A.16 KDE-template fits for burnsample, using ln(A) predictions from GNN and QGSJet-II-04 as the hadronic interaction model. . . . 172
Figure A.17 Comparison of $\langle \ln(A) \rangle$ expectation from this work (using GNN predictions) and the previous ICoII publication with results from other observatories . . . 173
Figure A.18 Comparison of $\langle \ln(A) \rangle$ expectation from this work using EAS-observables, to those from GNN predictions from this work and the previous ICoII publication . . . 174
Figure A.19 KDE-template fits for burnsample, using MC-distributions of $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}} / \frac{GeV}{m}\right)$ . . . 176
Figure A.20 KDE-template fits for burnsample, using MC-distributions of Mean Radii. . . . 177
Figure A.21 Log-scale: Compatibility checks for composition-fractions obtained from GNN for $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}} / \frac{GeV}{m}\right)$ and comparison with fractions from other flux-models . . . 178
Figure A.22 Log-scale: Compatibility check for composition-fractions obtained from GNN for Mean-Radii (with maximum DOM-radii cut of 200 m) and comparison with fractions from other flux-models . . 179
Figure B.1 Energy loss along the segmented track. . . . 181
Figure B.2 Charge footprint and unfolded energy loss using MILLIPEDE. . 182
Figure C.1 Division of IceCube array for CNN-based implementation at IceCube . . . 186
Figure C.2 Homophily for Charge weighted y coordinate as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$, with maximum node connections = 10. . . . 186
Figure C.3 Homophily for Total Charge at a node (DOM), with maximum node connections = 40. . . . 187
Figure C.4 PointNet Architecture . . . 187

LIST OF TABLES

Table 5.1 Variety of available event generators . . . 50
Table 5.2 Details of CORSIKA datasets used for this work. . . . 56
Table 5.3 IceTop quality cuts . . . 63
Table 5.4 IceCube quality cuts . . . 65

Table 7.1 Passing Fraction (p and Fe) for different choices of maximum DOM radii. . . . 97
Table 8.1 Mean classification accuracy for low-homophily and high-homophily regime on a synthetic graph dataset 122
Table 9.1 p-value for compatibility test with $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$, using this work and other flux-models 149
Table 9.2 p-value for compatibility test with Mean-Radii (with maximum DOM-radii cut of 200 m), using this work and other flux-models 150

ACRONYMS

CR cosmic-ray

EAS Extensive Air-Shower

LDF Lateral Distribution Function

HE high-energy

MC Monte Carlo

UHECR Ultra-high-energy Cosmic Ray

CORSIKA COsmic Ray SImulations for KAscade

EGS Electron Gamma Shower system

NKG Nishimura Kamata Greisen

GSF Global Spline Fit

GST Gaisser-Stanev-Tilav

H4a-IT Variation of Gaisser's H4a flux using only four components

EM Electromagnetic

HEP High-Energy Physics

CS cross-section

DM Dark Matter

SM Standard Model

BSM Beyond Standard Model

CC Charged-Current

NC Neutral-Current

QCD Quantum Chromodynamics

mwe Meter Water Equivalent

PE Photoelectron

AGN Active Galactic Nuclei

GRB Gamma Ray Burst

FRB Fast Radio Burst

TDE Tidal Disruption Event

SNR Supernova Remnant

GZK Greisen–Zatsepin–Kuzmin

WIMP Weakly interacting massive particle

KIT Karlsruher Institut für Technologie

LHC Large-hadron Collider

AMANDA Antarctic Muon And Neutrino Detector Array

DUMAND Deep Underwater Muon and Neutrino Detector

KASCADE KArlsruhe Shower Core and Array DEtector

LHAASO Large High Altitude Air Shower Observatory

LIGO Laser Interferometer Gravitational-Wave Observatory

ICoII IceCube Collaboration

ICNO IceCube Neutrino Observatory

ICL IceCube Laboratory

IT IceTop

IC IceCube

DC DeepCore

IT-IC IceTop and IceCube

DOM Digital Optical Module

PROPOSAL Propagator with Optimal Precision and Optimized Speed for All Leptons

PMT Photo-Multiplier Tube

COG Center-of-Gravity

VEM Vertical Equivalent Muons

LC Local Coincidence

HLC Hard Local Coincidence

SLC Soft Local Coincidence

HG High-Gain

LG Low-Gain

DDDDR Data Derived Differential Deposition Reconstruction

SMT Simple Multiplicity Trigger

GBDT Gradient-Boosted Decision Trees

ML Machine Learning

MLP Multi-Layer Perceptron

NN Neural Network

DL Deep Learning

CNN Convolutional Neural Network
RNN Recurrent Neural Network
LSTM Long Short Term Memory
NLP Natural Language Processing
LLM Large Language Model
GMS graph mapped system
GNN Graph Neural Network
PPFConv Point-Pair feature convolution
KS Kolmogorov–Smirnov
WL Weisfeiler-Lehman
kNN k - Nearest Neighbours
RLHF Reinforcement Learning from Human Feedback
FOM Figure of Merit
NOA Normalized Overlap Area
burnsample 10% of 2012's real-data
KDE Kernel Density Estimation
TSE Total Stochastic Energy
PC Photograph or Illustration Credits

1 INTRODUCTION

...Yeh tara, woh tara, har tara
Yeh sab saath mein
Joh hai raat mein
Toh jagmagaya aasman sara.
(...This star, that star, every star
When they all come together
In the night
Then the whole sky shimmers.)

— Movie - Swades: We, the People (Homeland). Lyricist - Javed Akhtar

Cosmic-rays (CRs), a *"thin-rain"*[1] of ionized nuclei, regularly bombard the Earth's atmosphere. With energies surpassing those achievable by any earth-bound accelerator by a factor of over a million, these cosmic-ambassadors provide the most representative matter sample of the astrophysical sources they originate from. CRs were discovered by Victor Franz Hess during his series of balloon flights in 1911 and 1912. Through his balloon flights, Hess was able to detect ionization produced by an unknown radiation in an electroscope. He observed that ionization first decreases with altitude and then noticeably increases again at altitudes greater than 1000 m [2]. He finally attributed the increased ionization and hence increased radiation to sources from outer space[2]. This study was extended by Werner Kolhörster to altitudes exceeding 9 km, using improved instrumentation. He also observed an increase in radiation with an increase in altitude [3, 4], hence confirming Hess's hypothesis of extraterrestrial origin. Later studies [5, 6], by probing the absorption characteristics and latitude-variation, helped established the charged particle nature of CRs. Using high-altitude balloon flights, Marcel Schein and others [7] helped establish that the radiation primarily consisted of ionized protons. Present knowledge indicates that CRs consist of approximately 90% protons, about 9% α-particles, with the remaining fraction comprising heavier nuclei. The total particle flux drops steeply with increasing energies. Hence, the CRs impinging Earth are primarily low-energy CRs.

Mankind has made immense technological and scientific progress in the last century. Our capability to produce high-energy particles using giant accelerators has grown by leaps and bounds. This has also allowed huge international collaborations on scientific as well as political frontiers[3]. However, we still have to cover about $\mathcal{O}(10^6)$ in

1 *thin-rain*: Poetic wording borrowed from the Nobel-Prize Lecture of Cecil F. Powell (December 11, 1950).
2 Hess's highest flight went up to a height of 5.3km. This flight (17 April 1912) was done during a partial solar eclipse. No significant decrease in ionization hinted towards the sun as not the source of the radiation.
3 *CERN (Conseil Européen pour la Recherche Nucléaire)*, currently hosting the most powerful particle accelerator facility in the world, was also the birthplace of the World Wide Web, hence responsible for the advent of *information-age*.

energy-scale to reach the energy of highest-energy CR-event ever observed[4]. CRs pervasivenature and more than a century of research into them have helped mature the field. However, their sources and propagation histories through the interstellar and intergalactic media remain largely enigmatic. Chapter 2 presents our current knowledge about CRs. A comprehensive investigation into CRs presents the opportunity to not only study CRs, but also bears significant implications on other astrophysical messengers like neutrinos, gamma-rays and vice versa[5] [8]. IceCube Neutrino Observatory (ICNO), a cubic-kilometer astroparticle detector at the geographic South Pole, is very well suited to study these messengers and hence help provide a comprehensive understanding of the skies. The surface array of ICNO, IceTop (IT), is primarily used for CR-analysis, whereas the in-ice component, IceCube (IC), is primarily used for neutrino-related studies. Chapter 3 and Chapter 4 present a detailed summary of ICNO (and the planned enhancements), and the variety of ongoing/completed multi-messenger (and interdisciplinary) studies at ICNO respectively.

This work utilizes IT as well as IC to investigate the elemental composition of CRs detected at ICNO. An investigation into the elemental makeup of CRs holds the potential to demystify their sources, provide insights into their propagation mechanisms, and elucidate the source dynamics which facilitate the acceleration of CR-nuclei to extreme energies. In an integrated operation of IT and IC, the observatory is sensitive to CRs in an energy range between ≈ 1 PeV to ≈ 1 EeV. This is approximately the transition region between galactic and extra-galactic CRs. ICNO is currently the only CR-observatory which spans this whole energy range. It also holds the unique position of being one of its kind three-dimensional CR-observatory. This allows for physics-tests of interactions in extensive air-showers (CR-initiated cascade of secondaries in the atmosphere) using GeV and TeV muons and is closely connected to CR-composition, which currently no other CR observatory can do [9]. Being at the South Pole also gives it the additional benefit of viewing an almost fixed part of astrophysical skies. Since, ICNO is a multi-messenger observatory comparing the physics results from the variety of astrophysical messengers also becomes easier.

In order to study the elemental composition of CRs, this work uses Monte Carlo (MC)-simulations of extensive air showers generated using CORSIKA. Chapter 5 provides a detailed summary of simulations used for CR-analysis at ICNO. The chapter also summarizes the details of shower reconstructions and simulation-datasets used later in the analysis. In order to obtain the elemental spectra, the work uses two approaches. The first one focuses on developing CR-composition-sensitive parameters, which are primarily reliant on charge deposits by TeV-muons in IC. The details of the composition-sensitive parameters developed and used in this work are given in Chapter 7. The second approach, which is a major part of this work, focuses on establishing a framework for an air-shower physics-inspired Graph Neural Network (GNN)-based approach, to estimate logarithmic mass for each air-shower observed at ICNO. Chapter 6 provides a brief overview of graphs and GNNs, and discusses the motivation to use GNNs at ICNO. The GNN-based approach maps the Cherenkov-tanks at IT and DOMs (PMTs with electronics) in IC, as the nodes of the input-graph for GNN. An air-shower physics-inspired

4 produced by astrophysical accelerators

5 Neutrinos and gamma-rays can be emitted as a consequence of CR-interaction with matter in the vicinity of the astrophysical source. Presence of neutrinos provide confirmed proof of hadronic progenitor. A multi-messenger study hence holds a huge potential. ICNO is also very well placed as a multi-messenger detector.

approach is used to introduce useful inductive biases in the GNN architecture. Beyond leveraging the low-level information contained within the shower footprint, acquired through the integrated operation of IT and IC, the work also capitalizes on the reconstructed shower observables, encompassing composition-sensitive parameters and others. For energy estimation of the air-shower, a GBDT-based approach is utilized. The intricate details of the mass and energy estimate are extensively elaborated in Chapter 8.

In order to obtain the elemental composition, a template-fitting approach is utilized. Using 10% of 2012's real-data (burnsample), the elemental fractional contribution as a function of reconstructed energy is presented in Chapter 9. Finally, a summary of the work, physics-interpretations, and outlook for this work are discussed in Chapter 10.

Appendix A, Appendix B and Appendix C present additional information relevant to the work. Appendix D presents the explorations by the author of this work in science-communication.

2 COSMIC RAY PHYSICS

Toutes choses sont dites déjà ; mais comme personne n'écoute,
il faut toujours recommencer.
(All things are already said; but since no one is listening,
you always have to start over.)

— André Gide [10]

Cosmic-rays (CRs)[1]are energetic nuclei of extraterrestrial origin that regularly bombard the Earth's atmosphere[2]. CRs are predominantly composed of approximately 90% protons, about 9% α-particles, with the residual fraction comprising heavier nuclei. With energies encompassing over 11 magnitudes in energy, from around a GeV to 100 EeV, CRs have fascinated and occupied physicists since their discovery in 1912. The astrophysical sources of these nuclei can accelerate them to extreme energies, which can be about a million times higher than those achievable by any accelerator on Earth. This was also the reason that in particle-physics infancy, CR-initiated cascades in the atmosphere (air-showers) served as the natural lab to find subatomic particles. The most prominent among them was the discovery of positrons by Carl D. Anderson in 1932 [13]. Over the decades CR-physics has not only stayed an exciting research area on its own but has also helped make leaps into other fields like archaeology [14], biology [15], extraterrestrial life [16], information field theory [17] and more. This chapter will briefly summarize our current knowledge of CRs, extensive air-showers, and their detection principles.

2.1 THE LANDSCAPE OF COSMIC RAYS

Since their discovery multiple space, balloon or earth-bound observatories have directly or indirectly measured CRs in specific energy ranges. The combined measurements span almost the entire energy range (eleven orders of magnitude) of CRs, providing the opportunity for researchers to develop a comprehensive understanding of CRs. Figure 2.1 presents an overview of the energy spectra obtained from the measurements done by multiple observatories. As is indicated in Figure 2.1, the flux of CRs falls steeply with increasing energies. For energies $\lessapprox 10^{15}$ eV direct detection of CRs is still possible by using satellite or balloon-borne detectors. However, for higher energies, the steeply diminishing flux necessitates a larger detection area. Current cost and

1 Robert Millikan is regarded as the nomenclator of CRs. He called them rays since he believed CRs were electromagnetic rays. Arthur Compton subsequently established that Cosmic-rays are 'particles' rather than 'rays'. However, the name stayed.

2 Like meteorites and stardust, CRs are direct samples from astrophysical sources. With an energy density of about 0.83-1.02 ev/cm^3 in the local interstellar medium [11], it sits very close to the magnetic field energy density of 0.25 eV/cm^3 (for a magnetic field of 3 µg). Hence, the interaction between CRs and magnetic field helps shape the magnetic fields and vice versa [12].

technological constraints limit such detectors/observatories to be earth-bound. However, an earth-bound observatory can only detect CRs indirectly[3], through the cascade of secondaries it produces an interaction with atmospheric nuclei. The spectrum for proton has also been shown separately for a few observatories. The spectrum for other high-energy messengers like gamma-rays, neutrinos (and anti-neutrinos), e^-/e^+, and anti-proton is also depicted. The intensity (y-axis) is scaled with the square of energy. The spectrum can be very well described by inverse power laws over large energy ranges. The power-law nature of the spectrum also hints non-thermal origin of cosmic rays[4]. A closer look at the spectrum shows that instead of a single power law, the spectrum show breaks in the behavior, at energies generally referred to as *"Knee"* and *"Ankle"* of CRs. These energies are characterized by the change in spectral-index γ of the parametrization of the all-particle spectrum given by

$$\frac{d\Phi}{dE_0} = \frac{dN}{dt \cdot dA \cdot d\Omega \cdot dE_0} \approx 1.8 \cdot E_0^{-\gamma} \frac{\text{nucleons}}{s \cdot cm^2 \cdot sr \cdot GeV/A} \tag{2.1}$$

From 10 GeV to 1 Pev (10^6 GeV), $\gamma \approx 2.7$[5]. From 10 PeV to 1 EeV (10^9 GeV), $\gamma \approx 3.1$. Above 1 EeV it shifts again to $\gamma \approx 2.6$, followed by a cutoff at around 100 EeV. The location of the *Knee* is generally accepted to be between 3-5 PeV [52, 53] and at around 4 EeV for *Ankle* [54–56]. Recently few other features have been noted between *Knee* and *Ankle*. The spectrum hardens at about 20 PeV [57–60], followed by two softening at $\approx$ 80 PeV [57, 58] and $\approx$ 317 PeV [54, 59–62] (generally referred as *Second Knee*). Very recent work has revealed another feature in the spectrum at about 13 EeV called *"Instep"* [18, 63], characterized by flattening of the spectra. The end of CR-spectrum is at about 60 EeV and is generally attributed to *Greisen–Zatsepin–Kuzmin (GZK)-cutoff* [64, 65]. While the factors contributing to these spectral-index shifts remain to be fully elucidated, they are generally interpreted as the manifestation of various aspects of CR-production, their source distribution, and their propagation through Galactic and extragalactic medium. The generally accepted explanations and other relevant details are:

- **Knee and Second Knee**: *Knee* has been generally suspected as a consequence of the limit in energy-quota of Galactic CR-sources [66] or the maximum energy of Galactic magnetic confinement [67–69]. Both effects grow linearly with the charge (Z) of the nuclei. An idea proposed by Bernard Peters [66], generally referred to as *Peters Cycle* predicts that we should see charge-dependent knee positions for the nuclei[6]. Hence, the *Knee* should be the limit for the lightest nuclei (proton), and the *Second Knee* should be the end of highest-Z galactic CRs[7]. Above the *Second Knee*, results from Pierre Auger Observatory and Telescope Array suggest

3 Space-bound experiments like JEM-EUSO, and POEMMA are planning to make indirect observation.
4 because the spectrum doesn't follow Planck's law
5 Below 10 GeV solar-modulation plays a crucial role in shaping the CR-flux.
6 The underlying reason being that rigidity(R) is the sole determinant of acceleration and propagation in models assuming collision-less diffusion in magnetized plasma. Here $R = \frac{E_{total}}{Ze} = \frac{pc}{Ze}$ where Ze is the charge of a nucleus, with total-energy $E_{total} = pc$. Rigidity determines the gyro-radius ($r_L = R/B$) in a magnetic field B. When an acceleration process reaches its limits and the gyro-radius becomes larger than the radii of the source, the corresponding rigidity (R^c) determines the characteristic energy where the drop will be seen. This is given by (using rigidity-equation) $E^c_{total} = Ze \cdot R^c$. Hence, the drop will be based on the charge (Ze) of the nuclei. The first evidence for this was observed at KASCADE [33].
7 Other ideas intended to describe the features depend on ideas of increasing galactic-leakage (and early onset of extra-galactic component) [70], single source models [71], Cannonball models [72], change in the

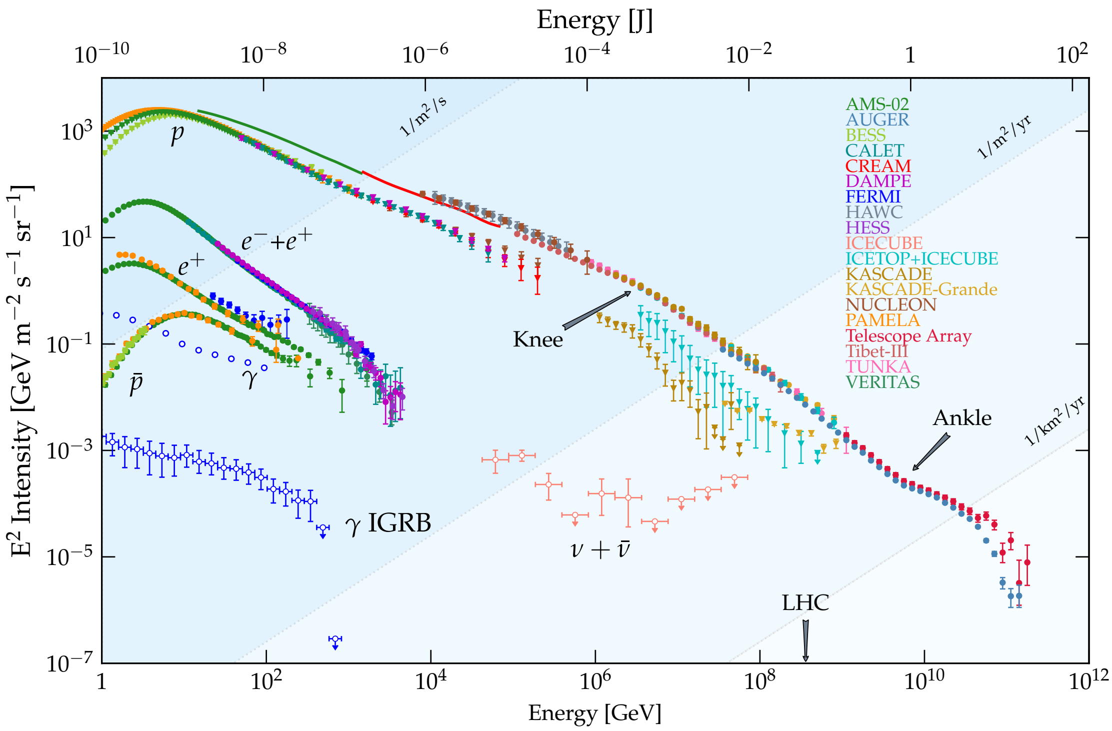


Figure 2.1: The energy spectrum of cosmic rays over eleven orders of magnitude in energy. The spectrum is scaled with the square of energy. The spectrum for other high-energy messengers like gamma-rays, neutrinos (and anti-neutrinos), e^-/e^+, and anti-proton is also depicted. All particle flux from Auger [18], HAWC [19], IceTop and IceCube [20], KArlsruhe Shower Core and Array DEtector (KASCADE) [21], KASCADE-Grande [22], NUCLEON [23], Telescope Array [24], Tibet [25] and TUNKA [26] is included. Proton data from AMS-02 [27], BESS [28], CALET [29], CREAM [30], DAMPE [31], IceTop and IceCube [20], KASCADE-Grande [32], KASCADE [33], NUCLEON [23], PAMELA [34] is included. Lepton data from AMS-02 [35], CALET [36], DAMPE [37], FERMI [38] and VERITAS [39]. Anti-proton data from AMS-02 [40], BESS [41] and PAMELA [42] is included. Positron data from AMS-02 [35], PAMELA [43], FERMI [44] is included. Total neutrino-flux data is from IceCube [45]. Gamma data is from FERMI [46]. The underlying data is from [47–50]. maximum attainable energy at Large-hadron Collider (LHC) is also marked. Plot adapted from [51].

a shift to lighter nuclei, between $10^{17.2}$ eV to $10^{18.33}$ eV [76]. This is seen as the transition from galactic to extra-galactic CRs. Above $10^{18.33}$ eV the composition becomes heavier again. Above 1 EeV, a scaling with A is seen again [77].

- **Ankle** : *Ankle* is generally considered as an onset of extragalactic CRs. Although The *Ankle* is generally explained by the Dip Model [78, 79]. It considers extragalactic CRs as protons with a mixture of 10% Helium, and the *Ankle* is a consequence of pair-production losses of protons on CMB photons. However, CR composition results disfavor this at Pierre Auger Observatory [80, 81][8]. The model also creates excess neutrino-flux expectation, which is in tension with experimental limits [82]. Other models which fit the observed spectra predict a gap between Galactic and extra-galactic CRs [83–85]. [86] proposes a model to explain the composition results and neutrino-flux using a single source class model, where the acceleration source is considered to be the core of Active Galactic Nucleis (AGNs). However, this model is also in tension with X_{max}[9] measurements at Pierre Auger Observatory. A clear picture of the reasoning behind the emergence of the *Ankle* might become available in the years to come.

- **GZK Cutoff**: GZK-cutoff [64, 65] is a steepening observed at the highest energies and marks the end of the CR-spectrum. It is considered a consequence of the interaction of the highest energy CRs with CMB photons. As a result of the interaction, only sources within a distance of $\approx$ 100 Mpc can contribute to the flux observed at the highest energies in CRs. The shape of steepening observed at the highest energies is however difficult to distinguish from if the sharp steepening was emerging from the limit of energy quota to produce CRs at the extreme energies. Hence, despite the observation of steepening in 2007 [55], the debate persists about whether the phenomenon is because of GZK-cutoff or constraints because of the energy-budget. A concrete test [87] of GZK presence is currently difficult because of limited event statistics at the highest energies.

[88] presents a detailed overview of other models used to explain multiple features in the CR-spectrum shown in Figure 2.1. Anti-protons and positrons, predominantly generated by the collision of CRs during their propagation in the intergalactic medium, are also shown in Figure 2.1. The majority of the electrons originate from astrophysical sources, however, their overall characteristics are substantially influenced during propagation. It is evident that, after more than 110 years CR research is mature but still open to producing surprises[10].

In order to describe and compare the CR-composition measured at observatories multiple flux models have been proposed. The ones used to make comparison in this analysis are H3a/H4a [90] model, Gaisser-Stanev-Tilav (GST)-3 [91] model and Global Spline Fit (GSF) [92] model. H3a is a flux model which assumes three populations of cosmic rays [90][11], an idea proposed by Hillas [93]. The three populations originate

propagation of Galactic CRss [73, 74] and more. All these models show a rigidity-dependent location of *Knee*. Any change in hadronic interactions was ruled out as the reason behind the origin of the knee [75].

8 It supposes that current extensions of particle-physics to highest-energies are correct i. e. hadronic-models (See Table 5.1).

9 Atmospheric depth at which the particle number in an air-shower (cascade produced because of CR interaction with air) is maximum.

10 The line is adapted from [89].

11 The model was proposed by T.K. Gaisser.

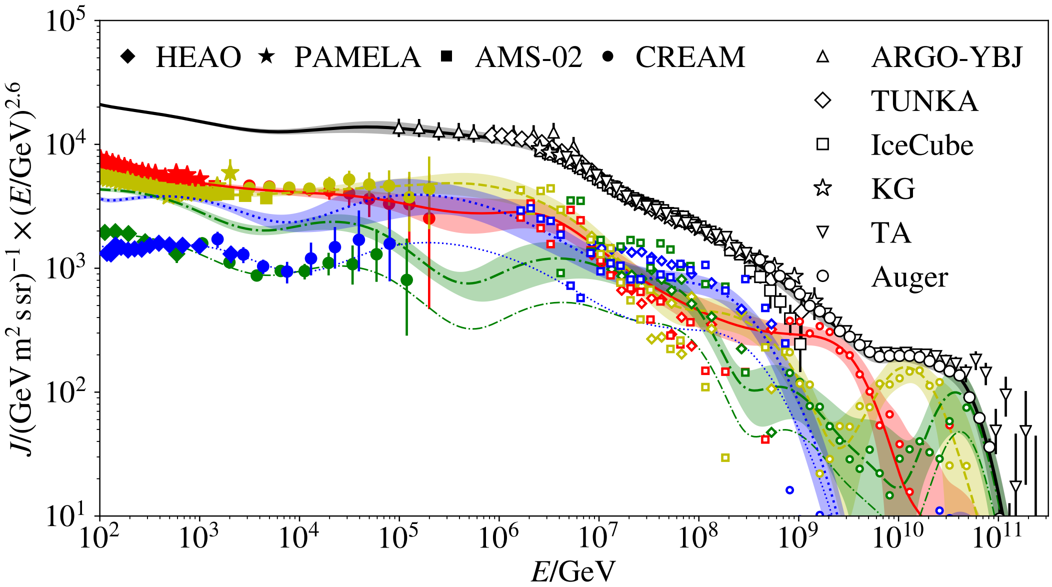


Figure 2.2: Total (black solid line) and individual primaries (Red solid line = proton; Yellow dashed line = Helium; Green dashed-dotted line = Oxygen group; Blue dotted line = Fe group) flux-parametrization by GSF Model. Bands around the lines show 1σ. Error bands represent combined statistical and systematic uncertainties. KG stands for KASCADE-Grande and TA stands fro Telescope Array. Illustration from [92].

from the acceleration of supernova remnants[12], an unknown galactic-source[13], and an extra-galactic component. The model groups the elements into five groups of nuclei, namely p(proton), He, CNO, Mg-Si, and Mn-Fe. The model assumes charge-dependent knee positions for the nuclei (*Peters Cycle*). In H3a the highest energy CRs are purely Iron, with the cutoff at the end explained by energy-budget constraints (instead of GZK-cutoff). In H4a, the extragalactic component is considered to be purely proton. The two different models were motivated by results from Auger [56] and HiRes [94] respectively. GST is pretty similar to H3a/H4a in most of its assumptions. However, it uses different characteristic rigidity cutoffs to fit CREAM data [95]. This results in a substantial prevalence of heavier primaries at the highest energies. GSF is a very recent effort to have data-driven parametrization of CR-flux and the corresponding mass-composition, with minimal dependence on any theoretical assumptions [92]. It utilizes measurements done using satellite (direct)[14] as well as from earth-bound (indirect)[15] CR-observatories. In addition to providing a smooth parametrization of total particle flux and flux from primaries, the model also established a common energy scale between different observatories. Figure 2.2 presents the smooth parametrization (total and individual primaries) from the Global Spline Fit (GSF) model.

12 with energy cut-off around the knee, proportional to the charge of nuclei
13 with the capability to accelerate particles to energies exceeding those from first-population
14 of individual elements
15 of mass groups

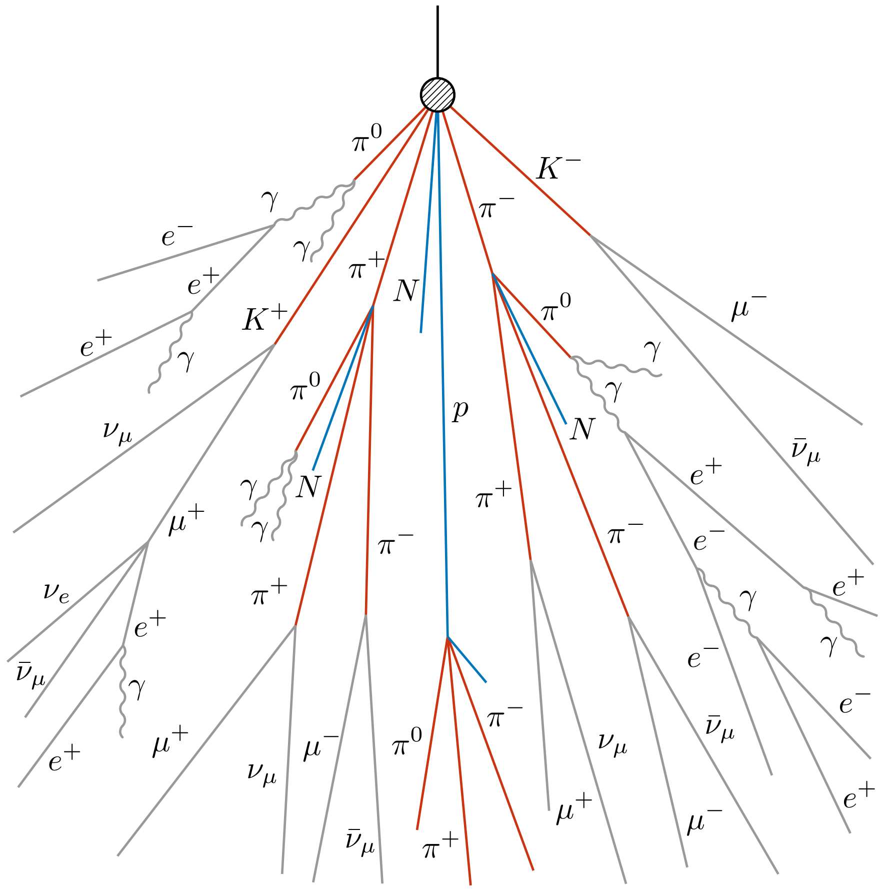


Figure 2.3: Illustration of Extensive Air-Shower, and its constituents. The cosmic-ray primary (depicted by the black line at the top) interacts with atmospheric nuclei producing a cascade of secondaries. The leptons, mesons, and baryons are represented using grey, red, and blue lines respectively. Illustration from [96].

2.2 EXTENSIVE AIR-SHOWERS

As depicted in Figure 2.1, for energies exceeding about 10^{15} eV, the flux of CR-primaries drop to about $1/(\mathrm{m}^2 \cdot \mathrm{year})$. At still higher energies the flux can be as low as $1/(\mathrm{km}^2 \cdot \mathrm{century})$. Making any quantitative assessment of CRs features at these energies hence requires large areas to collect sufficient flux. This requirement can be fulfilled by building large earth-bound observatories like IceCube Neutrino Observatory, Pierre Auger Observatory, and Telescope Array. However, on earth only indirect detection of CRs is feasible. This is because CR after bombarding Earth's atmosphere, interacts with particles in the atmosphere. On interaction with atmospheric nuclei, it produces a cascade of secondaries, which produces an avalanche of other particles. This is generally referred to as Extensive Air-Shower (EAS)[16]. An illustration of an EAS is depicted in Figure 2.3.

In a simplistic model, primarily charged ($\pi^{\pm}$) and neutral pions (π^0) are produced in the interaction of the CR-primary with atmospheric nuclei, as first-generation of secondaries in EAS. π^0 decay immediately into two photons (γ). These photons can further produce $e^{\pm}$ pairs. These can further create EM cascades by bremsstrahlung or ionizing other atmospheric nuclei. This avalanche of electromagnetic processes continues

16 "Extensive": Because at highest energies the footprint of the air-shower can be many kilometers wide.

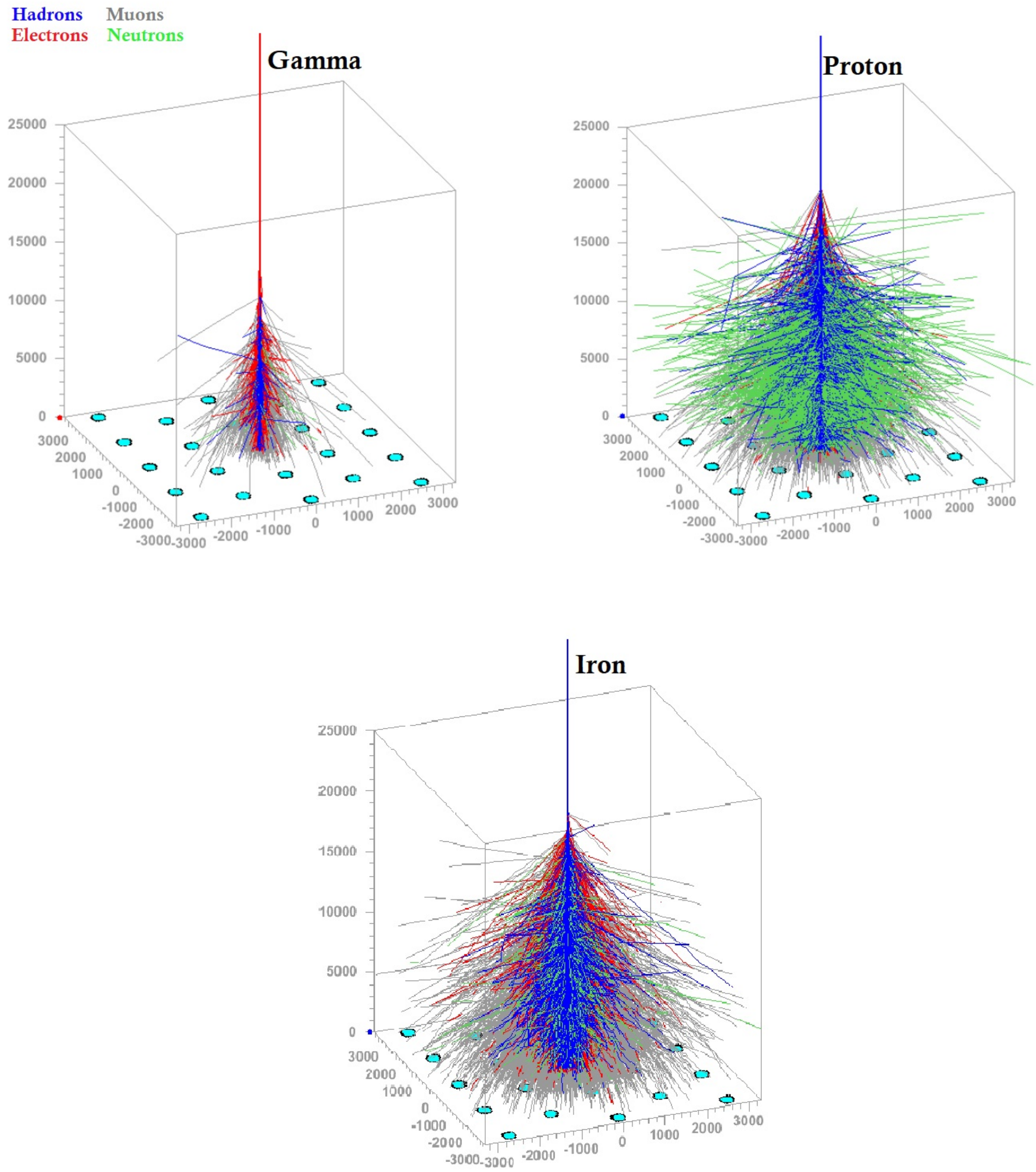


Figure 2.4: CORSIKA simulations of extensive air showers initiated by gamma, proton, and Iron primaries. These are vertical showers with an energy of 1 PeV each and have been adapted from animations of CORSIKA simulations generated at the Institute for Astroparticle Physics (IAP), Karlsruher Institut für Technologie (KIT) Germany.

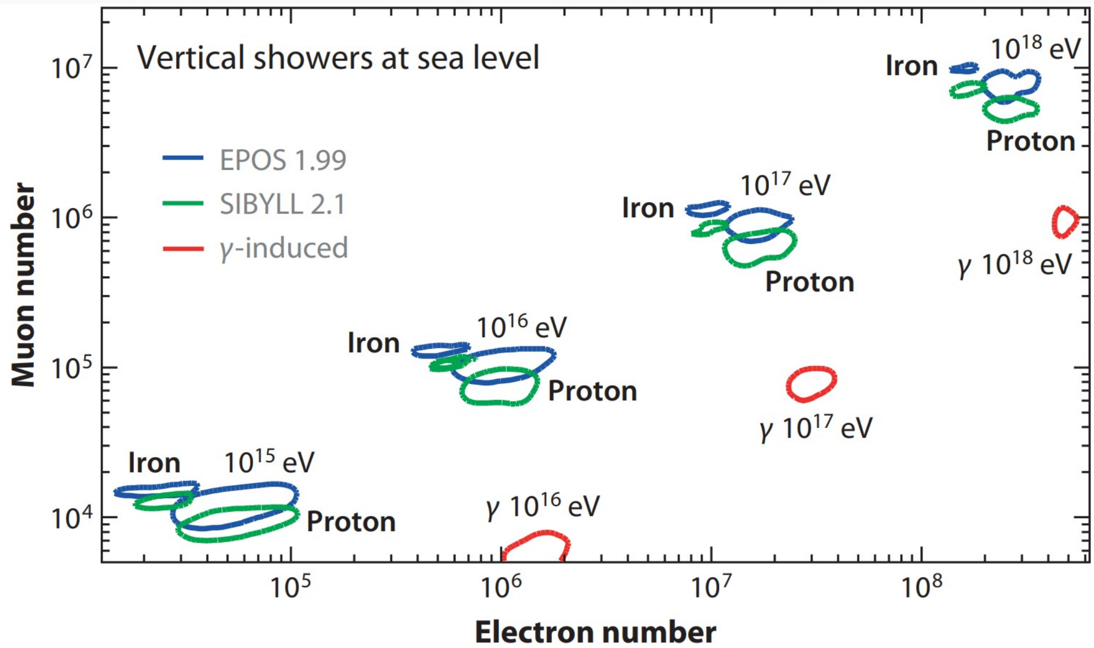


Figure 2.5: Correlation between abundance of muons and electrons in EAS at sea-level, for different hadronic-interaction models. An energy threshold of 0.25 MeV for γ and $e^{\pm}$, 0.1 GeV for muons and hadrons was chosen. Plot from [99].

feeding more photons and $e^{\pm}$ into the EAS, until the energy drops and reaches a level where ionization energy-loss dominates the radiative energy-loss. This constitutes the *Electromagnetic (EM)* component of EASs. On the other hand, $\pi^{\pm}$ produced in the first collision interact again before finally decaying[17] ($\pi^{\pm} \to \mu^{\pm} + \nu_{\mu}/\overline{\nu}_{\mu}$) to produce muons and neutrinos. The produced muons can decay ($\mu^{\pm} \to e^{\pm} + \nu_e/\overline{\nu}_e + \overline{\nu}_{\mu}/\nu_{\mu}$) into $e^{\pm}$ (hence feeding into the EM-component) and neutrinos. Since most of the muons in EAS are relativistic, they reach the sea level in spite of having short life-times[18]. This builds up the *muonic* component of EASs. Finally, the *hadronic* component of EAS is build up by the baryons produced during the cascade interactions. Understanding the intricate details of first interaction and baryon-interactions is a continuing research endeavor, and is the predominant source of particle-multiplicity uncertainties in hadron-induced cascades [76, 97]. EM fluorescence and radio emission are also produced in EAS. [98] provides a comprehensive overview of the intricacies of EAS.

For most CR-analysis from EAS measurements, detailed MC-simulations need to be performed to understand the response of EAS to detector components. For the case of ICNO, the details of the MC-simulations done using CORSIKA will be given in Chapter 5. It is essential to simulate EAS for making quantitative statements on energy and composition of CRs, as well as other shower observables. As can be seen in Figure 2.4, even at the same energies, depending on the choice of the primary the EAS response can look significantly different to each other. It is important to notice that the reliability of these simulations is currently among the main contributor to systematic uncertainties in interpreting cosmic ray data at most observatories [20, 33, 56, 100, 101].

17 when $E_{\pi} \lesssim 30$ GeV
18 because of time-dilation

Figure 2.5 presents the correlation between abundance of muon and electrons in EAS. For the case of hadronic primaries i. e. proton and Iron, it is evident that the muon number is a good discriminator between the two. Since this analysis is focused on studying the CR-composition at ICNO, the following text will discuss a few relevant details of muons in EAS.

2.2.1 *Muons*

Muons are considered as a powerful discriminator between CR-primaries [102]. In addition to that muons can be observed with a duty cycle of almost 100%. This is much greater than the almost 15% duty cycle for another prominent composition-sensitive observable i. e. depth of shower maximum (X_{max}) [97] of the EM-component. It is however very important to notice that X_{max} has much smaller systematic uncertainties than those from muon number [103]. ICNO currently lacks the capability to measure X_{max}[19]. In contrast, prior work at ICNO [20] has successfully used charge deposit by TeV muons to estimate CR composition at ICNO. In the same spirit, the following text provides details of HE-muons in EAS, which will be relevant for discussions later in this thesis.

2.2.1.1 *Multiplicity: High-energy Muons*

The approximation for the number of high-energy (HE) muons above threshold energy of E_μ was parametrized by J.W. Elbert [104, 105][20] and is given by:

$$\langle N_\mu(>E_\mu, E_0, A, \theta)\rangle \approx \frac{14.5\mathrm{GeV}\cdot A}{E_\mu \cos(\theta)}\left(\frac{E_0}{A\cdot E_\mu}\right)^{0.757}\left(1-\frac{A\cdot E_\mu}{E_0}\right)^{5.25} \tag{2.2}$$

where A is the mass of the primary of energy E_0, incident at a zenith-angle of θ. The first term, $\frac{14.5\mathrm{GeV}\cdot A}{E_\mu \cos(\theta)}$, encapsulates the probability associated with the decay of pions and kaons. The second term, $\left(\frac{E_0}{A\cdot E_\mu}\right)^{0.757}$, addresses the increase in muon multiplicity with an increase in energy and includes the scaling from superposition approximation (treating incident nuclei as A independent nucleons, each of energy E_0/A). The final term addresses the vanishing inclusive cross section for pion production, close to E_0. The muon number expectation, given by Equation 2.2, for proton and Iron initiated shower for muons with energy greater than 800 GeV (the average energy required to reach the bottom of the IceCube detector) for different true MC-energy is shown by solid-lines in Figure 2.6(a)[21]. The dashed lines indicate the expectation without the final term. Figure 2.6(b) represents the muon multiplicity for different values of the muon-energy threshold for primaries of 4 discrete MC energies. It can be seen that for $\mathcal{O}(\mathrm{PeV})$ energy showers the muon-multiplicity for the p-initiated shower becomes

19 The planned upgrade of the IceCube Neutrino Observatory (ICNO), known as IceCube-Gen2, includes the deployment of surface radio-antennas. This will provide the capabilities to measure X_{max} at ICNO in future (for $E > 10^{16-17}$ eV). Read Section 3.2 for details into IceCube-Gen2.

20 The parametrization was obtained using MC simulations and uses superposition principle for nuclear primaries (since the number of muons is not effected by nuclear fragmentation model [106]). Similar results using MC simulations were obtained by T.K. Gaisser and T. Stanev [107], and C. Forti [108] et al. Analytical approach [109] also gives comparable results.

21 A comparison of the parametrization with the expectation from MC results for protons and iron initiated EAS is shown in [110] and shows remarkable overlap over the whole energy range (above threshold).

greater than that for the Fe-initiated shower. This leads to the expectation of seeing more high-energy muons carrying a large fraction of primary energy, at low energies. A parameterization for low-energy muons has been presented in [12]. A recent work [111] generalizes the Elbert formula to obtain the distribution of slant depths over which the muons are produced.

2.2.1.2 *Energy Loss in Matter*

It is crucial to understand the energy loss of muons in matter, since they are among the sole survivors in EAS[22] with enough penetrating-power to deposit signal after traversing more than a kilometer of Antarctic Ice[23]. Muons can lose energy in the matter by multiple channels: namely ionization and atomic excitation [113–115]; bremsstrahlung [116, 117], electron pair-production [118], photo-nuclear interactions [119, 120]. The mean energy loss or stopping power can be parameterized quasi-linearly by:

$$\left\langle -\frac{dE}{dx} \right\rangle = a(E) + b(E) \cdot E \tag{2.3}$$

where $a(E)$ denotes the ionization losses. As can be seen from Figure 2.7, it exhibits a weak logarithmic increase with muon energy. $b(E)$ is a combination of the radiative losses (bremsstrahlung, pair-production, photo-nuclear interactions). For energies exceeding ≈ 1 TeV, the radiative losses in-ice dominate the ionization losses.

IONIZATION LOSS For moderately relativistic charged particles, the mean rate of ionization-loss is given by the Bethe-formula [121, 122]:

$$\left\langle -\frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \tag{2.4}$$

where z is the charge of the particle[25], Z and A are the atomic number and mass number of the matter being traversed. $\gamma = 1/\sqrt{1-\beta^2}$ is the Lorentz factor, with $\beta = v/c$ denoting the velocity (v) of the particle in comparison to the speed of light in vacuum (c). m_e is the electron's rest mass, and $K = 4\pi N_A r_e^2 m_e c^2 \approx 0.3071 \mathrm{MeV cm^2 g^{-1}}$ where $r_e = e^2/4\pi\epsilon_0 m_e c^2$ is the classical radius of the electron and N_A is the Avogadro's Number. I is the material-dependent mean ionization-energy and T_{max} is the maximum transferable kinetic energy to the electron (in a single collision), given by:

$$T_{max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma m_e/m + (m_e/m)^2} \tag{2.5}$$

with m being the incident particle's mass[26]. $\delta(\beta\gamma)$ is a correction factor for the density effects in the matter [123, 124].

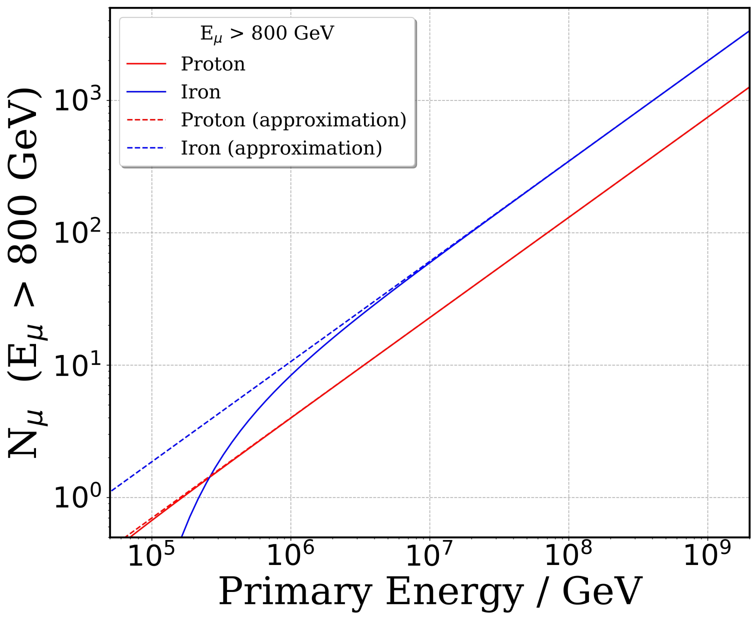
---

22 with neutrinos

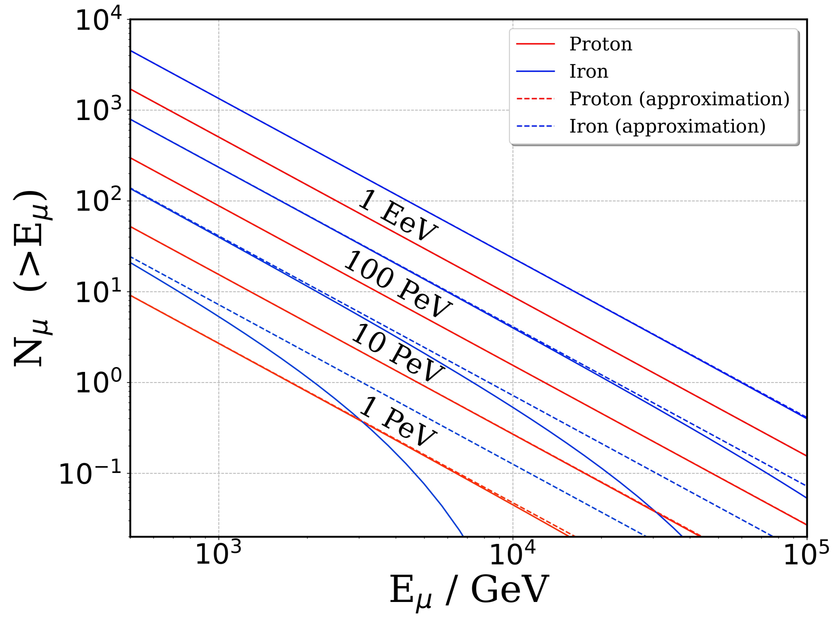
23 This is in contrast to other constituents like $\pi^\pm$, K, $e^\pm$ or hadrons. $\pi^\pm$ and K[24] are unstable particles and have a much higher probability to interact than decay, in a denser medium. The number of surviving $e^\pm$ and hadrons in a denser medium is greatly reduced with increasing depths because of cascade processes and nuclear interactions respectively.

25 For Muons: $z = 1$

26 here $m = m_\mu$



(a) Multiplicity (for $E_\mu > 800$ GeV) as a function of primary energy.



(b) Multiplicity as a function of muon-energy threshold.

Figure 2.6: Muon multiplicity expectation from Elbert formula (Equation 2.2) for p and Fe initiated showers. The dashed lines represent the expectation without the last term i. e. $\left(1 - \frac{A \cdot E_\mu}{E_0}\right)^{5.25}$ in the formula.

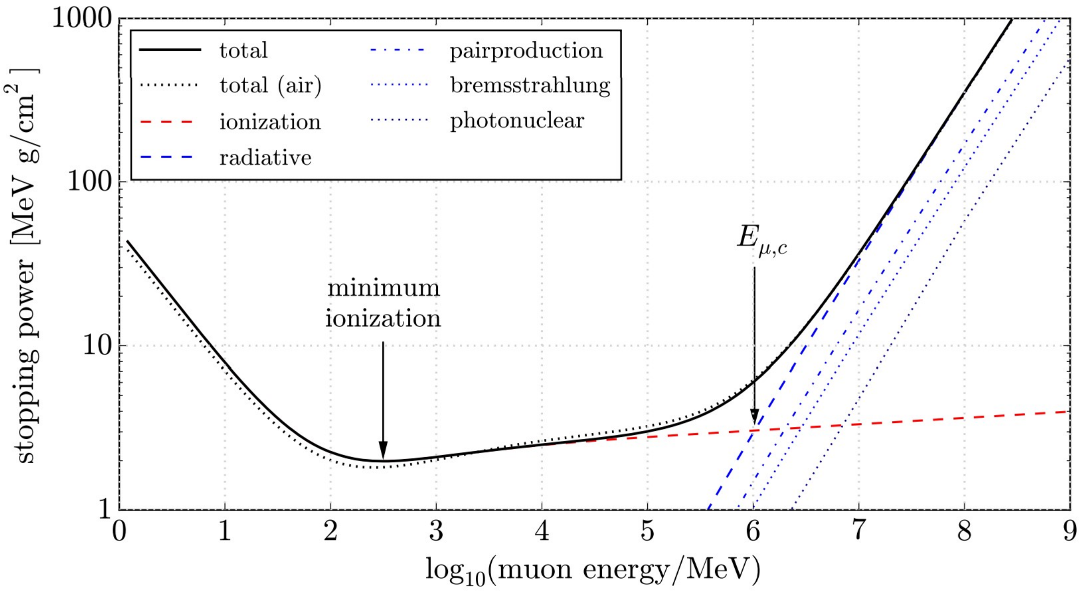


Figure 2.7: Stopping Power or mean energy-loss of muons in ice and the contributions from different underlying processes, as a function of muon energy. Plot from [112].

RADIATIVE LOSSES As can be seen from Figure 2.7, for energies in excess of about 1 TeV the radiative losses dominate the energy-loss landscape. The underlying constituents of radiative losses are:

1. **Bremsstrahlung**: In the presence of an electromagnetic field of another charge (matter), a muon (or a charged-particle) gets decelerated and releases radiation. This is termed as *bremsstrahlung*[27]. It has significant contributions from the elastic scattering of the muon with the atomic nuclei. Two additional contributions from the inelastic scattering of muons with electrons of atomic nuclei, and nuclear excitation constitute as the secondary contributions. In contrast to the ionization losses, bremsstrahlung is characterized by large fluctuations in its energy deposits (by electromagnetic cascades). The cross-section scales with fractional energy loss (ν) as ν^{-1} [122].

2. **Pair-Production**: High-energy muons can also produce virtual photons which can create e$\pm$ pairs in the presence of the electromagnetic field of nuclei. The cross-section scales with ν^{-2} to ν^{-3} [125], with influences of nuclear screening [126] and its finite extent [118]. This leads to softer energy losses than bremsstrahlung, and the deposits are almost continuous [122].

3. **Photo-nuclear**: It is a subdominant energy loss process that is caused because of inelastic scattering of muons with nuclei. Even though it is a subdominant process, since the energy losses are hard, the electromagnetic and hadronic cascades produced can cause large fluctuations within a medium [116, 122].

27 The etymology of the term traces its roots to the German language. Specifically, *Bremse* = "a brake," and *Strahlung* = "radiation." It was coined by Arnold Sommerfeld and was later adopted into English by Arthur Edward Ruark and Harold Clayton Urey. It is also referred to as *free-free* radiation sometimes. This is because it is identical to the phenomenon to transitions between unbound states of the electron in the field of the nucleus (atomic physics).

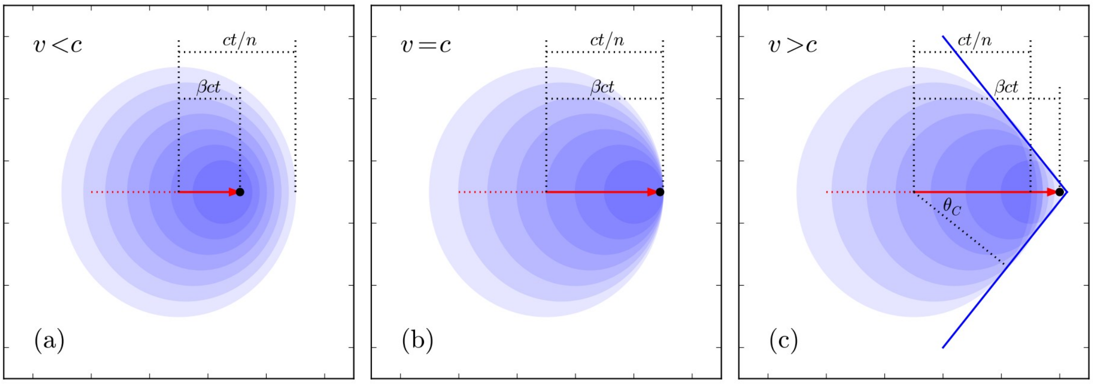


Figure 2.8: Wavefront (based on Huygens-Fresnel principle [131]) created by the propagation of a subatomic particle in a dielectric medium, for a) phase-velocity less than the speed of light b) equal to the speed of light and c) greater than the speed of light. For speeds exceeding the speed of light a light cone at a characteristic angle (based on the particle's velocity and index of refraction of the medium) i. e. Cherenkov angle (θ_c) is produced. Plot from [112].

2.2.1.3 *Cherenkov Radiation*

First observed by the Soviet scientist Pavel Alekseyevich Cherenkov (1958 Nobel Prize in Physics along with Ilya Frank and Igor Tamm) in 1934 [127][28], Cherenkov radiation is electromagnetic radiation emitted when a charged subatomic particle propagates through a dielectric medium with a speed greater than phase velocity of light in that medium. The charged particle polarizes the medium, leading to an emission when polarized atoms return to their ground state [130][29]. The emission happens at a characteristic angle, θ_c, depending on the index of refraction (n) of the medium and is given by:

$$\cos(\theta_c) = \frac{1}{n\beta} \tag{2.6}$$

where $\beta = v/c$ gives the ratio of the particle's speed (v) to that of light in vacuum (c). Figure 2.8 compares the Cherenkov emission for three cases, for a) phase-velocity less than the speed of light b) equal to the speed of light, and c) greater than the speed of light. Constructive interference happens when the speed of the charged subatomic particle traversing the dielectric medium exceeds the phase velocity.

The phenomenon is crucial in enabling the detection of EASs and neutrinos with IceTop and IceCube[30]. The South-Pole Antarctic ice with its thickness of about three kilometers, serves as an efficient and economical medium for the generation of Cherenkov radiation. Because of the directed nature of Cherenkov-radiation, it also plays an important role in directional reconstruction at IceCube Neutrino Observatory(ICNO).

28 The existence of Cherenkov radiation was theoretically predicted at the turn of 20th century by Oliver Heaviside [128] and Arnold Sommerfeld [129], but was largely dismissed.
29 If the speed is less than the phase-velocity, the emission interferes destructively.
30 The particles traversing IceTop as well IceCube, emit Cherenkov light which is collected by detectors at IceTop and IceCube.

ICECUBE OBSERVATORY 3

IceCube Neutrino Observatory (ICNO), a successor to AMANDA, is the world's first giga-ton and a cubic-kilometer astroparticle detector located close to the geographic South Pole. It detects particles from cosmic ambassadors in the endeavor of understanding the dynamics of such sources. This allows analysis in a variety of research areas such as cosmic rays and neutrino physics and more. The following text will detail the rationale behind the construction of ICNO, the building blocks of the detector (DOMs) and the component arrays of the observatory (ICL, IT and IC). Finally, a brief summary of the next-generation instrument (IceCube-Gen2) will be given.

3.1 ICECUBE : THE CUBIC-KM EYE ON THE COSMOS

Neutrinos were considered to be an interesting probe into stellar atmospheres as early as 1960 [132–134]. Their charge-less nature along with very-rare interaction with the interstellar medium at source [135] meant that they travel in straight trajectories carrying the most unperturbed information of their sources. However, the difficulty in detection[1] accompanied by low astrophysical flux pushed scientists towards building a large neutrino observatory [136–139]. The efforts started by the people at Deep Underwater Muon and Neutrino Detector (DUMAND) [140] persuaded and paved way for other groups to use photomultipliers for particle detection [141–145]. Antarctic Muon And Neutrino Detector Array (AMANDA), which operated from 2000 to 2009, finally heralded the way for astronomy using cubic-kilometer observatory like **ICNO**.

3.1.1 *Building IceCube - The Rationale*

The results from Kolar Gold Fields neutrino experiment in India and the results from East Rand mine in South Africa laid the first steps towards establishing the possibility of searching astrophysical neutrinos using detectors buried deep underground [146, 147]. This was followed up by MACRO [148] and Frejus experiment [149]. It was suggested [134] that a bigger Cherenkov detector is required. After the failures of experiments like DUMAND, the initial encouraging results from AMANDA showed the feasibility of building such a detector at the South Pole and finally paved the way for constructing the IceCube Neutrino Observatory (ICNO).

IceCube Neutrino Observatory (ICNO), a cubic-kilometer eye on the cosmos, was built to enable us to detect the extraterrestrial source of neutrinos, and hence provide a complimentary view of the sky using particles. Neutrinos because of their charge-less and minimal-interaction nature can propagate and reach us without any deflection (in contrast to cosmic rays) and from the depths of the universe which is generally opaque

1 because of low cross-section

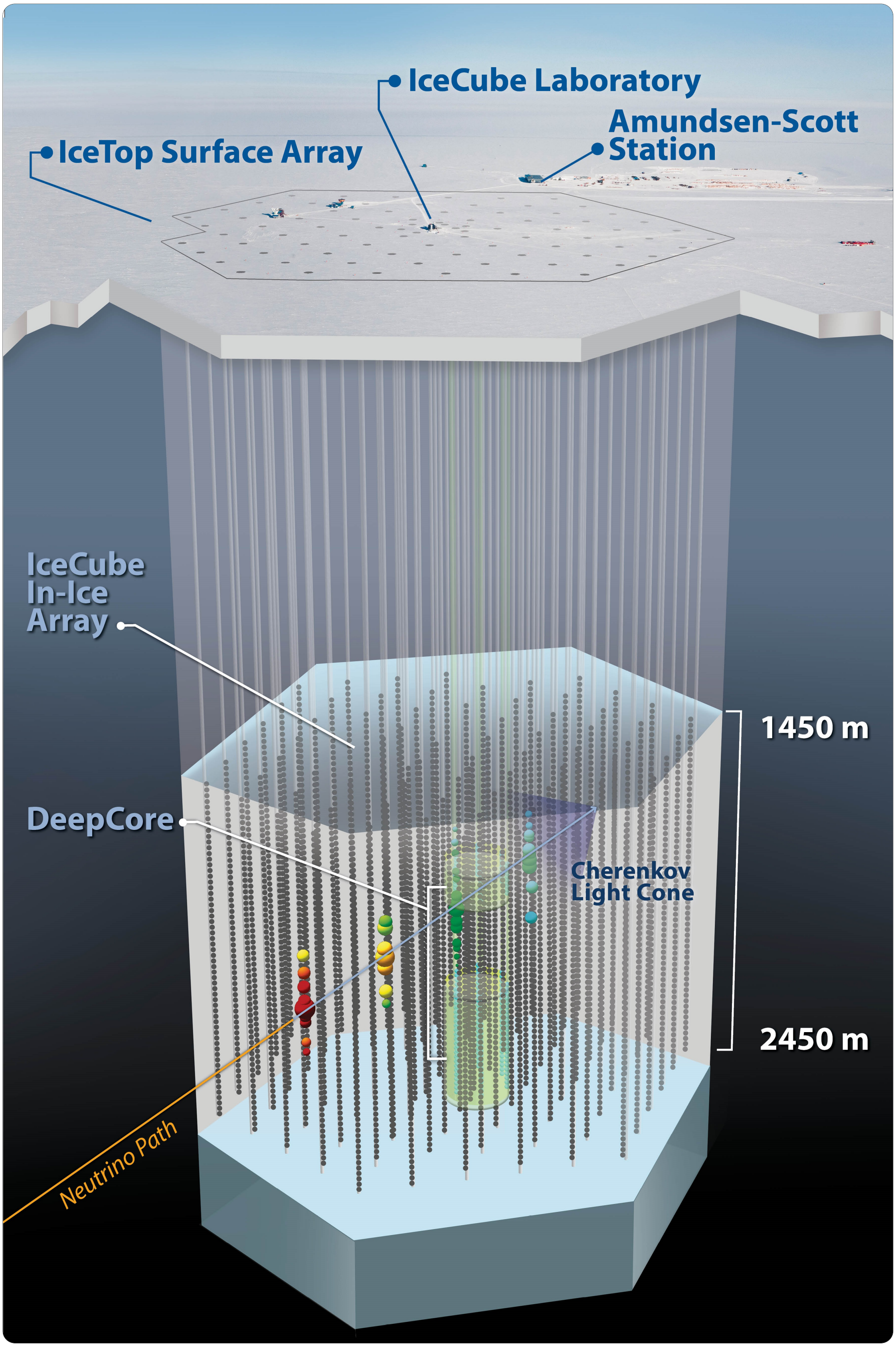


Figure 3.1: A Schematic of IceCube Observatory. An example neutrino event is also depicted. Photograph or Illustration Credits (PC): IceCube Collaboration (ICoII).

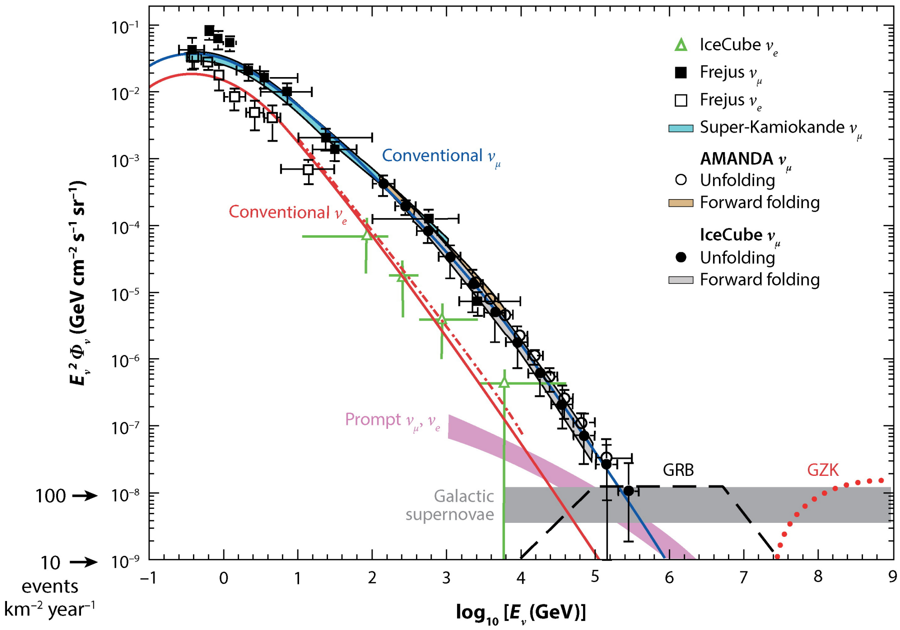


Figure 3.2: Anticipated astrophysical neutrino fluxes at ICNO, and comparison with other experiments. Plot from [159].

to other cosmic messengers (like photons). The limited flux and small cross-section meant that a large detection volume was required (in order to have enough statistics). Initial estimates (Figure 3.2) predicted about $\approx$ 10-100 ν_μ events per year [150]. It is expected that, at the highest energies, the acceleration at the astrophysical sources and their subsequent interaction with the surrounding medium at sources can produce the HE ν-flux. Hence, the HE neutrinos at ICNO might also help to map the yet unidentified sources of galactic and extra-galactic CR sources and the origins of GZK limit [151–154]. An important observation that motivated the construction of ICNO was the observation of a CR excess from the Galactic Plane. At the time, this hinted towards neutrons as the primaries[2]. This led to the expectation of about 20 $\bar{\nu}$s [155]. Another observation by HESS of Supernova Remnant (SNR) RX J1713.7 - 3946 (G347.3 - 0.5) observed an increase in the direction of molecular clouds [156]. It was suggested that these could be the potential location of proton acceleration [157]. The possibility of performing extensive studies in the galactic and extra-galactic regime pushed towards building a first-of-its-kind cubic-kilometer observatory [158]. Chapter 4 will detail more scientific goals which were laid out by ICNO and were successfully carried out or are ongoing.

3.1.2 *Coordinate System*[3]

ICNO follows a local right-handed Cartesian coordinate-system, centered near the middle of the in-ice array. The origin is approximately at a depth of 1947 m, almost in the center of the in-ice instrumented volume. The y-axis is aligned with Prime Meridian (pointing towards Greenwich, UK). The x-axis points 90 degrees clockwise from

2 A neutral primary has a greater chance of producing the directed signal.
3 *"The study of geometry is a petty and idle exercise of the mind, if it is applied to no larger system than the starry one." - Henry D. Thoreau [160]*

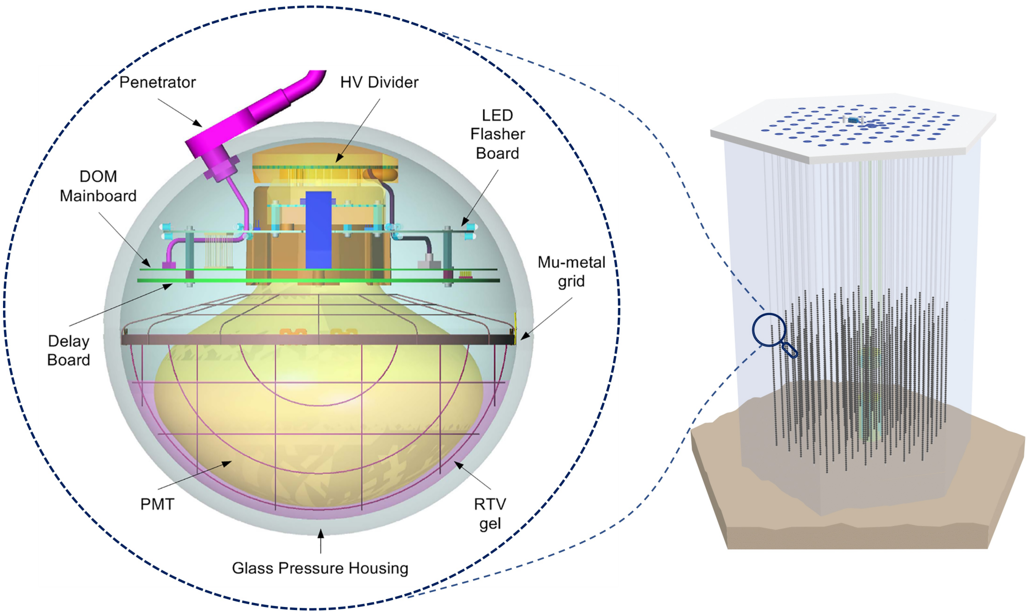


Figure 3.3: Schematic of Digital Optical Modules (DOMs) used at ICNO.

the y-axis and the z-axis points up (normal to the earth's surface)[4]. A volume of ice with high dust concentration is also located in-ice (In ICNO coordinates: Between -150 m to -50 m. See Figure 3.9 for details.). Since the glacier on which ICNO is located is slowly moving, the coordinate system shifts by about 10 m per year (along the 40° west longitude meridian).

3.1.3 *Digital Optical Module - DOM*

ICNO uses Digital Optical Modules (DOMs) in order to collect the Cherenkov radiation of traversing particles. For EAS detection at ICNO, the DOMs collects Cherenkov emission from particles traversing the surface-detector (IceTop), and the in-ice detector (IceCube). The deposit at IceTop is primarily by the EM-component of EASs and the low-energy muonic component in EASs. However, the in-ice DOMs collect light primarily from TeV muons. A DOM consists of a 10"-diameter downward-facing Photo-Multiplier Tube (PMT) [162] enclosed in a spherical glass housing. To serve as an autonomous unit, the PMT is accompanied by other electronics responsible for data acquisition, digitization, control, calibration, and more. In addition to this, the DOM also has communication and data-transfer channels to the neighboring DOMs, as well as the data-acquisition systems on the surface. The design of DOMs is engineered to ensure long-term reliability while withstanding the extreme cold conditions of the South Pole and their ability to survive the pressure of cubic kilometers of ice, while providing precise charge measurements with nanosecond time resolution. The schematic representation of a ICNO DOM's mechanical layout, complete with labeled components, is depicted in Figure 3.3. DOMs are equipped with multiple digitizers that have overlapping dynamic ranges and varying sampling speeds. This enables the detection of particles with energies spanning multiple orders of magnitude. As a result, waveform storage is possible for

4 The coordinate system uses the same directions as ICNO's antecedent experiment AMANDA [161]

amplitudes ranging from 1mV to approximately 2V and widths from 12 ns to around 1500 ns. Brief description about each DOM component is as follows (details in [163]):

- **Glass Sphere**: It has an outer diameter of 13" and a thickness of 0.5". It is built using two hemispheres that meet at the equator and are held together using an aluminum waistband with rubber gaskets.

- **PMT**: Each DOM consists of a 10" - diameter Hamamatsu PMT (Peak quantum efficiency ≈ 25 % . DC PMTs have higher quantum efficiency of ≈ 34 %). The PMTs are operated at a gain of 10^7 and are maximally sensitive in the wavelength range between 300 nm - 650 nm.

- **Gel**: Each downward facing PMT is anchored using high-resilience silicone gel. It has a refractive index of 1.41. In addition to providing structural support to the DOM constituents, it also serves as a good optical coupling agent. At normal incidence, the transmission ranges from 65% at 300 nm to 91% at 340 nm and reaches 97% at 400 nm.

- **Magnetic Shield**: It order to prevent the effects of ambient magnetic-field on the collection efficiency of PMTs[5], the PMTs are embedded in a Mu-metal[6] cage.

- **Mainboard**: The mainboard is the central data-acquisition unit of the DOM. The design [164] constitutes multiple components that carry out multiple processes, including controlling high voltage power supply, flasher-board, pressure, and temperature sensors. It is also responsible for waveform digitization and provides a computational interface for gain calibration, compression, storage, and data packing. It also serves as the channel which allows the DOM components to communicate with the ICL on the surface.

- **Flasher Board**: The flasher board in most DOMs consists of multiple single-colored LEDs used for calibration [58, 165]. These can be used to emit controlled flashes deep under the Antarctic ice. The flashes can then be detected by other neighboring DOMs, which can be used to calibrate the charge and time response of the DOMs. In addition to that it can also help improve the preciseness of DOM locations and improve our knowledge of the optical properties of ice. It can also help record any movement of the South Pole glacial ice over the years.

- **Cable Connections**: Three wire-pairs come out of the DOM. One of these pairs is responsible for powering the DOM and carrying a bidirectional digital communication pathway. This finally terminates at the ICL. The other two pairs connect to the DOM directly above and below. This is depicted in Figure 3.4. The connection to the neighboring DOMs is used to determine if the condition for Local Coincidence (LC) is met or not (details in Section 3.1.7).

- **High Voltage Supply and Divider**: Each PMT consists of a high voltage divider circuit with a maximum capability of 2047 V. It has an efficient power consumption of 300 mW at full load.

5 It can cause 5-10 % decrease in collection efficiency.
6 It is a ferromagnetic alloy of nickel and iron, with small quantities of copper and chromium or molybdenum.

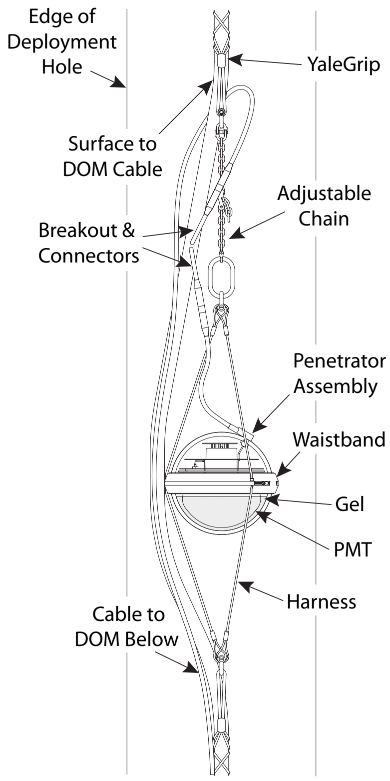


Figure 3.4: DOM on an in-ice cable. Illustration from [163].

3.1.4 *IceCube Laboratory*

IceCube Laboratory (ICL) is a building on the surface of ICNO and serves as the central control system of the whole observatory. The cables from all the detector components are finally channeled to ICL. ICL hosts powerful computing servers which are used to send real-time alerts to other observatories around the world[7], for higher-level data acquisition, event filtering, detector monitoring, and more. An outline of ICL is depicted in Figure 3.5.

7 This is very important to look for multiple cosmic messengers from the same astrophysical source. For details read Section 4.3.

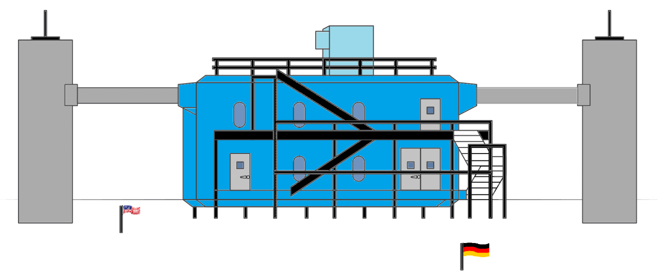


Figure 3.5: Illustration of IceCube Laboratory. PC: Adapted from ICoII.

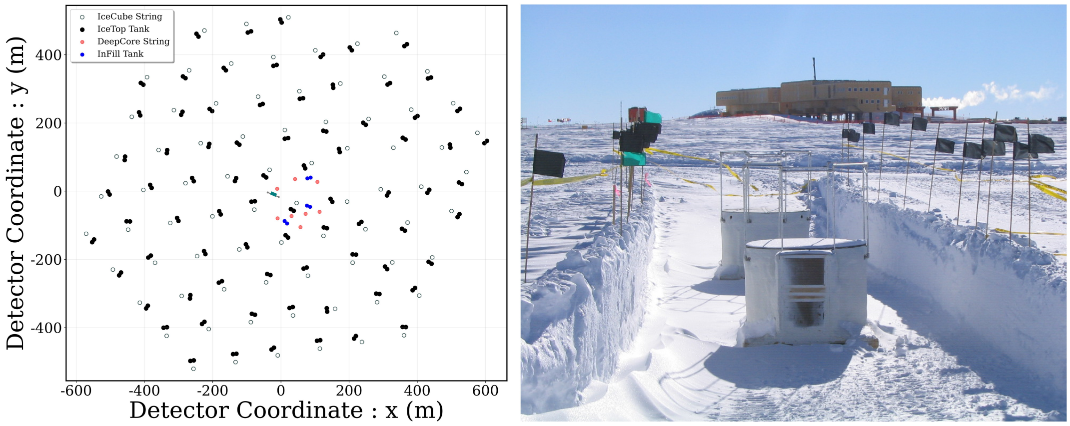


Figure 3.6: **Left:** A schematic of IceTop array. The IC strings, DC strings, and ICL(sea-green colored structure in the middle) locations are also depicted. **Right:** An IceTop station consisting of two ice-Cherenkov tanks, during deployment. PC: IColl.

3.1.5 *IceTop*

IceTop (IT) is the surface array of the ICNO. It is approximately a 1 km^2 air-shower detector with a focus on studying cosmic rays around the knee in the CR-spectrum i. e. in the transition region between galactic & extragalactic origin of cosmic rays. Recent work has been able to lower the energy threshold to 250 TeV, also bridging the gap between direct and indirect detection of cosmic rays [166]. IT is located at an altitude of 2835 m (approximate yearly-average atmospheric depth = 690 g/cm^2). It consists of 81 stations, with each station consisting of two cylindrical ice-Cherenkov tanks. The two tanks in a station are about 10 m apart. Each tank is equipped with two DOMs running at different gains in order to increase the dynamic range. The stations are arranged in a triangular grid with an approximate spacing of 125 m, generally located on top of IC strings. This is depicted in Figure 3.6. The denser in-fill array allows for finer shower sampling as well as lowering the energy threshold for air-shower triggering [167][8]. The tanks lie on a slightly tiled plane. The lowest tank lies approximately 1945 m above the center of ICNO, whereas the highest is approximately 1950 m above it. This is depicted in Figure 3.7. The whole array is covered by a few meters of the South-Pole snow. The snow height above the tanks is generally not the same and also changes over the years (because of snow-drifting by wind and other environmental factors). The change in snow height is one of the major sources of systematic error in calibration (and hence quantifying observed signal) at IceTop-tanks [169]. Figure 3.8 depicts the variation in snow height on IT-tanks over the years. Cubic interpolation is used to get a smooth mapping in the area between the IT tanks.

IT was built with three science goals in mind, namely: veto, composition, and calibration of IC. IC was primarily built to detect up-going (at the South Pole) neutrinos. This was done in order to benefit from the automatic vetoing of atmospheric muons (the major background), using earth's surface [171–173]. IT in conjugation with IceCube (IC)

8 A primary of approximately 300 TeV is enough to trigger hit three IceTop stations. In-fill reduces this to approximately 100 TeV [168].

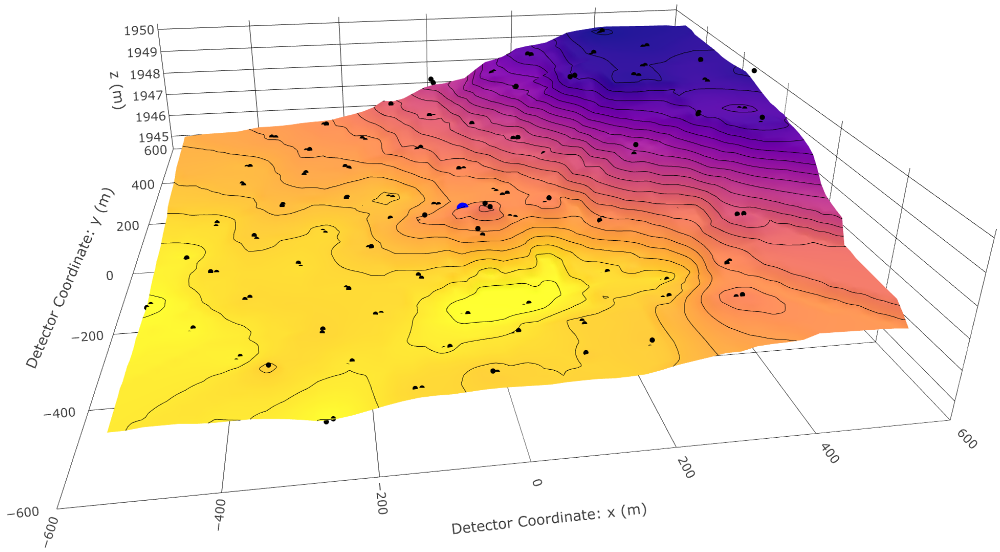


Figure 3.7: Height variation among IT tanks (black-dots), without snow coverage. Inverse Distance Interpolation [170] is used to estimate the IT-surface (from the sparse measurements at IT-tank location), and the contours. ICL is denoted by the blue-circle in the center.

(the in-ice array of ICNO) is an ideal detector to study the mass composition of cosmic rays. In addition to this it has also been used to study various other areas like testing hadronic models [174], cosmic ray anisotropy [175–180], the density of GeV muons [181], high-energy neutrons [182], high-p_T muons [112], PeV Gamma-Rays [183], solar-phenomenon [184, 185], seasonal variation of high-energy atmospheric neutrinos [186], veto for neutrino detection [187–189] and more. Although, IT-array was planned as an air-shower detector it has also proved to act as a very useful veto for neutrino detection [187–189]. This has enabled neutrino detection in the down-going direction too. Ongoing work is trying to improve the vetoing capabilities even further [190]. As a cosmic-ray detector IT has served as a unique probe into cosmic rays around the knee-region of CR-spectrum [20, 58, 166]. IceTop is also located at an atmospheric depth closer to the mean shower maximum of a 100 PeV proton shower (and a Fe-shower for higher energies) [191][9]. It can also help calibrate the in-ice array of ICNO by providing complementary measurements for muon-bundles at the surface. As mentioned earlier IT has already proved its potential in many more research areas [112, 174–186]. Section 4.2 will give a detailed overview of the variety of CR analysis done ICNO.

3.1.6 *In-Ice Array*

IceCube (IC), the in-ice array of ICNO consists of 86 strings, with 60-DOMs/string located in glacial depth between 1450 m to 2450 m. The DOMs in a string are situated 17 m apart from each other. As depicted in Figure 3.6 and in Figure 3.1, the strings are located in approximately a hexagonal array. The average separation between two neighboring

9 The shower maximum is a powerful composition-sensitive parameter, with minimal effects by intrinsic shower fluctuations.

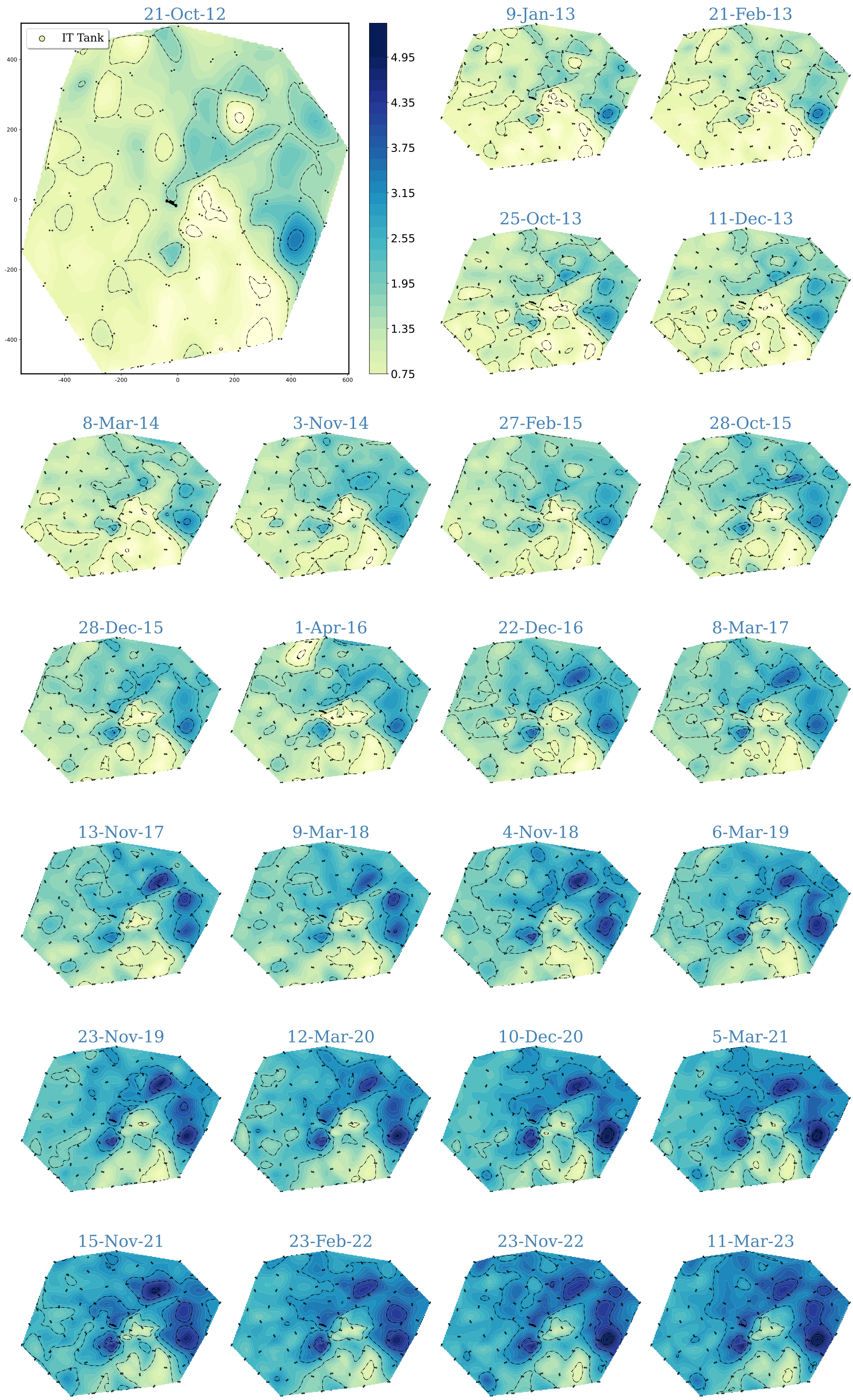


Figure 3.8: Change in snow-accumulation on IT-tanks over multiple years. IT is considered as a plane surface here.

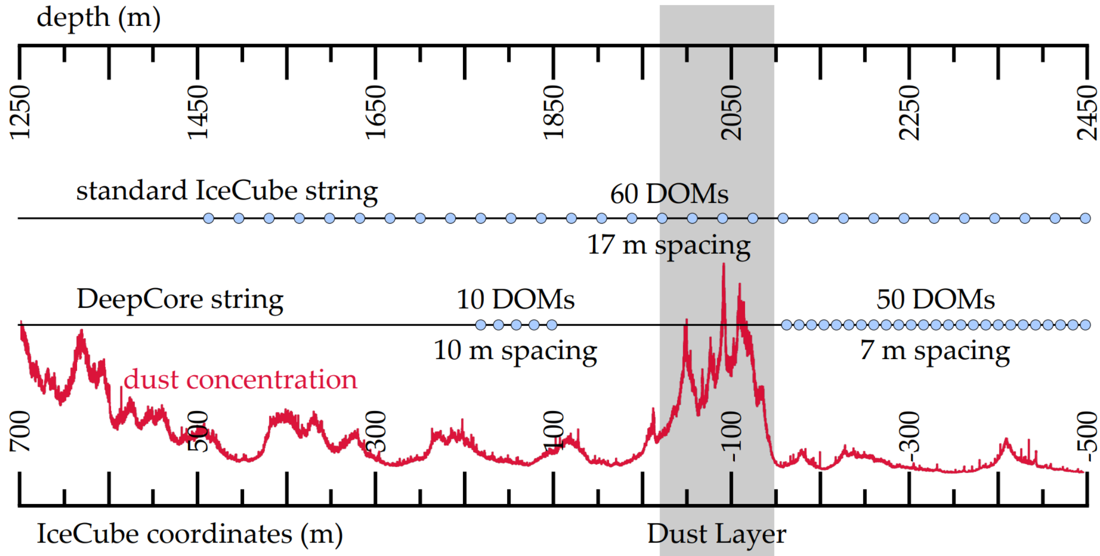


Figure 3.9: Dust concentration variation as a function IceCube-coordinates(z-axis) and depth. Comparison between standard IC and DC strings is also shown. Plot borrowed from [196].

strings is 125 m. This results in an energy threshold of 100 GeV for most IC analysis. It is well below the envisioned astrophysical neutrino observation range of $\mathcal{O}$(TeV)-$\mathcal{O}$(PeV). In order to lower the energy threshold of the observatory for ν-oscillation studies, the denser DeepCore (DC)-array is also instrumented. It is depicted by the green cylindrical outline in Figure 3.1. The inter-string distance is reduced to 72 m. The DC is segmented into two parts, separated by a dust layer. The top-part(inter-DOM distance = 10 m) of DC is generally used as veto for the rest(inter-DOM distance = 7 m) of the DC array. The dust concentration[10] as a function of depth (and ICNO's coordinate) is given in Figure 3.9. It also depicts the difference in inter-DOM spacing between IC and DC strings.

IC detects neutrino events that primarily belong to one of the three event topologies, namely: tracks, cascades, and double bang. The event topologies (and the underlying processes involved) for track, cascade, and double bang are shown in Figure 3.10[11]. A few details about the event topologies are as follows:

- **Tracks**: They are a consequence of the muons produced from the Charged-Current (CC) interaction of ν_μ with matter/ice. It is generally easy to reconstruct the direction of such events because of the long lever-arm. Above an energy of few TeV, the muon track direction is almost the same as its parent ν's direction [197]. However, the energy reconstruction for such through-going events is generally difficult because of incomplete knowledge of interaction vertex of ν_μ[12].

10 The glacial ice at the bottom of the South Pole is estimated to be about 165000 years old [192]. The ice layers hence manifest a natural record of South-Pole winds [193], and other changes in atmospheric conditions over the millennia. The atmospheric metamorphosis is imprinted in the South-Pole ice as a variation in dust concentration among ice layers, variation in ice-crystal structure, and isotope presence. The dust concentration is among the most crucial to reduce systematic uncertainties [194]. A recent work [195] helped improve our understanding of ice crystals at the South Pole.

11 The track and the cascade deposits shown are from real events. The double-bang is from a simulation.

12 The mean energy loss, i.e. $\left\langle \frac{dE}{dX} \right\rangle$, can be used to estimate the energy, since $\left\langle \frac{dE}{dX} \right\rangle \propto E$.

- **Cascades**: Cascade like events are a consequence of CC interaction of ν_e and ν_τ. In addition to that, they can also be caused by Neutral-Current (NC) interaction of any ν-flavor type [159]. The absence of a lever-arm and a near-spherical point-source-like emission makes the direction reconstruction difficult[13]. The energy reconstruction is easier than track-like events since these events are generally contained in IC.

- **Double Bang**: Double bang events are typically indicative of ν_τ. They are generated when on the neutrino via CC-interaction produces a hadronic shower, accompanied by a production of a τ-lepton. This τ-lepton generally propagates further (leaving a track-like deposit). It finally decays, producing a particle cascade [199]. Hence, such events look like a combination of the two topologies mentioned earlier.

The event topologies (shown in Figure 3.10) can hence be used to identify the type of ν[14] responsible to produce a signal at IC. Section 4.1 will give a brief overview of the variety of analyses done at ICNO, using IC and DC. Muons (from EAS) also show very similar charge deposits as from the muon-neutrino-induced tracks.

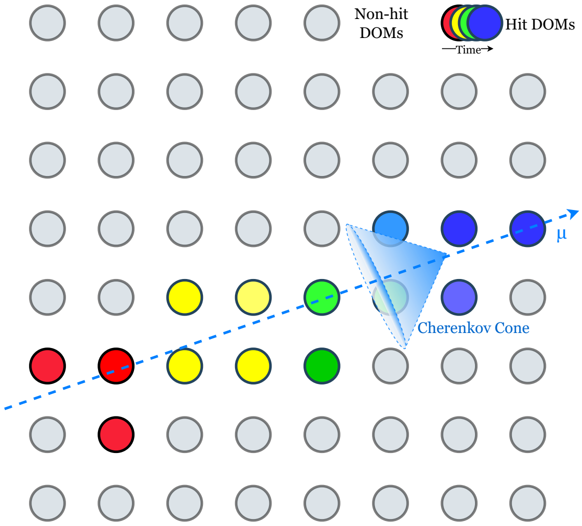
3.1.7 *Data Acquisition*

As mentioned in Section 3.1.3, there are three wire pairs that come out of a DOM. One of them is responsible to carry power to the DOM and allow the bidirectional communication pathway (finally terminating at ICL). The other two connect each DOM to the DOMs immediately above and below it. These are responsible to test if the LC condition is met or not. The LC condition ensures that the data is only digitized (at DOMs) and sent to ICL, if a coincident signal in neighboring DOMs was measured. This enables us to reduce the noise rate from electronics, the luminescence of glass sphere used in the DOMs or more of similar phenomena. IC and IT follow a slightly different variation of the LC condition.

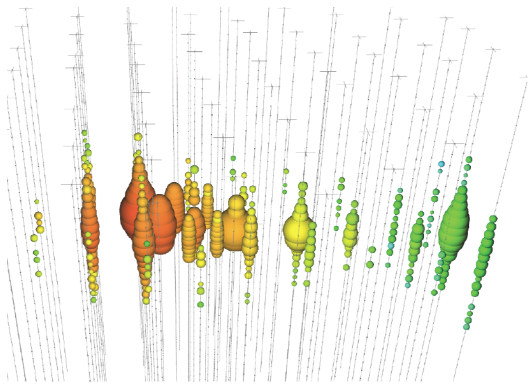
For IC, if a DOM and its nearest or next-to-nearest DOM is triggered within a time of $\pm 1\mu s$, the LC is considered satisfied and then the digitized waveform is sent to the surface and finally at ICL. Otherwise, the electronics in DOM is reset to reduce dead time. The total charge and time stamp are however still kept. The digitized hits are then called Hard Local Coincidence (HLC)-hits, whereas the others are referred to as Soft Local Coincidence (SLC)-hits. The HLCs are used to construct events from the individual hits[15] and perform spatiotemporal clustering and correlation to look for different possible event topologies (as discussed in Section 3.1.6).

The LC condition at IT is implemented station-wise, rather than DOM-wise. This is done in order the data-rate because of the lower threshold for triggering and a greater chance of random coincidences. An IT-station consists of two tanks, with two DOMs each. One DOM in each tank is a high-gain DOM, whereas the other one is a low-gain DOM. The wiring done to check LC among the DOMs in a station is depicted in Figure 3.11. For IT a hit is considered a HLC-hit if both tanks in the IT-station trigger within a time-window of $\pm$ 1 μs. However, it is considered as an SLC if only one

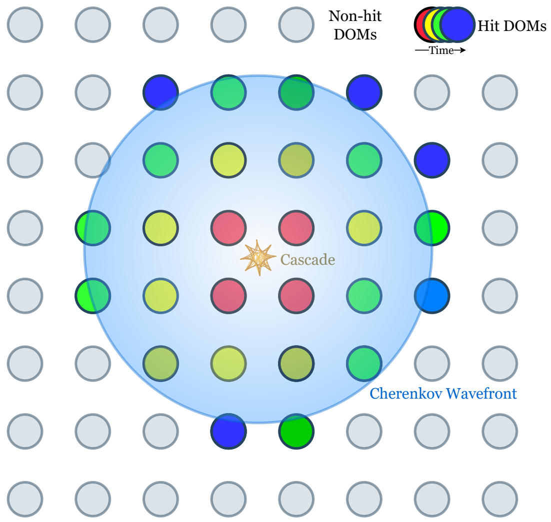
13 CNN-based methods have shown promise in the reconstruction of such events [198].
14 or a μ from an air-shower
15 within a time window of $10\mu s$



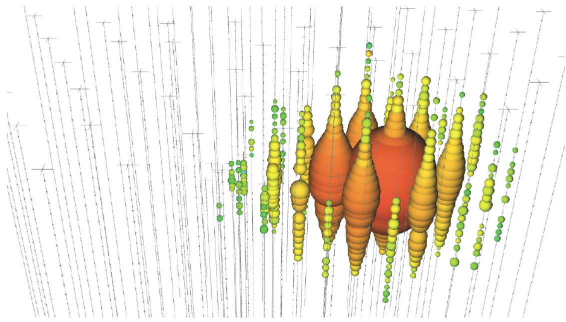
(a) Deposit by μ



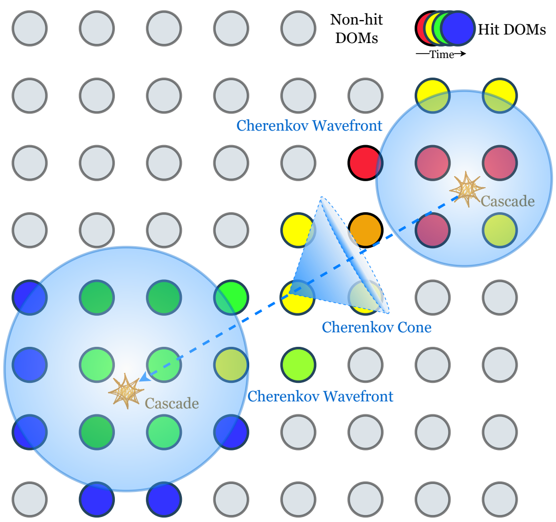
(b) Track Signature. PC : ICoII



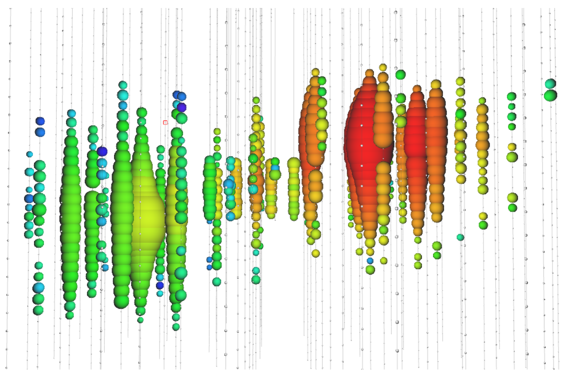
(c) Deposit by cascade event



(d) Cascade Signature. PC : ICoII



(e) Deposit by τ event



(f) Double Bang Signature. PC : ICoII

Figure 3.10: Neutrino event topologies at IceCube.

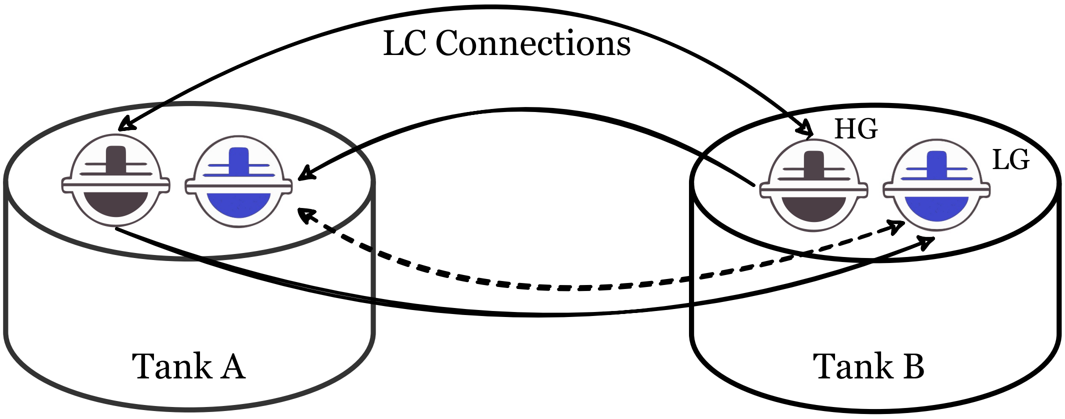


Figure 3.11: Cabling done to check LC at an IT station. Each IT station consists of two tanks with two DOMs each (high-gain = HG and low-gain = LG). The connection between the LG-DOMs is only used if the HG-DOM dies.

tank observes a hit. In the HLC-mode, waveforms of all the stations with successful LC is sent to the ICL, along with a mean timestamp and integrated charge information. In the SLC-mode only the timestamp and integrated charge information from all the triggered DOMs are transferred to the ICL.

After the successful fulfillment of LC-condition and data intake at DOM-level or station-level, a global trigger ensures that individual triggers from the sub-detectors are merged together. This allows for the identification of an individual event within the time window. The constructed events are then sent for further processing and filtering. Few important triggers at ICNO are:

- **Simple Multiplicity Trigger (SMT):** IceTop SMT requires 6 HLC-hits (usually from 3 stations) at IT within a time window of 6 µs. The readout window to test for the trigger starts 10 µs before the trigger and runs for the same time after the last 6 HLC-hits. On a successful trigger, the whole detector is read out and recorded. The trigger has an energy threshold of 300 TeV. A modification of this trigger requiring hits in at least 5 stations (**IceTopSTA5**) or hits in at least 3 infill tanks(**IceTop_InFill_STA3**[16]) is also implemented. For IC, if 8 HLC-hits occur in a time window 5µs, the IceCube SMT is satisfied. For IC, the readout window extends to 4 µs before and 6 µs after the hits.

- **InIceSMT_IceTopCoinc:** This trigger is employed to identify events characterized by a predominantly single HLC-hit at IT, while simultaneously meeting the IC LC-condition. This trigger plays an important role in vetoing fake neutrino signals.

- **SLOP:** This trigger is used to detect the presence of hypothesized sub-relativistic particles like magnetic monopoles [200].

Another trigger called Fixed Rate Trigger (FRT) is used to read the whole detector within the 10 µs window. Over the years, based on the requirements of different analyses multiple different string/area/volume level triggers have been implemented or updated. For this work only the IT-SMT and IC-SMT triggers are relevant. The IT-SMT

16 The threshold for this trigger is about 100 TeV

trigger condition is modified to only record if more than 5 stations-hits are recorded (instead of 3). The IC-SMT trigger condition remains the same. Another trigger ensures that the events recorded are coincident events between IT and IC. [201] provides a detailed summary of the steps involved in cleaning, filtering, and preparing the data before its ready to be analyzed.

3.1.8 *Calibration*

The charges at IT are expressed in terms of Vertical Equivalent Muons (VEM) units, in order to obtain a consistent measure of charge-deposited among the various DOMs in IT-tanks[17]. VEM is the charge deposited by a vertical muon in an IT-DOM[18]. Active calibration is also needed because the snow height over tanks keeps changing. The difference in snow height is responsible for the change in attenuation of the EM component of EASs. Atmospheric muon flux is used to actively get calibration measures between PE and VEM since these muons are mostly "minimum ionizing muons" and deposit constant energy in the detector. The waveform in High-Gain (HG) DOMs is used[19]. The distribution is then fit using a combination of normalized Landau and Fermi function for characterizing the muonic contribution and an exponential for the EM component [203]. This is given by:

$$f(x) = \underbrace{p_0 \left[L(x; p_1, p_2) + \frac{1.85}{p_1} \cdot \frac{1}{\exp\left(\frac{x - p_1}{p_2}\right) + 1} \right]}_{f_\mu} + \underbrace{p_3 \cdot \exp(p_4 \cdot x)}_{f_{EM}} \tag{3.1}$$

where in the muonic part the first part describes the through-going muons (using the Landau distribution) and the second part describes the non-edge clipping muons (given by the Fermi-like function). The second term is for the EM component. p_0 is the total number of non-edge-clipping muons. p_1 and p_2 are the location and the width of the Landau peak, respectively. p_3 and p_4 are the parameters of the EM-fit. Figure 3.12 shows the plot from tank 61A at IT during a calibration run and the fit obtain using Equation 3.1. 1 VEM is defined as 95% of the the maximum obtained from the fit. The calibration of Low-Gain (LG) DOMs is done using cross-calibration of the DOMs in a tank. It is assumed that the DOMs in a tank should record a proportional light yield per particle and hence the signal ratio of the DOMs in a tank should be approximately constant. The fit is then used to obtain the charge in VEM units for the LG-DOMs. A detailed summary of the calibration procedure at IT can be found in [202].

For IC, the calibration is done by using the LEDs on the flasher board (details in Section 3.1.3). In contrast to IT, this cannot be done actively during data-taking. To limit detector downtime, the calibration is done on a yearly basis. Moreover, the calibration is done in sections in order to still detect a transient event. After correcting for temperature, electronic and other effects, the IC calibration finally is used to get the charge measure in units of PEs. The timing calibration for DOMs at IT and IC is done

17 Since every DOM can have a different charge response to incident particles, using Photoelectron (PE) as the charge units are impractical.

18 For standard IT-tanks 1 VEM ≈ 125 PE for high-gain DOMs and ≈ 110 PE for low-gain DOMs [202].

19 To enhance contribution from single muons, only hits without LC are considered. Every 8192$^{\text{nd}}$ hit is taken (if LC is satisfied, the next hit is used), leading to a rate of ≈ 0.2 Hz.

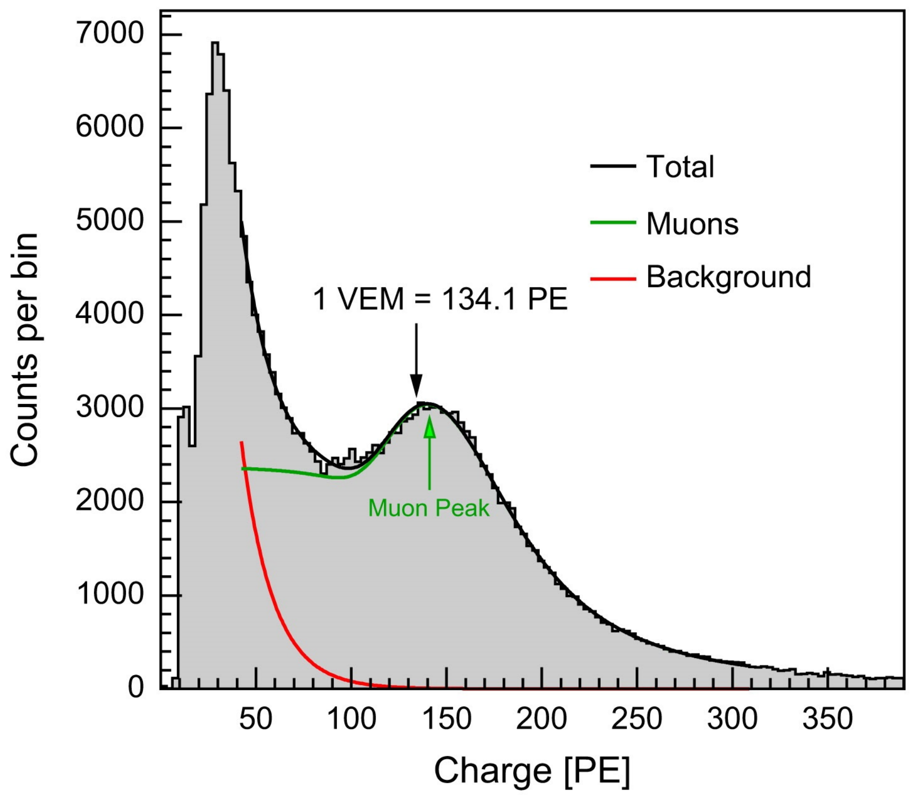


Figure 3.12: Charge spectrum at HG-DOM in tank 61A at IT, and fit to obtain EM and muonic component. 1 VEM = 95% of muonic peak-value. Plot borrowed from [202].

by calibrating and synchronizing them to clocks at ICL. This has an $\mathcal{O}$(ns) accuracy. A detailed summary of the charge-calibration procedure for IC and timing calibration for IT and IC can be found in [163].

3.1.9 *Environmental Characteristics @ South Pole*

Understanding the environmental conditions at the South Pole is crucial to predict and finetune measurement expectations at IT & IC. There are multiple factors that need to be considered namely,

- **Seasonal Variations**: During South-Pole winters (April to September) the surface temperature at Pole can range between -20 deg C to - 90 deg C [186]. The change in temperature leads to a difference in the density profile of the atmosphere[20], hence affecting the lepton production rate. The IC muon rate closely follows the density changes with seasonal variations, with an approximate $\pm$ (8-9)% variation. For IT, the variation is about $\pm$ 5% [204]. Similarly, a study for ν_μ observed more than 10 σ correlation between atmospheric neutrinos and stratosphere temperatures [186].
- **Snow Accumulation**: Due to drifting, snow accumulation over the IT tanks changes regularly. The snow depth can only be measured during Antarctic summer. For

20 This is responsible for the change in ground pressure and atmospheric overburden.

the rest of the months, a physics and data-driven approach is used. Using the fact that snow accumulation mainly affects the absorption of EM component of the air-shower (leaving the muon spectrum nearly unaltered) a calibration between snow heights and the ratio of muonic and EM component is used (calibration equation: Equation 15 in [202]).

3.2 ICECUBE-GEN2: THE WINDOW TO THE EXTREME UNIVERSE[21]

In the upcoming years, the ICoII plans to extend the current detector. This is done in order to directly or indirectly pursue the science objectives laid out in Astro2020 Decadal Survey [206] in the field of neutrino astronomy [207], fundamental physics with cosmic neutrinos [208], CR-Physics [209–216] (and their sources), and multi-messenger astronomy (in collaboration with other observatories; γ-rays [217–220] and gravitational waves [221–223]). To meet the scientific objectives the next detector will need to have a neutrino point-source sensitivity at least five times better than the current detector, with a near real-time reconstruction (with sub-degree resolution) for multi-messenger followups [205]. At least an order-of-magnitude increase in collection rate for all neutrino flavors in the 100 TeV to 10 PeV range (with improved flavor identification) is also needed. This should be accompanied by the ability to extend the energy sensitivity range beyond 10^{18} eV. IceCube-Gen2, the planned extension of the current IceCube Observatory, is designed to meet these requirements[22]. It is planned to be fully operational by 2033.

To meet the scientific objectives, the detector will add four new components: an extended in-ice optical array, a low-energy core (IceCube-Upgrade), a surface air-shower array [224][23], and an extended radio detector array. These are shown in Figure 3.14 (except the extended radio detector array). The figure presents a *Top-view* of the envisioned optical array of the IceCube-Gen2 Neutrino Observatory, with the current IceCube Array also shown. The dots represent the string locations. The string separation for IceCube-Gen2 ($\approx$ 240 m) is more than the current IceCube array ($\approx$ 125 m). IceCube-Gen2 will have a different design of the optical sensors (current design: Figure 3.3). The new optical sensors, mDOM [227, 228] and D-Egg (shown in Figure 3.13), will be tested in the IceCube-Upgrade. Instead of a single 10″-diameter PMT, the mDOM will consist of 24 smaller 3″ PMTs. mDOM, with PMTs pointing in almost all-directions (in contrast to bottom-facing in DOM), will contribute to an increase in photocathode area by a factor of 2.2. D-Egg [229] is also a multi-PMT sensor. It will use two 8″ PMTs, facing up and down. Although providing slightly worse directional information than mDOM, D-Egg needs lower power consumption and a smaller borehole (hence helping reduce operational and deployment costs). IceCube-Gen2 will add 120 new strings to the current optical array. Each string is planned to have 80 optical sensors (mixture of mDOM and D-Egg), located between the depth of 1325 m and 2575 m.

The surface array is a planned extension of the IT-array. Initially, scintillators and radio-antennas will be placed on top of IT-array. This is termed IceCube Surface Enhancement. The schematic of the planned surface array is depicted in Figure 3.15. As

21 title borrowed from [205].
22 Read [205] for details of how each of the objectives will be met.
23 Prototype stations for surface-array are already being tested at the South-Pole [225, 226]. The stations will finally be integrated into the surface array.

Figure 3.13: Prototype of mDOM and D-Egg. These will be tested in IceCube-Upgrade[230]. PC: ICoIl.

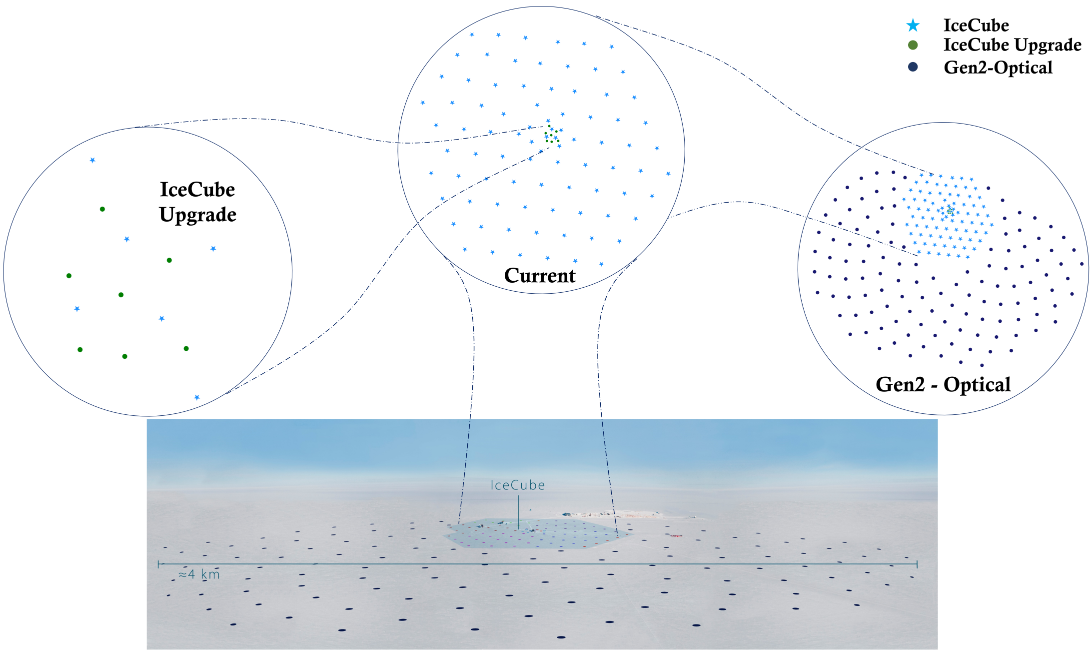


Figure 3.14: Planned design (*Top View*) of the extended IceCube-Gen2 optical component (*right: Dark-Blue Dots*) with the current IceCube Array (*center: Light-Blue Stars*) with the planned seven low-energy IceCube Upgrade strings (*left: Green Dots*). The extended radio array (not shown here) extends farther beyond the planned IceCube-Gen2 Observatory.

mentioned in Section 3.1.9, over the years we see an overall increase in the snow-accumulation over IT-tanks. This leads to an increased detection threshold [169], because of greater attenuation of the EM-component before deposit in the IT-tanks. Surface-Array is planned to avoid snow accumulation over future surface detectors and hence help reduce systematic uncertainties in the calibration (and the signal-measured). It will consist of elevated scintillator-panels and radio-antennas. The detector components can be elevated over years to prevent snow accumulation. Deployment of scintillator panels is also more cost-effective than ice-Cherenkov tanks (like the ones used for IT). By giving an independent measure of a shower footprint, scintillators can help improve the calibration of the current IT-array [231, 232]. Moreover, the complementary measure from IT and scintillator-panels can also be used to study CR-composition [231]. The scintillators are accompanied by radio antennas. The radio-antennas will provide enhanced sensitivity to inclined-showers [224], which is generally difficult with scintillators. In addition to this, radio-antennas are capable of providing independent X_{max} estimate[24]. X_{max} is an excellent composition-sensitive parameter, and is commonly used to perform composition estimates in other CR-observatories too [80, 233–238]. For IceCube, it can provide important insights into the CRs in the Galactic-to-extragalactic transition range [239, 240]. The complementary information provided by IT-tanks, scintillator panels and radio-antennas will also enable testing hadronic interaction models [224, 239], PeV-photon searches [240, 241], looking for mass-dependent anisotropy [241], probing particle physics and more. Prototype surface-array stations are already being tested at the South-Pole [225, 226]. This is depicted in Figure 3.16. Hybrid measurements have already shown promise in the potential benefits of the surface array [242]. In addition to serving as a CR-detector, the surface array will also act as a powerful veto against atmospheric neutrinos, and hence help identify astrophysical neutrinos.

An extended radio array is also planned. It will be used to detect neutrinos up to EeV energies by detecting Askaryan emission of neutrino-induced cascades in-ice. The array will cover an area of approximately 500 km^2, providing possibilities to study astrophysical and cosmogenic neutrinos. The complete array will consist of 313 stations. Out of the 313 stations, 169 will only consist of detector components located at shallow depths. Whereas the rest will be hybrid, consisting of both shallow and deep detectors. It will allow studying transient astrophysical neutrino events with unprecedented sensitivity and help extend the multi-messenger capabilities of the detector [244]. The layout of a hybrid radio-array station is depicted in Figure 3.17.

24 It is the atmospheric depth at which the number of particles in an air-shower is maximum.

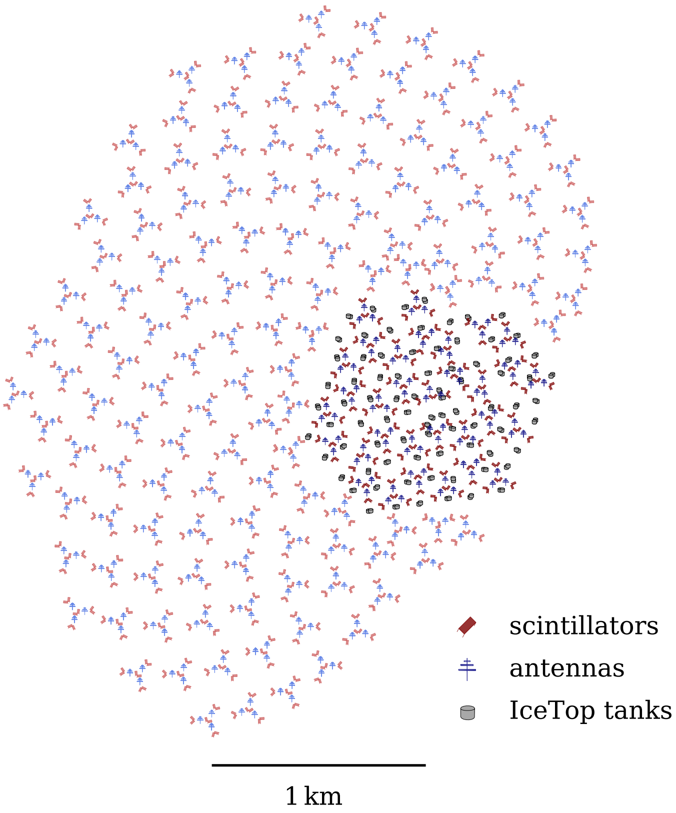


Figure 3.15: Planned layout of the Surface-Array component of IceCube-Gen2 Observatory, with the current IceTop Array. The surface array will constitute multiple stations, each constituting 8 scintillation detectors and 3 radio antennas. Plot from [243].

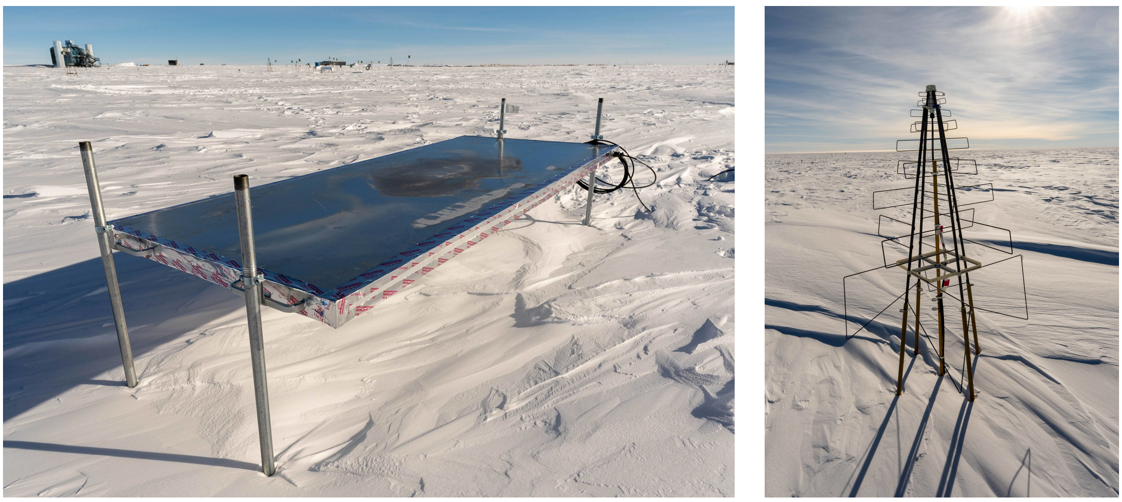


Figure 3.16: Elevated scintillator-panel (left) and radio-antenna (right) at the South-Pole. PC: ICoII.

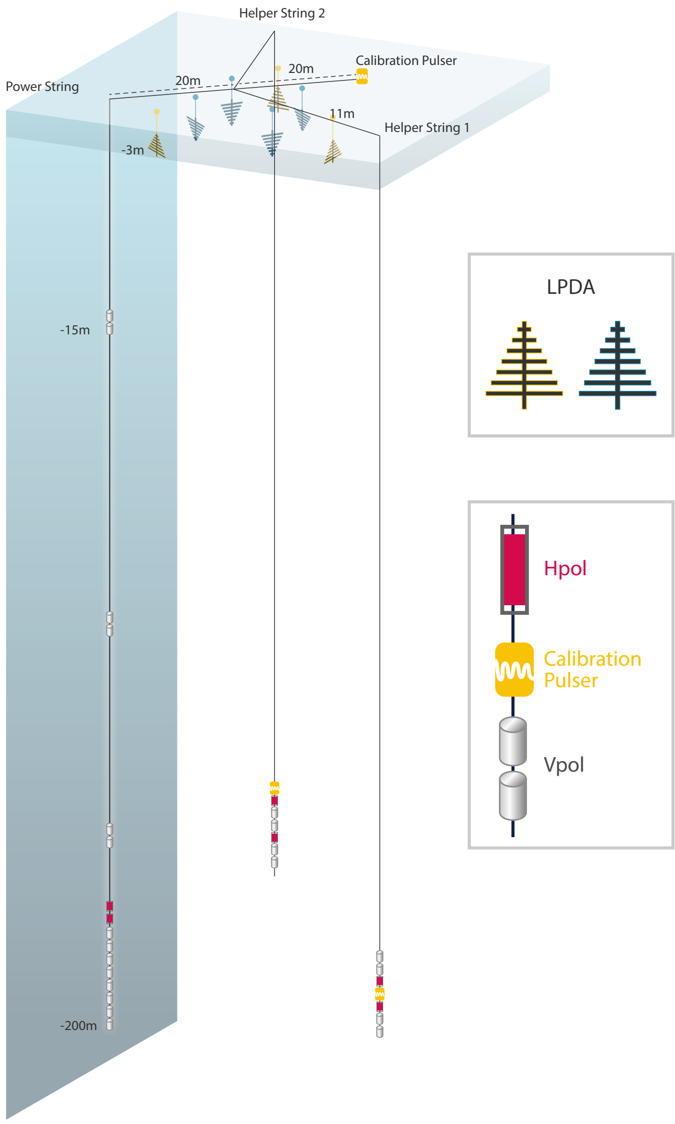


Figure 3.17: Hybrid radio station for the planned extended radio array. Picture from [244].

4 SCIENCE @ ICECUBE

ICNO is involved in research areas in multiple frontiers. In the field of neutrino-physics, the topics range from looking for neutrino point sources or diffuse flux to as wide as looking for new physics. This is further accompanied by collaboration with multiple cross-collaboration partnerships to perform multimessenger astronomy. On the cosmic-ray (CR) frontier, the collaboration is the leading observatory by being a unique 3D-detector in the transition region between galactic and extra-galactic CRs. This has enabled testing for CR-anisotropy, composition studies, and more. In addition to that the observatory has also contributed to interdisciplinary efforts in the field of glaciology, and heliophysics. The following text will give a brief overview of this variety of research works carried out at ICNO.

4.1 NEUTRINO ASTRONOMY

ICNO, as envisioned, has proved to be a very crucial probe to study astrophysics. To date ICNO has achieved many first-of-its-kind and the world's-best scientific results, in addition to the many scientific milestones. The most prominent among them in the field of neutrino astronomy are:

- ICNO was the first observatory to provide hints [245] (two events named *Bert* and *Ernie* - Figure 4.1) and finally confirm [246] the detection of high energy astrophysical neutrinos[1]. This helped open new eyes to see the universe, which was otherwise opaque to us in the EM-spectrum. This was soon followed up by the detection of the highest energy ν ever observed (*Big Bird* - Figure 4.1), with E $\approx$ 2 PeV [248]. The energy limit was recently surpassed (Hydrangea - Figure 4.2) by the observation of Glashow resonance (interaction of HE electron antineutrino with an electron) [249].

1 The conclusion was confirmed once more by observation of Astrophysical ν-flux from northern sky [247].

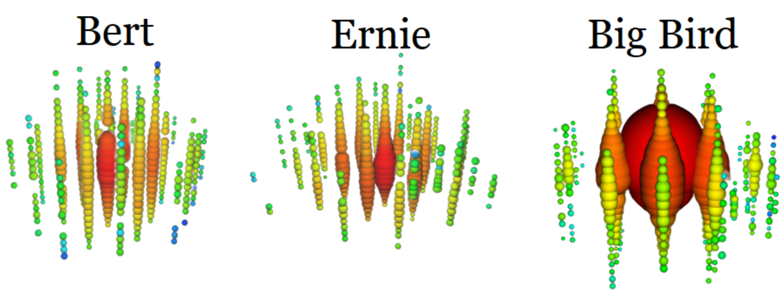


Figure 4.1: Bert, Ernie and Big Bird. See text for details. PC : ICoIl.

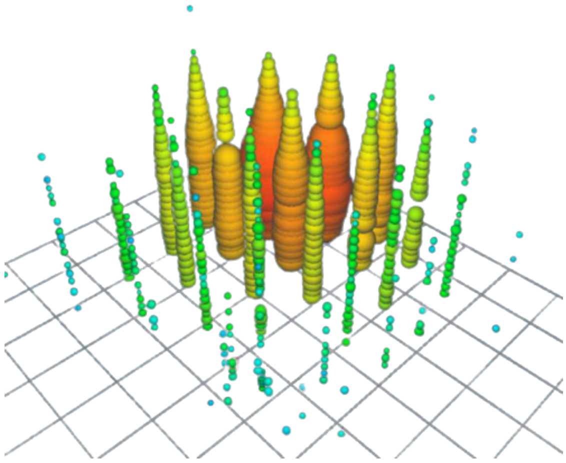


Figure 4.2: Glashow resonance event observed at IC, named *Hydrangea*. PC : ICoII.

- ICNO was the first observatory to provide evidence and point out the source of an astrophysical neutrino [250]. The astrophysical accelerator was a blazar titled TXS 0506+056. This alert sent by ICNO was soon followed up by other observatories around the world. The detection was corroborated by multiple observatories around the world[2] (across the EM-spectrum) that the blazar was indeed the likely source [251]. The containment region is depicted in Figure 4.3. Evidence for HE-neutrino emission from NGC 1068 (Messier 77) was also reported recently [252].

- Recently, ICNO was also able to establish very strict upper limits on the contribution of core-collapse supernovae to the HE neutrino emission [253], from transients [254], from Fast Radio Burst (FRB)s [255], galaxy clusters [256], time-dependent neutrino sources [257]. ICNO has identified neutrino emission from the Galactic plane with a 4.5 σ level of significance. The results from this study are shown in Figure 4.4.

- ICNO was able to measure all-flavor neutrino cross-section measurements [260]. The results are consistent with Standard Model (SM) of Particle-Physics. Another analysis was performed to test for extension of the "3+1"[3] neutrino flavor model, by testing for if the fourth state decays into lighter invisible particles. The effect was tested by looking for influences on the flux of muon neutrinos. It was found that the unstable sterile neutrino model is more compatible (than the standard model or 3+1 model) with the data observed [261]. The results are not statistically significant yet.

2 The space missions AGILE, INTEGRAL, and Fermi, as well as ground-based telescopes such as HAWC in Mexico, H.E.S.S. in Namibia, MAGIC in Spain, and VERITAS in the U.S., have detected gamma-rays. Additionally, space missions like MAXI, NuSTAR, and Swift, along with ground-based observatories such as ASAS-SN in Chile and the U.S., GTC, and Kapteyn in Spain and the U.S., Kanata and Kiso in Japan, Liverpool in Spain, OVRO in the U.S., SALT in South Africa, Subaru in Japan, and VLA in the U.S., have observed X-rays, optical, and radio radiation. Finally, neutrinos were detected by ANTARES in France.

3 ν_e, ν_μ, ν_τ accompanied by massive sterile neutrino

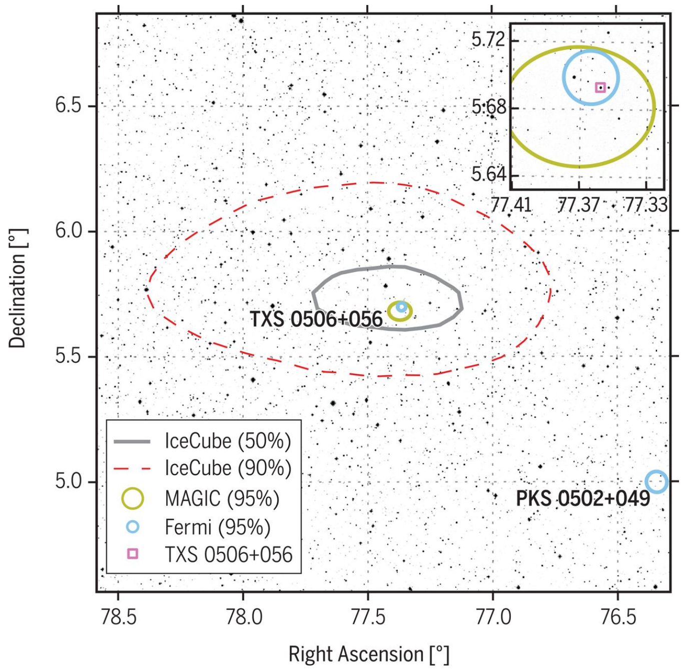


Figure 4.3: Containment region for event observed at ICNO, corresponding to observation of blazar TXS 0506+056. The confidence region from the follow-up campaign by MAGIC Telescope and Fermi spacecraft are also shown. Plot from [251].

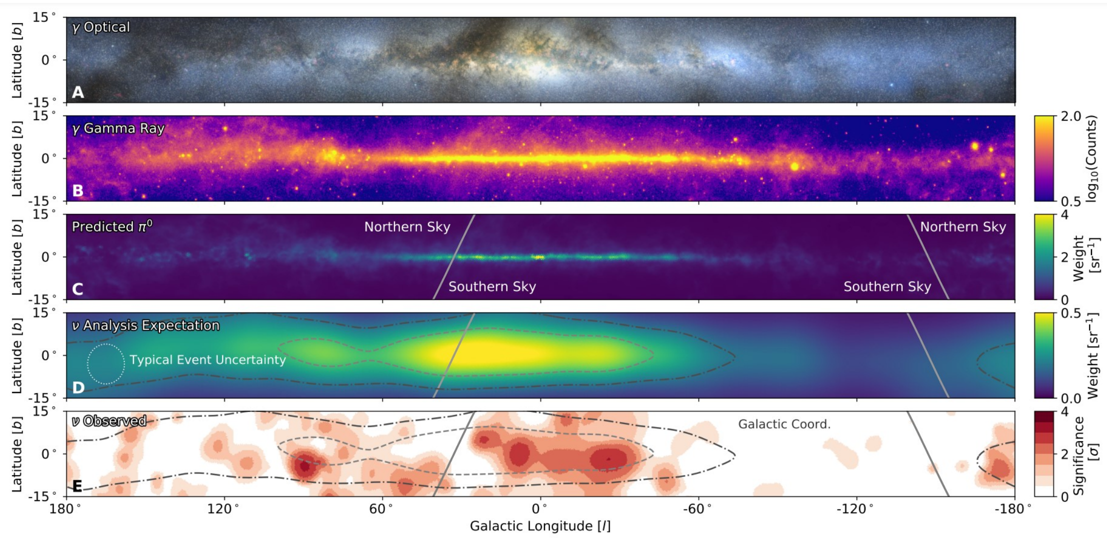


Figure 4.4: Milky Way galactic plane in photons and neutrinos. The panels are in galactic coordinates. The first panel presents an optical view of the galactic plane [258]. The second panel presents the integrated gamma-ray flux from *Fermi*-LAT (for E > 1 GeV) 12-year survey [259]. The third panel depicts the emission template for the expected neutrino flux (derived from π^0 templates from *Fermi*-LAT observations of diffuse γ emission). The fourth panel shows the emission template accounting for the detector sensitivity and 20% and 50% containment contours for diffuse neutrino emission. The last panel shows the pre-trial significance of the IC neutrino observations. The grey line depicts the Northern-Southern sky horizon line at ICNO.

- Ultra-high-energy Cosmic Ray (UHECR)-sources are considered as a possible sources for astrophysical neutrinos too. ICNO in a cross-collaboration effort (with ANTARES, Auger, and Telescope Array) looked for correlations between neutrinos and UHECR sources [262]. Although no significant correlation was detected, the analysis was able to constrain the neutrino flux in the direction of UHECR sources[4].

Leveraging DeepCore

DeepCore has more denser string (and DOM) density and also has higher quantum efficiency DOMs, than the standard IC-array. This allows to reduce the detection threshold to about 10 GeV (from about 100 GeV for the standard IC-array) [263]. The lower threshold allows the DC to perform systematic studies in the following research areas:

- **Neutrino Oscillation:** An energy threshold of approximately $E_\nu > 10$ GeV, allow DC to study the phenomenon of neutrino-oscillation using atmospheric neutrinos [264, 265]. The best-fit oscillation parameters obtained by analyzing muon neutrino disappearance are competitive and in agreement with other fully earth-bound[5] long-baseline experiments, such as T2K, MINOS or NOνA-experiment [266]. The observation of atmospheric ν_τ [267], provides probes for Beyond Standard Model (BSM)-physics, which is generally not possible at other experiments. It might also help shed new light on the possible solution to the neutrino mass ordering problem [268].

- **WIMP Searches:** Weakly interacting massive particle (WIMP) are among the most promising Dark Matter (DM) candidates [269]. DC can detect neutrinos from the annihilation of WIMPs (in Earth's center, in the Sun, Galactic Center, and the galactic Halo). DC has already set competitive WIMP-proton scattering cross-section estimates and more [270–273], and see Section 4.4.2 for more details.

- **Other:** Looking for slow-moving monopoles, supersymmetric stau pair production [274, 275] and low-energy neutrinos from astrophysical sources [276] are other research areas where DC can play an important role. A recent analysis was able to establish first-ever observational upper limits on neutrino emission from novae [277].

4.2 COSMIC-RAY PHYSICS

In addition to being an efficacious neutrino observatory, ICNO holds the unique position of being a very powerful CR detector in the transition region between galactic to extra-galactic CRs. IT is regularly bombarded with CR initiated EAS. The EM-component and GeV-muons are primarily responsible for deposit at IT. The TeV-muons present in the EAS can however propagate deeper and deposit signal in IC. Both IT and IC have been used to study various aspects of CR-Physics. Most prominent CR-research directions probed at ICNO are:

4 An improvement on mass-estimate on an event-by-event basis for CRs is expected to improve the analysis substantially in the future.

5 i. e. neutrino source also on Earth (generally an accelerator or nuclear reactor)

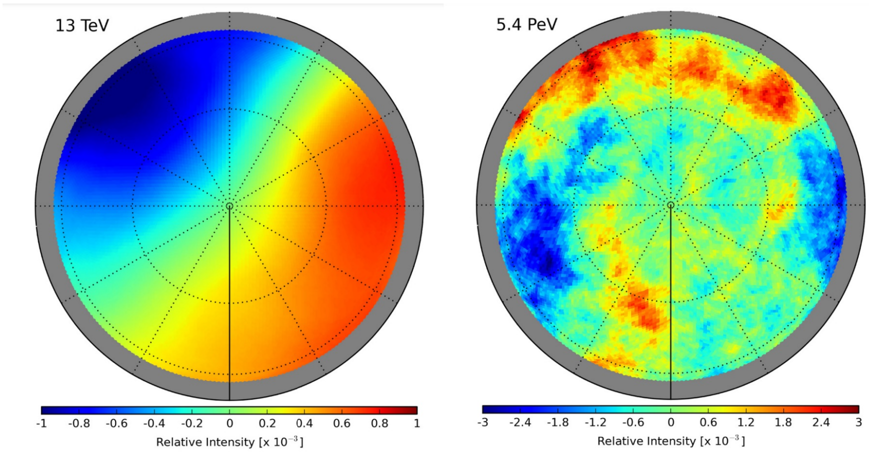


Figure 4.5: Relative CR-anisotropy map at IC, using 6 year of IC data. Plot from [278].

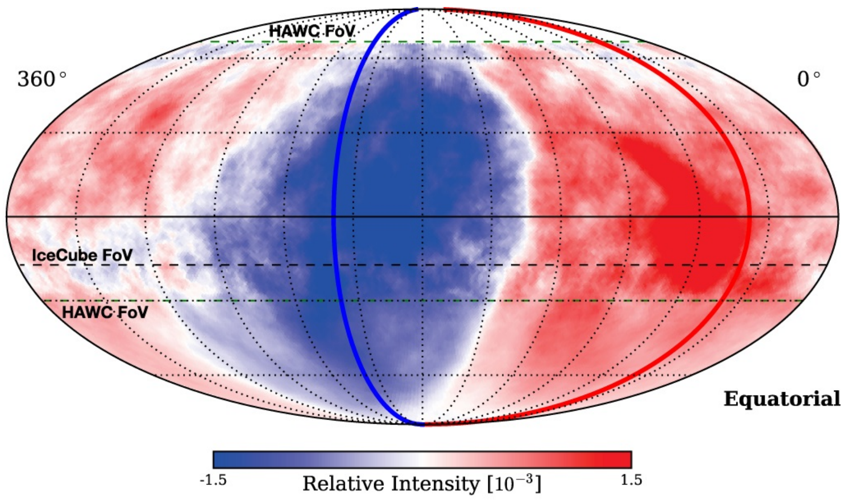


Figure 4.6: Relative intensity map of CR at a median energy of 10 TeV, using HAWC and IC data. Plot from [281].

- **CR Anisotropy:** CR anisotropy can provide useful hints to the location of CR-sources[6], and the magnetic fields they propagate through. ICNO was the first observatory to report on an anisotropy in the CR arrival-direction distribution for the Southern Hemisphere [278]. This is depicted in Figure 4.5. The result found a phase-shift in anisotropy with increasing energies, consistent with results from other observatories (Pierre Auger Observatory [279], KASCADE-Grande [280]). No significant anisotropy variation over time was observed. The work was recently updated using 9 years of ICNO-data, in the TeV-PeV energy range. Similar to the previous results, large and small-scale structures were observed. Similarly, a phase transition in the CR-dipole was observed. A recent full-sky combined analysis between HAWC and ICNO to study CR anisotropy at median energy of 10 TeV [281] was also performed.

6 Knowing the CR-composition can also provide hints about the acceleration mechanism at such sources.

- **CR Spectrum and Composition:** IT, in conjunction with IC, was used to probe CR-spectrum and composition in the transition region between galactic to extra-galactic CRs [20]. The total-particle spectra was found to be in good agreement with previous work. The composition expectation was also found to be generally consistent with the phenomenological models (well agreement with H3a and H4a [90]; and within systematic uncertainties with GST-fit [91] and GSF [92]). Ongoing work is trying to improve on the composition results by looking for more composition-sensitive parameters and using a GNN-based method for mass prediction [282–285]. This work is a detailed summary of the publications mentioned (Chapter 6 - Chapter 10). A recent work [166] augmented the HE CR-spectrum results by lowering the energy threshold of the detector range to about 250 TeV [7]. This work also allowed us to reduce the energy gap with space or balloon bound CR detectors. The flux was found to be within the systematic uncertainty of the HE-spectrum results. Ongoing work is trying to also fill in the composition gap at these energies [284, 286].

- **Muon Puzzle:** There is a known discrepancy in the muon-number expectation in simulation and the one observed at the CR-observatories. A recent work by ICNO shed light on the observed discrepancy between the observed muon number and the expected muon-number from simulations, for GeV muons. It was observed that IT-data agrees well with pre-LHC hadronic models for any realistic composition mixture. However, the results found out that the observed muon density was lower than the one expected by the post-LHC model (e.g. EPOS-LHC). This is in tension with observations from EAS-observatories at EeV-energies, which see higher muon density in the data than the simulations predict. The results however confirmed that the discrepancy between theory and data increases with energy. Ongoing works to test similar discrepancies for TeV muons found results consistent with expectations for all hadronic models [9]. A cross-collaboration effort is also working on trying to resolve the issue [287].

- **γ-Rays:** Locating γ-ray sources in our galaxy can provide strong hints towards the identification of galactic CR-sources. ICNO performed an analysis to search for PeV γ-ray sources in our galaxy [288]. Even though no significant excess of γ-rays was observed, the result was able to establish the most stringent upper limits on PeV γ-ray emission. Ongoing works are trying to improve the possibility of finding such sources in the future [289].

- **Sun Shadow:** CR are charged-particles. Studying the Sun's shadow of CRs can hence be used to test and improve solar magnetic field models[8]. Similarly, Moon also blocks CRs, with a well-known solid angle[9]. Data from ICNO was used to test Sun's and Moon's shadow [295]. Sun's shadow was found incompatible with the geometrical shadowing, with a 7.3σ confidence level. Moon shadow was found compatible with the models.

7 By implementing a dedicated trigger, using an in-fill array of IT.
8 Sun shadow is created by blocking of CRs by the sun's coronal magnetic field [290].
9 This is generally used at CR-observatories to study angular resolution, absolute pointing, and absolute energy scaling [291–294].

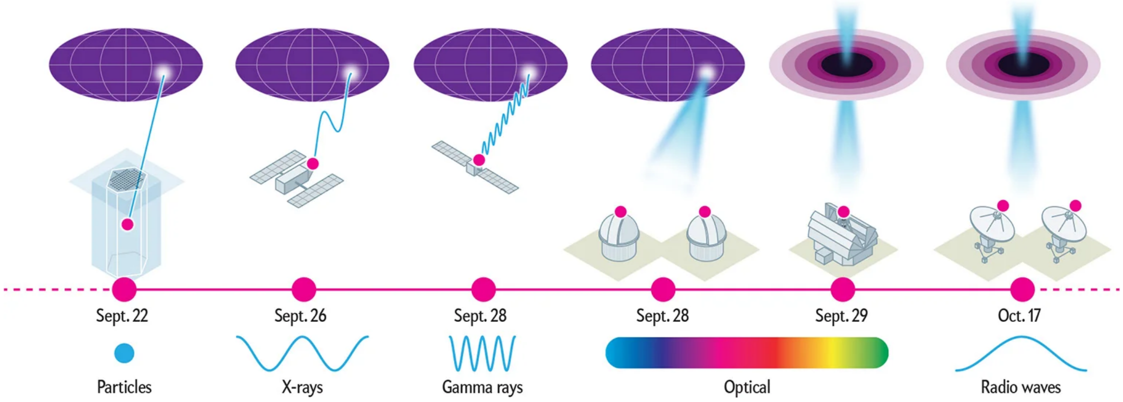


Figure 4.7: Illustration of neutrino detection from blazar TXS 0506+056 at ICNO and the subsequent followup at other observatories, using other cosmic messengers. PC : Nigel Hawtin, Scientific American [302].

- **Other:** Other works are trying to test hadronic models [174]; look for composition sensitive parameter at IT [296]; extend the zenith-range [297]; extend the capability of IT using radio-antennas [242], scintillators [225], Cherenkov telescopes [298]; two component LDF (EM and muonic); improve veto capabilities [190]; 3D event-reconstruction [299] and more.

4.3 MULTIMESSENGER ASTRONOMY

Multimessenger astronomy[10] is the field of studying cosmic sources using various astrophysical messengers together, like γ-rays, neutrinos, gravitational waves, cosmic rays, etc. The complementary information provided by the sources allows us to benefit from the unique information provided by each of the messengers. The identification of the first possible source of an astrophysical neutrino [251] from the blazar TXS 0506+056, was among the most prominent observations done using multiple messengers. An alert sent (on September 22, 2017) by ICNO, from neutrino detection, was followed up by very-HE γ-ray observations by MAGIC, H.E.S.S., and VERITAS; HE γ-ray observations at Fermi-LAT and AGILE; X-ray observations by Swift XRT; in optical spectrum at ASAS-SN, Kiso/KWFC, and Kanata/HONIR; by radio observations at OVRO and VLA. An outline of these is depicted in Figure 4.7. Another event from the same source was found in the archival dataset(from 2015) [250]. In 2020, another multimessenger observation [301] was reported. It was an observation of a neutrino that was spatially and temporally coincident with a Tidal Disruption Event (TDE) observed by Zwicky Transient Facility. Some other multimessenger efforts going at ICNO, corresponding to specific cosmic-messengers, are:

- **Gravitational Waves:** There are ongoing efforts to look for neutrino sources correlated with sources responsible for the detection of gravitational-wave events detected By Laser Interferometer Gravitational-Wave Observatory (LIGO)/Virgo [303, 304]. The studies were able to establish limits on the maximum viable neutri-

10 "*In cosmic messengers we find, a universe of wonders intertwined.*" - ChatGPT[300] prompted by me; Prompt "Write a 2 line poem about multimessenger astronomy.".

nos linked to each gravitational wave source and on the total energy discharged by neutrinos.

- **Gamma Ray:** Even though ICNO was unable to find any significant correlation between detected neutrinos and gamma-ray sources observed by Large High Altitude Air Shower Observatory (LHAASO) [305], it was able to set constraints [306]. ICNO was also able to set the limits on neutrino emission from Gamma Ray Burst (GRB) 221009A [307][11]. Another work was able to set upper limits on the contribution of MeV GRBs to astrophysical neutrino flux [308].

4.4 OTHER SCIENCE

The research work at ICNO is not limited to looking for neutrino sources and CRs. ICNO has also contributed to multiple areas of research which are trying to look for new physics or are in the field of interdisciplinary research. The following text will give a brief overview of some of them.

4.4.1 *Beyond Standard Model (BSM) Physics*

Neutrinos are considered as one of the most promising candidates to look for BSM physics. BSM physics-based models hypothesize the existence of additional interactions (than otherwise) in the interaction of neutrinos with matter. ICNO was able to put limits on all the parameters used to describe these non-standard interactions [309]. The limits put by ICNO individually, were comparable and compatible to the combined limits put by all other research works worldwide.

4.4.2 *Dark Matter*

ICoIl recently performed an analysis [272] to observe monochromatic-neutrinos (of same energy) from DM annihilations or decay into neutrinos[12]. Although no significant excess was found, new upper limits on DM annihilation and lower limits on DM lifetime were established.

4.4.3 *Quantum Gravity*

ICNO looked for imprints of quantum gravity in the observed fractions neutrino flavors of astrophysical neutrinos. The work used simulations generated for the astrophysical neutrino flavor model which include the quantum gravity effects. This was then compared with data observed at ICNO. No evidence of effects by quantum gravity on neutrino flavors was found [310]. However, the work was able to establish stringent limits on the parameterization of the space-time defects and can be used to set limits on BSM Physics [311], long-range force [312], neutrino–dark energy coupling [313], neutrino–dark matter scattering [314], violation of equivalent principle [315] etc.

11 It was observed on October 9th, 2022. It is one of the closest and brightest GRB ever observed. With an energy exceeding 10 TeV, it is the first such GRB detected by earth-bound gamma-ray observatories.

12 The monochromatic energy can give hint to the mass of the DM particle.

4.4.4 *Glaciology*

The knowledge of the detector is crucial for understanding the systematic effects in most analyses. ICoII has led the initiative by performing the most detailed measurement of Antarctic ice properties and their effects on light propagation. Several properties like ice-layer stability, tilt, and shear have been studied [316]. Using high-resolution particulate profiles, ICNO was also able to reconstruct a detailed climate record of the last glacial period. In addition to that, the measurement was also able to find evidence of Toba volcano super-eruption (in Sumatra, Indonesia), which happened about $74,000$ years ago [317]. A recent work [195] helped improve our understanding of ice crystals at the South Pole.

4.4.5 *More*

ICNO has looked for Lorentz symmetry [318], neutrinos in solar flares [319]. In addition to that another work tried to study the solar magnetic field by looking at its imprint on the shadow of cosmic rays [295]. ICNO has also conducted searches for the presence of magnetic monopoles, which could potentially exist in the relic flux stemming from the Big Bang. However, these searches have thus far yielded no conclusive results [320].

5 SIMULATION AND RECONSTRUCTION

Simulations are the foundation stones, as well as crucial towards understanding the complex dynamics of underlying processes in EAS physics. The simulations play a vital role in linking EAS observables with the properties of CR-primaries. Moreover, they also play a vital role in design decisions for EAS observatories. Reconstructions meanwhile allow extracting useful physics information (like energy, direction, composition, etc.) from EASs which can then be used to compare the experimental data with simulated EAS profiles as well as understand possibly observed discrepancies. This chapter will give the underlying details for producing EAS simulations, and their detector response at ICNO. Finally, reconstructions and quality cuts used for this analysis will be discussed. Quality cuts ensure that only high-quality events with good reconstructions are used further in the analysis.

5.1 EAS SIMULATIONS

In order to perform a realistic EAS simulation, understanding of multiple components is crucial. The following text will detail the important building blocks for producing relevant CR MC-simulations at ICNO.

5.1.1 *CORSIKA*

Detailed theoretical understanding and modeling are crucial for establishing quantitative expectations for drawing out physics from measurements performed in any experiment. For modelling EASs, COsmic Ray SImulations for KAscade (CORSIKA)[1] [322] is globally used for this purpose. CORSIKA is a Monte Carlo (MC) simulation software that enables detailed simulations of EASs initiated by HE CRs. In order to make estimates of air-shower observables at the observatories, it explicitly tracks billions of particles as they interact with air nuclei or decay during their propagation in the atmosphere. The detailed simulations account for ionization losses, multiple scattering, as well as the deflection of charged particles under Earth's magnetic field. It has been refined and extended multiple times to be used by multiple groups, ranging from the use case in the Cherenkov telescope to the highest energy cosmic-ray observatories. Multiple experiments have used it to simulate EASs initiated with varieties of primaries like p, α, O, N, Al, Fe, γ, etc as well as to understand interactions and decays of nuclei, hadrons, muons, electrons and more. The detailed simulations have allowed multiple groups to probe multiple EAS properties like particle number, density, lateral distributions and

1 It was developed as a simulation tool for KASCADE experiment [321] and its first version was made public on October 26, 1989.

e^+ *vs* e^-	l *vs* p (A)	p *vs* p	h *vs* A	A *vs* A
PYTHIA [327]	PYTHIA	PYTHIA	PYTHIA	PYTHIA
HERWIG [328]	HERWIG	HERWIG	QGSJET	QGSJET
Phojet [329]	RAPGAP [330]	SHERPA [331]	SIBYLL	SIBYLL
EPOS [332]	CASCADE [333]	Phojet	DPMJet	DPMJet
	Phojet	DIPSY [334]	HIJING [335]	AMBT
	EPOS	QGSJET [336]	DIPSY	HIJING
		SIBYLL [337]	EPOS	DIPSY
		DPMJet [338]		Hydro Models [339]
		EPOS		EPOS

Table 5.1: Event generators are available for a variety of projectiles and targets. Here e^- = electrons, e^+ = positrons, l = leptons, h = hadrons, p = protons and A = Nuclei. Table adapted from a talk given by Tanguy Pierog at Topical Lectures "Cosmic Rays" at NIKHEF, Amsterdam, The Netherlands (October 2021).

so on. The validation against observed data has established it as a preferred tool for the purpose. CORSIKA recognizes over 50 elementary particles for tracking during EAS propagation. In order to perform an EAS-simulation, the primary particle has to be given a type, a predefined energy, and an angle of incidence, along with the choice of hadronic-interaction models. Section 5.1.2 will discuss the details of the choice of the various hadronic interaction models. In addition to this, the atmosphere needs to be defined [2]. At ICNO, an atmosphere model [201] based on data collected from balloon flights conducted by the Antarctic Meteorological Research Center (AMRC) [323] and satellite observations made using the Atmospheric Infrared Sounder (AIRS) [324][3]. [322] gives a detailed summary of CORSIKA. It is being upgraded again to make it compatible with modern coding standards and flexible to choice of propagation media [325, 326].

5.1.2 *Hadronic Interaction Models*

Researchers have developed a variety of event generators in order to describe the physics of interactions between a variety of projectiles and their targets. The event generators help physicists explain and explore the possible interactions happening (and their rates) between the target and projectile [4]. Table 5.1 provides a non-exhaustive list of some standard event generators. The EM interaction in CORSIKA is examined using Electron Gamma Shower system (EGS) or with Nishimura Kamata Greisen (NKG) formula. These can be used to understand and estimate the spatial and kinematic distribu-

2 CORSIKA allows users to select multiple representative atmospheres. It consists of a five atmospheric layers model. Read Section A.1 for details of the layered structure.

3 Earlier ATMOD-12 (based on July 01, 1997, South Pole atmosphere (MSIS-90-E)) was used. However, it had inconsistencies with the simulations. Moreover, work was needed to improve the modeling of the layer near the surface of the Antarctic surface. Read [201] for details.

4 In addition to this it can also steer researchers to design new detectors and analysis strategies [340].

tion of the EM particles[5]. For this analysis EGS4 [341] is used as the EM event generator. For e^-/e^+, it accounts for annihilation, Bhabha scattering, bremsstrahlung, Møller scattering, and multiple scattering. For γ-rays, it accounts for Compton-scattering, e^-e^+ pair production and photoelectric reaction are accounted for. Furthermore, in spite of having very small cross-sections (CSs), $\mu^-\mu^+$ pair production and the photo-nuclear reaction of atmospheric nuclei with protons and neutrons are also incorporated. [322] provides a detailed summary of EGS4 and the modification done to it to account for barometric density dependence, Earth's magnetic field, pressure correction, computational approximations, and more.

In contrast to the EM-interactions, the hadronic interaction models for multiparticle production are currently one of the major sources of theoretical uncertainty in expectation for most of the EAS observables [342, 343]. The current hadronic event generators generally rely on the extrapolation of laboratory/accelerator measurements. The extrapolation is needed since the highest energy CRs have energies which are multiple orders of magnitude higher than any current earth-bound accelerators, and at energies exceeding 10^{15} eV only indirect observation of CRs is feasible. These are generally the source of the largest systematic uncertainties. Over the years multiple hadronic event generators and their versions (and their corresponding improvements using LHC data), which extend to the highest energy CRs, have been proposed. Among these DPMJET [**fedynitch2015cascade**, 344, 345], EPOS [346–349], QGSJET [350–355], SIBYLL [356–360] are the most prominent ones. The models updated and tuned to agree with the LHC-data are generally regarded as post-LHC models. SIBYLL 2.3, QGSJET II-04 and EPOS-LHC are the post-LHC models. These are also accompanied by low-energy models (with a particle production threshold of $\approx$ 200 GeV) like GHEISHA [361], FLUKA [362], UrQMD [363]. Similar to EM event generators, a final spatial and kinematic distribution of all the particles generated is produced.

The hadronic interactions depend on strong interactions which can be explained using Quantum Chromodynamics (QCD) [364, 365]. These models rely on data from accelerators to model and explain multi-particle production at high energies. Accelerators primarily measure interaction processes with large momentum transfer, which can be modeled using perturbative-QCD. However, the development of EAS is predominantly influenced by particles generated in the forward region. Hence a multi-pronged approach using perturbative-QCD, theoretical constraints, and phenomenological modeling is employed to provide a comprehensive understanding of multi-particle production at accelerator energies and extend it to the highest-energy CRs observed at EAS-observatories. Based on the choice of hadronic model the particle content can be very different. This is because of different modeling of hadronic interactions and needs to be considered when interpreting any physics results. This work uses SIBYLL 2.1, SIBYLL 2.3c, EPOS-LHC and QGSJET II-04 as the high-energy models ($E > 80$ GeV), for calculating expectations and making comparisons between multiple EAS observables. FLUKA is chosen as the low-energy model. A few important aspects regarding the high-energy hadronic models are:

- **SIBYLL 2.1 & 2.3c** : SIBYLL is based on dual parton model [366, 367], Lund Monte Carlo algorithms [368], and the minijet model [369–372]. It uses Glauber

5 EGS provides and estimate of momentum, spatial position as well as propagation time. NKG is an analytic method and is used to only get quick estimates of electron densities.

scattering theory [373] for calculation of total interaction cross sections and semi-superposition model [356] for explaining multi-nuclei interactions. It is optimized for describing interactions in EASs and uses minimum assumptions[6]. SIBYLL 2.1 treated the issue of wide expectation of secondary particle multiplicity, particle-multiplicity growth rate (with energy) and other discrepancies with data, as seen in its previous versions [374]. [358] provides details of SIBYLL 2.1. SIBYLL 2.3 was the recent update of the model using LHC-data [375]. Improved production of baryon-antibaryon pairs, the phenomenological model for the production of charm particles were among the major additions to this model. The charm-production channel introduced in SIBYLL 2.3 is responsible for the production of high energy muons and neutrinos[7], and hence it is among the standard hadronic-model used in most air-shower observatories. However, the results from ICNO [9, 174, 181], seem to favor SIBYLL 2.1 as the model most compatible with real-data. This is in contrast to measurements from observatories like Pierre-Auger, which seem to prefer post-LHC models [379][8]. Ongoing studies are trying to understand the reason behind this [174, 380]. This work also uses SIBYLL 2.3c which was tuned to minimally violate the Feynman scaling of leading particle distributions, to obtain a better description of NA49 data [381].

- **EPOS-LHC** : EPOS is based on microscopic semihard Pomeron model [346, 347, 382, 383] and use Reggeon Field Theory framework [384] to model non-perturbative interactions using explicit calculations [385, 386]. The model parametrizations are further tuned to accurately reflect data gathered from numerous accelerator facilities. EPOS-LHC (based on EPOS 1.99 [347]) was fine-tuned to fit the public data release from LHC experiments in 2009 [348]. It improved on the previous models by introducing an effective flow parametrization in the core-corona model used earlier [348]. It shows very good agreement with transverse momentum measurements, however, doesn't use the charm hadrons as SIBYLL[9]. A summary of EPOS-based models can be found in [347, 349].

- **QGSJET II-04** : QGSJET was originally based on QGS model [387] and has been updated to use semihard Pomeron model [388, 389] and uses higher order calculation for multi-Pomeron interactions. It has the least number of free parameters among the hadronic models[10]. Similar to EPOS-LHC it also doesn't treat charm hadrons. [351, 354] provides a brief overview of the hadronic model.

Based on the choice of a hadronic interaction model, the expectation values of shower observables can be significantly different. This can be expected from Figure 5.1. The figure presents the average muon-multiplicity of post-LHC hadronic-models w.r.t. SIBYLL 2.1. As can be seen the differences can amount for an increase in multiplicity of as big as 60%[11]. Similar discrepancies among-hadronic models can also be seen for larger

6 It is hence regarded as a minimum bias model.

7 It can hence also shed light onto discrepancy seen between simulated and observed muon-number at multiple observatories (and also for other hadronic models) [33, 360, 376–378].

8 Similar conclusions are also drawn from LHC data [359].

9 A recent version of EPOS has introduced charm production as well.

10 The choice is made to reduce uncertainties due to extrapolation. However, as a consequence, the final stage hadronic interactions are less detailed.

11 A recent update of SIBYLL 2.3c predicts much lesser muon-multiplicity [390], although still greater than other post-LHC models.

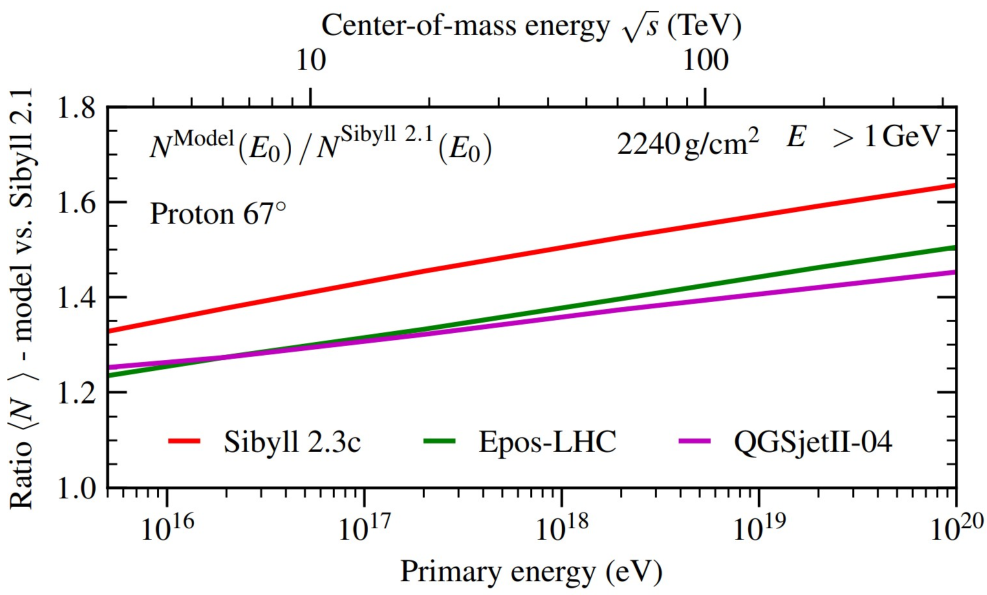


Figure 5.1: Ratio of average muon-multiplicity prediction for post-LHC hadronic-models w.r.t. SIBYLL 2.1. Plot from [360].

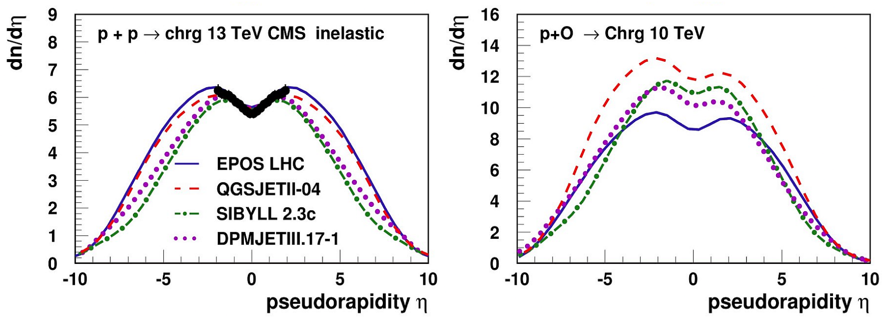


Figure 5.2: Spectra for charged-particle as a function of pseudorapidity, for the choice of different hadronic-models. Data from [391] is also shown. Plot from a talk given by Tanguy Pierog at International Symposium on Very High Energy Cosmic Ray Interactions, Nagoya, Japan (May 2018) [392].

pseudorapidity (η). This is depicted in Figure 5.2. Based on the choice of rapidity, the difference among hadronic models for p-O collisions at 10 TeV can grow as big as 50%. Since, this works probes multiple EAS-shower observables which depend both on muon-multiplicity and their lateral spread (details in Chapter 7). It is crucial to understand how the choice of hadronic models affects the expectation from shower-observables. It is crucial to interpret cosmic-ray composition at ICNO. This will be discussed in detail in Chapter 7 and Chapter 9.

5.1.3 *Thinning*

For CR primaries with energy exceeding 10^{16}eV the computational and memory requirement to track all secondary particles in CORSIKA becomes excessive. As an example, a 10^{20} eV shower has $O(10^{13})$ particles[12]. To overcome this issue *thinning* (also called as *thin sampling*) is used [393]. It allows us to select a representative sub-sample of all the particles. The standard thinning process involves the following steps:

- **Step 1**: Define a thinning level ($\epsilon_{Th} = E/E_0$, where E is a selected thinning energy)
- **Step 2**: For all j secondaries with sum-energy less than thinning energy (i.e. $\epsilon_{Th}E_0 > \Sigma_j E_j$), select at random only one secondary with probability $p_i = E_i/(\epsilon_{Th}E_0)$
- **Step 3**: Give the surviving secondary a weight factor $w_i = 1/p_i$

All the surviving secondaries after thinning follow the standard CORSIKA particle propagation. Thinning enables the tracking of a nearly uniform number of particles, rather than an exponentially increasing one. The work by M. Kobal et. al. [394] improves this further by introducing separate weight limitations for different components of EASs. This is termed as the optimized thinning procedure and also helps reduce the additional artificial fluctuations introduced in EASs because of the thinning procedure[13]. Thinning is useful to reduce the computational time, however, excessive thinning can lead to information loss leading to larger uncertainties in expectation of EAS-observables.

For EAS-simulations at ICNO, showers up to an energy of 100 PeV are simulated without any thinning. Above this energy range, only thinned showers are produced. In the energy range between 10 PeV to 100 PeV, both thinned and unthinned showers are produced[14]. A thinning-level for EM-component, upto an energy of $10^{8.4}$GeV is (i.e. ϵ_{Th}^{EM} =) 10^{-6} [395]. Above this energy $\epsilon_{Th}^{EM} = 273/E_0$ up to an energy of 273 PeV, after which it is kept constant at 273 PeV. The muonic and hadronic component is not thinned for EAS-simulations at ICNO. During detector simulation the thinned shower is unthinned. Section 5.2.1 will give a brief overview of the dethinning procedure.

5.2 DETECTOR SIMULATION

After the simulation of EAS using CORSIKA is completed, the detector response to the CORSIKA-showers need to be simulated. For this work, the particles at a height of 2837 m above sea level are read out. These are then used to propagate throughout the detector to understand their response on detector components. However, before propagating the particles and understanding the detector response (read Section 5.2.3 for details), two additional factors need to be considered, namely *Dethinning* and *Resampling*. The following text will give a brief summary of both.

5.2.1 *Dethinning*

In order to perform the detailed shower simulation the thinned showers (read Section 5.1.3 for details) need to be dethinned. This can be done using the weights (w_i)

12 Simulating fluorescence and Cherenkov light requires tracking significantly more photons than particles. The number of photons is at least three orders of magnitude higher.
13 In [394], an optimized thinning technique is demonstrated and the impacts on detector simulations are analyzed.
14 for consistency and validity checks

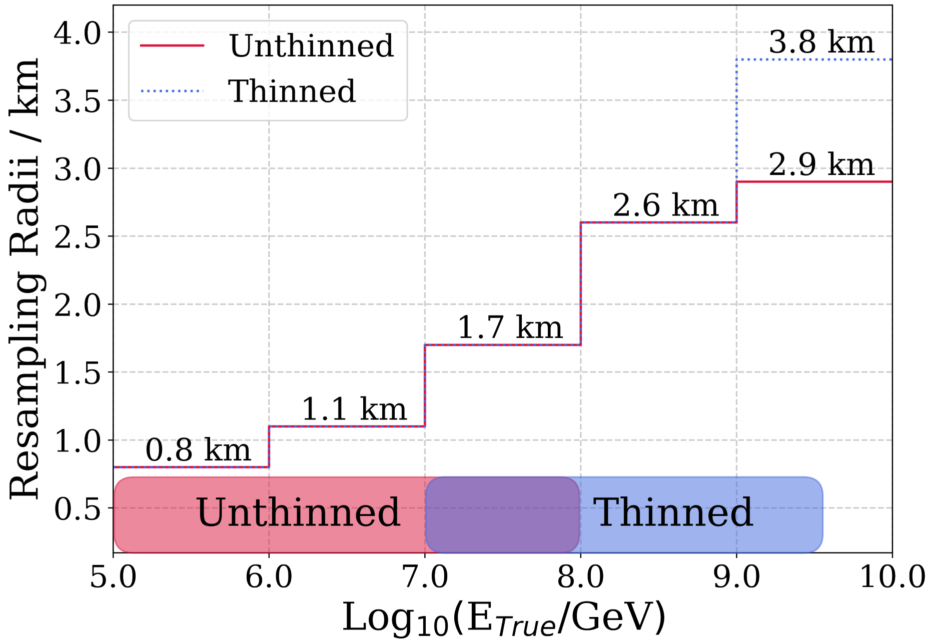


Figure 5.3: Resampling radii for IT as a function of true MC-primary energy, for thinned and unthinned showers. For $7 \leqslant \text{Log}_{10}(E_0/\text{GeV}) < 8$, both thinned and unthinned showers are produced. For CR analysis at ICNO, only events with $\text{Log}_{10}(E_{True}/\text{GeV}) \leqslant 9.5$ are produced.

assigned to the surviving during the thinning procedure. The algorithm detailed in [396] is used for the dethinning procedure. It involves artificially increasing the acceptance area around the tank (/detector). For this work, the increased sampling area for each tank at IT is obtained by minimizing the following term:

$$\min\left(\left|\frac{\Sigma_i w_i \cdot A_i^p}{A_{sampling}} - n_{sampling}\right|\right) \tag{5.1}$$

where $A_i^p = \pi \cdot R^2 + 2 \cdot R \cdot H \cdot \tan\theta^p$ is the area of the tank (with radius = R and height = H) seen by a particle incident at a zenith angle of θ^p. $n_{sampling}$ is the number of particles within the sampling area. [201, 395, 397] provide a detailed summary of the dethinning procedure.

5.2.2 *Resampling*

As already noted, producing MC CORSIKA-simulations is a computation and memory-intensive task. Simulation of a single $\mathcal{O}(10^4$ GeV) air-showers takes about a few minutes. Whereas for HE-EAS with energy $\mathcal{O}(10^8$ GeV), the simulation time extends to a couple of days for each shower. After detailed CORSIKA simulations, a detailed detector response simulation is also required. This limits the number of simulations that can be performed in a reasonable timeframe. However, in order to make statistically significant assertions, the majority of the analysis conducted within ICNO (as well as other observatories) require a substantially larger number CORSIKA simulations than

# Showers	Primary	$\log_{10}(E_0/\text{GeV})$	Zenith	HE Model	Thinned
20000	p	5-8	0-65	SIBYLL 2.1	No
20000	Fe	5-8	0-65	SIBYLL 2.1	No
20000	He	5-8	0-65	SIBYLL 2.1	No
20000	O	5-8	0-65	SIBYLL 2.1	No
24000	p	7-9.5	0-65	SIBYLL 2.1	Yes
24000	Fe	7-9.5	0-65	SIBYLL 2.1	Yes
24000	He	7-9.5	0-65	SIBYLL 2.1	Yes
24000	O	7-9.5	0-65	SIBYLL 2.1	Yes
6000	p	5-8	0-65	SIBYLL 2.3	No
6000	Fe	5-8	0-65	SIBYLL 2.3	No
6000	p	5-8	0-65	EPOS-LHC	No
6000	Fe	5-8	0-65	EPOS-LHC	No
6000	p	5-8	0-65	QGSJet-II.04	No
6000	Fe	5-8	0-65	QGSJet-II.04	No

Table 5.2: Details of CORSIKA datasets used for this work.

are available. However, detector simulation for an EAS at ICNO requires far less time than that required for CORSIKA simulation (1:5 time ratio). This motivates reusing the generated CORSIKA simulations. Hence, a working solution to overcome this situation was obtained by taking each CORSIKA shower and injecting it at random locations on IT a few times. These are then propagated through the detector simulation chain. Each simulated shower is resampled 100 times to limit biasing of the dataset by excessive resampling. The resampled showers are randomly injected within a circular region around the center of IT. The relationship between energy and resampling radius for both unthinned and thinned showers at IT is illustrated in Figure 5.3 (for showers with zenith angle < 40 deg). The dataset used for this analysis is depicted by the shaded boxes at the bottom. The resampling radii are chosen in such a way that the farthest located resampled showers in an energy bin can still trigger the detector. With an increase in the energy of the primary, the size of the shower footprint grows. Hence, with an increase in energy, the resampling radii are also increased accordingly. Since for showers with zenith angle > 40 deg the shower footprint at the detector can be even larger, a future extension of this analysis will need a revaluation of resampling radii.
Table 5.2 summarizes the details of CORSIKA datasets used for this work. Due to computational constraints, only p and Fe datasets have been simulated for hadronic models, with the exception of SIBYLL 2.1. New simulations with bigger shower libraries and more recent hadronic models are being produced and can be used to update this analysis in the future.

5.2.3 *Detector Response*

After dethinning and injecting the CORSIKA-generated particles' shower core at a random location within the resampling radii, a detailed detector simulation procedure is performed. For IT, only those particles which are expected to fall within 30 cm around any tank are kept. The surviving particles are then propagated through air and snow[15] using GEANT4 [398]. To simulate the response of the tank to the particles, a parametrization is used instead of simulating the full Cherenkov light response (and the corresponding photon propagation) of the particles which traverse through the tank [399, 400][16]. The time response is modeled as an exponential distribution[17]. After obtaining the charge and timing response, PMT effects (amplification, charge, and time smearing) and other effects (pre and after pulses, PMT saturation, electronic noise) are added [162, 164]. These are then used for triggering (read Section 3.1.7 for details).

To get the response of EAS at IC, only muons with $E_\mu > 273$ GeV are propagated to the depth of IC. It is the energy at which about 0.1% of the surface-muons reach the depths of IC. A brief derivation of this is given in Section A.2. The in-ice muon propagation for IC is done using Propagator with Optimal Precision and Optimized Speed for All Leptons (PROPOSAL) [402]. This is used to get the energy loss by the particles (primarily TeV muons) and obtain the generated secondaries. In order to determine the Cherenkov-photons flux (and their time distribution) produced by the propagated particles a hybrid approach is used. These are used to propagate the photons produced directly by the muons. The photons created during cascades use PHOTONICS's lookup tables [403]. The lookup tables are binned in six dimensions (space, time, incident photon angle, and photon emission angle), to determine the photon flux (and their time distribution)[18]. Ice properties for IC (like absorption, scattering, etc.) are also considered during this process[19]. After this thermal and non-thermal noise[20] is added to the simulations. Similarly to IT, the electronic effects are finally added before using them for further processing for a physics analysis.

5.3 RECONSTRUCTIONS

After an event has been simulated or detected several reconstruction methods and quality-cuts are used or developed by analyzers based on the requirement of the analysis. The reconstructions can then be used to compare observables between simulations and real data. They can also be used for understanding discrepancies between simulations and data, as well as for using them for high-level physics analysis. Within ICoII, this process occurs at multiple levels, following a hierarchical approach. Moreover, based on the requirement of analysis, it can be done in an offline as well as

15 As depicted in Figure 3.8, the snow height above each tank is different and changes over the years. Different simulation sets are created every three consecutive years.
16 This parametrization is slightly different for each tank, to account for differences in optical properties. Vertical muons are used to perform the calibration [401] (Read Section 3.1.8 for details.).
17 This also treats each tank differently, considering the reflectivity of the tanks.
18 A multi-dimensional penalized spline surface technique is used to efficiently compute interpolated values [404].
19 SpiceMie ice model is used [316].
20 To account for radioactive decay in PMT and glass sphere

online manner[21]. The following section will give a brief overview of the reconstruction methods relevant to this analysis. Finally, the quality cuts used in this analysis will be discussed.

5.3.1 *IceTop Reconstructions*

IT is primarily used as a CR-observatory. To conduct most physics analyses utilizing the extensive air-shower (EAS) footprint, several fundamental pieces of information are essential. These include details such as the shower core, direction, and energy of the primary cosmic ray (CR). The core location and the direction of an observed EAS can be reconstructed from the signal information from the IT tanks. Using a radial distribution of charges and arrival time serves as the stepping stone to deduce the Lateral Distribution Function (LDF) and shower curvature. From analysis and experience at other CR-observatories, the local density at a certain distance distances from shower-core is known to be a good CR-energy estimator, as well as provide hints of CR composition [181, 405–409][22]. At IT a first-guess of core position and shower direction is performed. It then serves as a seed for LDF and curvature reconstructions.

5.3.1.1 *First Guess*

A fast and robust estimate of core location and direction is important for almost every likelihood minimization algorithm discussed later. In addition to this, they also play a crucial role in realtime-alerts sent by ICNO [190]. The following text will give a brief summary of the first-guess reconstructions.

CORE-POSITION: Being one of the most simplest and rational solutions, the first guess of core position is obtained using Center-of-Gravity (COG) of IceTop hit-tanks. This can be represented as:

$$r_{COG} = \frac{\sum_i^{nTanks} r_i Q_i^w}{\sum_i^{nTanks} Q_i^w} \tag{5.2}$$

where in Equation 5.2, r_i refers to the position of tank i; Q (VEM units) represent charge measured with tank i and weighted with factor w. $nTanks$ definition can vary depending on the tanks we want to select (e.g. all hit-tanks, hit-tanks passing a predefined threshold). In [410], choosing 7 tanks with the largest signals and $w = 0.5$ was determined as the most optimal choice for obtaining the best core-resolution[23].

SHOWER-DIRECTION: A first guess of shower direction can be obtained by minimizing the time difference between measured hit-times (t_i^{meas}) and the expectation

21 A detailed summary of the hierarchical steps can be found in [201].
22 The reference distance can change, depending on the observatory and the energy-range.
23 A limitation of this method is that it is unable to provide a good reconstruction for triggered showers with shower-core outside the IT array. However, these showers are already removed by the quality cuts used in this analysis. Ongoing efforts are trying to improve the reconstructions for uncontained EAS showers at IT.

(t_i^{plane}) from a plane shower-front approximation. This can be done using a chi-squared (χ^2) minimization, represented as:

$$\chi^2 = \sum_i \left(\frac{t_i^{meas} - t_i^{plane}}{\sigma_i} \right)^2 \tag{5.3}$$

where the summation is done over all HLC Tanks. A constant time fluctuation σ of 5 ns is used as an approximation. Additionally, in the first iteration, all tanks are assumed to be at the same height[24]. The expected time at a tank i with coordinates (x_i, y_i, z_i) is given by:

$$t_i^{plane} = t_0 + \frac{(\mathbf{r}_i - \mathbf{r}_c) \cdot \mathbf{n}}{c} \tag{5.4}$$

where c denotes the speed of light. An approximation of light-speed travel for shower-front propagation is taken. $\mathbf{r}_c$ and $\mathbf{r}_i$ are position-vectors to shower-core and tank-position respectively. $\mathbf{n}[= (-\sin\theta\cos\phi, -\sin\theta\sin\phi, -\cos\theta)]$ denotes the shower-direction (θ : zenith, ϕ : azimuth) in IceCube-coordinates. After the first iteration, a height correction is applied to all signals (details in [410]).

5.3.1.2 *Lateral Distribution & Curvature*

A multi-component likelihood is used to improve on the first-guess methods. The full likelihood is given by:

$$\begin{aligned} \mathrm{Total\ Likelihood} \equiv\ & \mathrm{Charge\ Likelihood} + \mathrm{Time\ Likelihood} + \\ & \mathrm{No-Hit\ Likelihood} + \mathrm{Saturation-Correction\ Likelihood} \\ L(\mathbf{r}_c, t_0, \mathbf{n}, S_{ref}, \beta) =\ & \mathcal{L}_q(r_c, S_{ref}, \beta) + \mathcal{L}_t(r_c, t_0, \mathbf{n}) + \\ & \mathcal{L}_{no-hit}(r_c, S_{ref}, \beta) + \mathcal{L}_{sat}(r_c, S_{ref}, \beta) \end{aligned} \tag{5.5}$$

where $\mathbf{r}_c = (x_c, y_c)$ are the shower-core's coordinates; t_0 is the time at which shower-front passes the core; $\mathbf{n}$ denotes the shower-direction; S_{ref} and β are LDF parameters. The details about the individual likelihood components are:

CHARGE LIKELIHOOD Charge Likelihood is given by:

$$\mathcal{L}_q = \prod_i^{nHitTanks} P(S_i | S_{exp,i}) \tag{5.6}$$

where S_i^{exp} is the expected charge at a distance R_i from shower-core; the summation runs over all unsaturated tanks. IT measures energy-deposits rather than particle densities. The detected signal has a nearly linear correlation with the energy deposited by shower secondaries in the tank [410]. Hence, instead of using NKG-function [411, 412], a dedicated analysis [413] was performed to get the LDF and is given by:

$$S_{exp}(r) = S_{ref} \left(\frac{r}{R_{ref}} \right)^{-\beta - \kappa \log_{10}\left(\frac{r}{R_{ref}}\right)} \tag{5.7}$$

24 There are differences in tank heights at IT, as can be seen from Figure 3.7.

where S_{ref} is the signal at reference distance R_{ref}; β and κ ($= 0.303$) are respectively the slope and curvature of the LDF at the reference distance.The normalization constant in Equation 5.7 i. e. S_{ref} is referred to as shower-size and was chosen to be a good-estimator of the true primary energy (performance in Section 5.3.4); with minimal fluctuations for showers with the same energy and minimal dependency on primary-type [414]. In [415] a perpendicular distance of 125 m from the shower axis was found to be the best choice of reference-distance (i. e. $R_{ref} = 125\ \mathrm{m} \implies S_{ref} = S_{125\ \mathrm{m}}$)[25]. Also, the fluctuation of signal expectation is given by [202]:

$$\log_{10}\sigma_{q,i} = \begin{cases} 0.283 - 0.078 \cdot \log_{10} S_{exp,i} & \text{if } \log_{10} S_{exp,i} < 0.340 \\ -0.373 - 0.658 \cdot \log_{10} S_{exp,i} & \\ +0.158 \cdot \log_{10}^{2} S_{exp,i} & \text{if } 0.340 \leqslant \log_{10} S_{exp,i} < 2.077 \\ 0.0881 & \text{if } 2.077 \leqslant \log_{10} S_{exp,i} \end{cases} \tag{5.8}$$

.

As shown in Figure 3.8, the snow height over tanks changes over time. The LDF (given by Equation 5.7) doesn't account for the snow height. However, variation in snow height is responsible for attenuation of signal. A greater snow height causes more attenuation of the EM-component of EAS (slightly to the muonic component), leading to a smaller deposit than expected from Equation 5.7. A snow correction is applied to the fit value of the signal. An exponential absorption model is assumed for the correction and the correction is given by:

$$S_{exp,i}^{corrected} = S_{exp,i} \cdot \exp\left(\frac{h_i^{snow}}{\lambda_{eff} \cdot \cos(\theta)}\right) \tag{5.9}$$

where h_i^{snow} is the snow-height over the i^{th}-tank and θ is the zenith angle. λ_{eff} is effective attenuation length. For this analysis $\lambda_{eff} = 2.25\ \mathrm{m}$ [20]. Ongoing work is trying to improve the uncertainty on the λ_{eff} (± 0.2 m).

The total charge likelihood is given by:

$$\mathcal{L}_q = \prod_{i}^{nHitTanks} \frac{1}{\sqrt{2\pi}\sigma_{q,i}} \cdot \exp\left(-\frac{\log_{10} S_i - \log_{10} S_{exp,i}}{\sqrt{2}\sigma_{q,i}}\right)^2 \tag{5.10}$$

TIME LIKELIHOOD Time Likelihood is given by:

$$\mathcal{L}_t = \prod_{i}^{nHitTanks} P(t_i | t_{exp,i}) \tag{5.11}$$

Studies showed that the fluctuation from plane wave approximation (given by Equation 5.4) can be best explained by a summation of a parabola with a Gaussian-nose, symmetric around the shower-axis [415]. The correction-factor to it is given by:

$$\Delta t(R_i) = a \cdot R_i^2 + b \cdot \left(1 - \exp\left(-\frac{R_i^2}{2\sigma^2}\right)\right) \tag{5.12}$$

25 Because low-energy showers have a smaller lateral spread, the reference distance is set at 80 m. However, this work doesn't use these showers. Ongoing work is trying to optimize the fit for low-energy showers [289]. Also, ongoing efforts are trying to have a multi-component LDF at IT, to describe the contribution for EM and muonic component separately [416, 417]

where $a = 4.823 \cdot 10^{-4}$ ns/m^2, $b = 19.41$ ns and $\sigma = 83.5$ m. Hence, expected time is given by:

$$t_{exp,i} = t_i^{plane} + \Delta t(R_i) \tag{5.13}$$

The time likelihood is given by:

$$\mathcal{L}_t = \prod_i^{nHitTanks} \frac{1}{\sqrt{2\pi}\sigma_{t,i}} \cdot \exp\left(-\frac{t_i - t_{exp,i}}{\sqrt{2}\sigma_{t,i}}\right)^2 \tag{5.14}$$

where $\sigma_{t,i} = 2.92 + 3.77 \cdot 10^{-4} \cdot R_i^2$ is the fluctuation in arrival times [418][26].

NO-HIT & SATURATION-CORRECTION LIKELIHOOD To not bias the fit by untriggered tanks a no-hit likelihood is used. It is given by:

$$L_{no-hit} = \prod_i^{nUnHitTanks} (1 - P_{hit,i}^2) \tag{5.15}$$

where $P_{hit,i} = 1 - P_{no-hit,i}$ and $P_{no-hit,i}$ is given by:

$$P_{no-hit,i} = \frac{1}{2}\left(\text{erf}\left(\frac{\log_{10} S_{threshold} - \log_{10} S_{exp,i}}{\sqrt{2}\sigma_{q,i}}\right) + 1\right). \tag{5.16}$$

Here $S_{threshold} \approx 0.1657$ VEM is the tank threshold.

Similarly, the saturated DOMs are treated using the following likelihood:

$$\mathcal{L}_{sat} = \prod_i^{nSaturatedTanks} \frac{1}{2}\left(1 - \text{erf}\left(\frac{\log_{10} S_{saturated,i} - \log_{10} S_{exp,i}}{\sqrt{2}\sigma_{q,i}}\right)\right) \tag{5.17}$$

After constructing the individual likelihoods, the total likelihood (given by Equation 5.5) is then maximized. In reality, instead of maximizing the likelihood, the negative log likelihood is minimized. The minimization is performed using *Laputop*-framework, in an iterative manner to ensure a good balance between achieving global minima and computational constraints. The shower-core location, shower-core time, direction, S_{125m}(signal at a distance of 125m from shower-axis), β (slope of LDF) are finally obtained as the fit parameters. The fit performance and the effects of multiple correction factors are shown in [410]. Figure 5.4 depicts and example of the fitting procedure on a custom MC-simulation.. As a final step the SLC-hits within a time window of (-200ns,+800ns) are included to ensure that the information from muons deposits located farther from shower axis is not removed.

5.3.2 *IceCube Reconstructions*

The IC component of ICNO makes it a unique three-dimensional CR-detector. IC primarily detects TeV muons from EAS, and can be used for CR-composition studies at IC. For CR-analysis, IC relies on a variety of reconstructions. The most common among them are MILLIPEDE (read Section B.1 for details) and Data Derived Differential Deposition

26 This is based on old detector configuration. Ongoing work is trying to improve it.

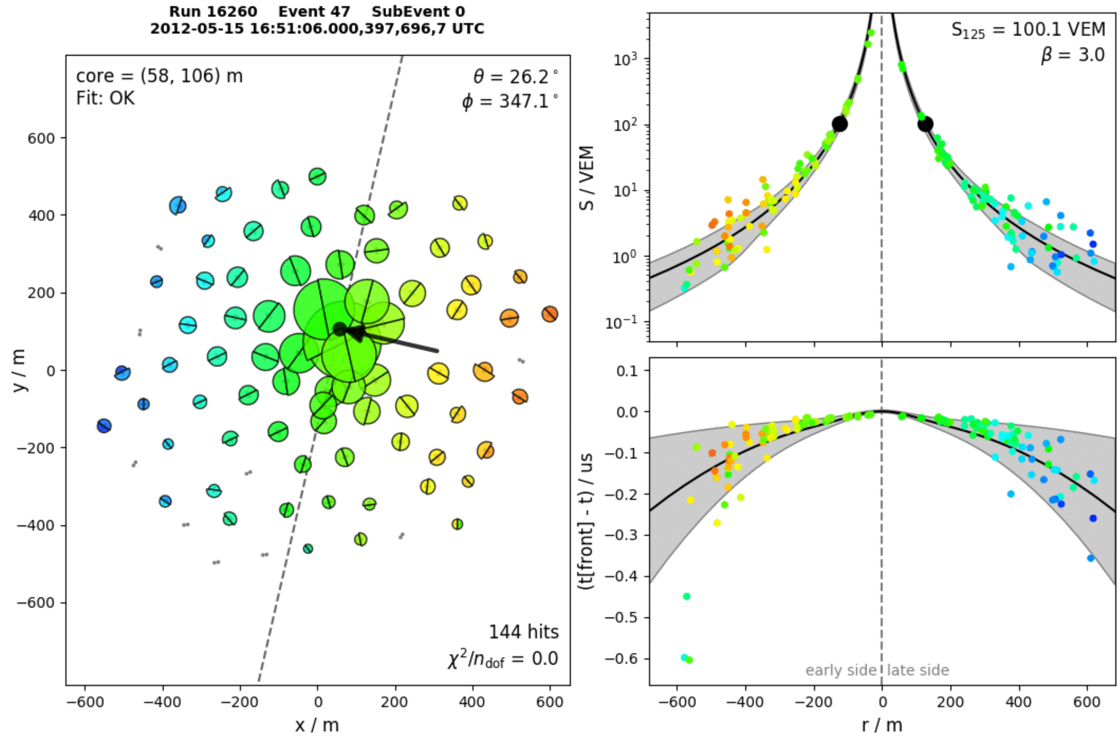


Figure 5.4: **Left:** Footprint for MC-simulation (HLC-hits) of a Proton event with true-energy of 92 PeV and $(\theta_{\mathrm{True}}, \phi_{\mathrm{True}}) = (26.2, 347.3)$ and true core at $(x_{\mathrm{True}}, y_{\mathrm{True}}) = (55.4, 104.8)\mathrm{m}$. The size of the semi-circles represents the charge deposit (cleaned and corrected for snow attenuation) in an IT-tank (grey dots). Early times are denoted by red and later times by blue. The black arrow denotes the axis projection on the IT-pane with the arrow-head at the reconstructed shower core. **Top Right:** Reconstructed LDF from Equation 5.7 with overlapped charge deposits. **Bottom Right:** Time residuals in comparison to plane wavefront, given by Equation 5.12.
The reconstructed core-location (+ zenith and azimuth), $S_{125\mathrm{m}}$ (+β) are also shown in the left and top-right plots respectively.

Quality Cut	Requirement
- IceTopSTA5 Filter	Passed
- Station with maximum charge deposit	Not on edge
- Maximum snow-corrected charge deposit in a tank	$\geqslant 6$ VEM
- Charge-deposit in the adjacent tank with the highest deposit	$\geqslant 4$ VEM
- Core containment fraction i. e. D_{it}/d_{it}(Refer to Figure 5.5)	< 0.96
- Station density of hit-tanks between shower-core and farthest station	> 0.2
- Slope(β) of LDF-fit at IT (given by Equation 5.7)	$1.4 < \beta < 9.5$
- EAS-Reconstruction (Laputop)	Succeeded
- $\text{Log}_{10}S_{125m}$/VEM	$\geqslant 0$

Table 5.3: IT quality cuts used for this analysis. For details read Section 5.3.3.

Reconstruction (DDDDR) (read Section B.2 for details). These are used for reconstructing energy losses by the muon/muon-bundle in IC. Brief summary of the two reconstructions is as:

- **MILLIPEDE:** It divides the track by muon/bun-bundle in IC and unfolds the energy loss profile using the charge deposits at DOMs.
- **DDDDR:** It uses data-driven approach to characterize photon propagation in IC, which is subsequently used to estimate energy. It is also the underlying reconstruction algorithm used to develop new CR composition-sensitive parameters discussed in Section 7.4.

Section B.1 and Section B.2 provide a more detailed summary of the two reconstruction methods. These are then used for deriving multiple composition-sensitive parameters at ICNO. The multiple CR-composition parameters primarily using the reconstruction methods above are discussed at length in Chapter 7.

5.3.3 *Quality Cuts*

To ensure that all the events used for this analysis are well reconstructed and are coincident between IT and IC, they have to pass multiple filters. They can be divided into IT and IC specific filters. They have been listed in Table 5.3 and Table 5.4. A brief description of them is as:

- **IceTop**:
  - **IceTopSTA5 Filter**: As mentioned in Section 3.1.7, at ICNO multiple triggers are used. These ensure that only relevant events are saved and processed, and hence help reduce storage and computation costs. IceTopSTA5 Filter

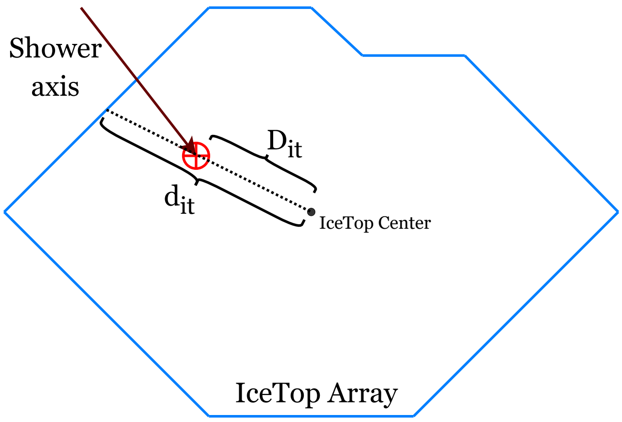


Figure 5.5: IT core containment cut requires $D_{it}/d_{it} < 0.96$.

ensures that after cleaning all events have at least 5 IT-stations containing HLC-hits.

– **Charge Deposits**: IT reconstructions detailed in Section 5.3.1 are based on charge-deposits at IT. To minimize false reconstructions, it is ensured that IT-station with the maximum charge deposit doesn't lie on the edge of the detector. This effectively contains the shower core inside the IT-array. Two further cuts ensure that the maximum charge is indeed because of the shower core. The maximum snow-corrected signal in an IT-tank has to be at least 6 VEM, with the adjacent tank detecting no less than 4 VEM. In addition to ensuring the core-locality at IT, the charge threshold also effectively removes low-energy as well as highly-inclined EASs. As a final shower-core containment filter, the distance of the reconstructed core-location from the IT-center should be no more than 96% of the distance to the edge (from the center) of IT-polygon in that direction. This is depicted by Figure 5.5, and hence $D_{it}/d_{it} < 0.96$ is required.

– **Others**: In order to prevent false-reconstructions of EASs with sparse-hits, the density of hit-tanks to total tanks with a circle centered at footprint's COG (and with radii = distance of COG to farthest tank) should be greater than 0.2. $\mathrm{Log}_{10}S_{125m}/\mathrm{VEM}$ should be greater than 0[27]. Another cut applied on slope (β) of LDF (given by Equation 5.7), ensures that it is between (1.4 , 9.5). It ensures good-quality reconstructions. The Laputop reconstruction also needs to be successful.

- **IceCube**:

  – **IC SMT-Filter**: Similar to IT, an IC multiplicity trigger ensures that there are atleast 8 cleaned HLC-hits remaining in an event.

  – **Reconstructed Energy**: The energy loss for IC is carried out using MILLIPEDE (details in Section B.1). In order to ensure good quality tracks the reduced log-likelihood of the charge deposits should be less than 10^2 and the ratio of predicted to measured charge should be greater than 0.93 (for

27 This cut ensures that the efficiency of observation for all primary types is 100%.

Quality Cut	Requirement
- Laputop track and direction	Pass through IC
- Number of IC DOM-hits	$\geqslant 8$
- MILLIPEDE(details in Section B.1) reduced log-likelihood	$< 10^2$
- Ratio of predicted to measured charge	> 0.93
- No. of track segments with non-zero energy loss (using MILLIPEDE)	> 3
- Energy-loss fit	Succeeded

Table 5.4: IC quality cuts used for this analysis. For details read Section 5.3.3.

details read [410]). In addition to that, there should be at least three cascades with non-zero energy-deposits[28]. Finally, the reconstruction needs to be successful.

– **Other**: Section 7.4 will introduce another cut based on the distance of the DOM from the shower-axis. However, for this analysis, the cut doesn't remove any additional events.

The IT and IC quality cuts implemented together ensure that the analysis only uses high-quality coincident events. The following text will present some performance plots for the reconstructions and quality cuts detailed earlier.

5.3.4 *Performance*

As mentioned in Equation 5.7, S_{125m}/VEM i. e. signal at a perpendicular distance of 125 m from the shower-axis was chosen as a reference distance in the LDF-fit and was chosen to be a good estimator of primary-energy [414, 415]. Figure 5.6 presents the $\mathrm{Log}_{10}S_{125m}$/VEM as a function of true-MC energy, with SIBYLL 2.1 as the underlying hadronic-model. A linear correlation between the two can be observed[29]. The deviation from the linear correlation is generally small. In addition to that, a good overlap between the primary types can also be observed. Hence, $\mathrm{Log}_{10}S_{125m}$/VEM will be used as an energy proxy for the upcoming analysis. However, Section 8.2 will present another improved method that is used for the final analysis. Figure 5.7 presents the distribution of $\mathrm{Log}_{10}S_{125m}$/VEM for various primary types, unweighted to any flux-model. An additional cut of $\mathrm{Log}_{10}S_{125m}/\mathrm{VEM} \geqslant 0$ is also used. Also with increasing $\mathrm{Log}_{10}S_{125m}$/VEM (or equivalently energy), the number of available simulations reduces significantly. Figure 5.8 presents the zenith and azimuth distribution for the events which pass the quality cuts.

28 Dust-layer not included.
29 The distributions have not been weighted with any flux-model.

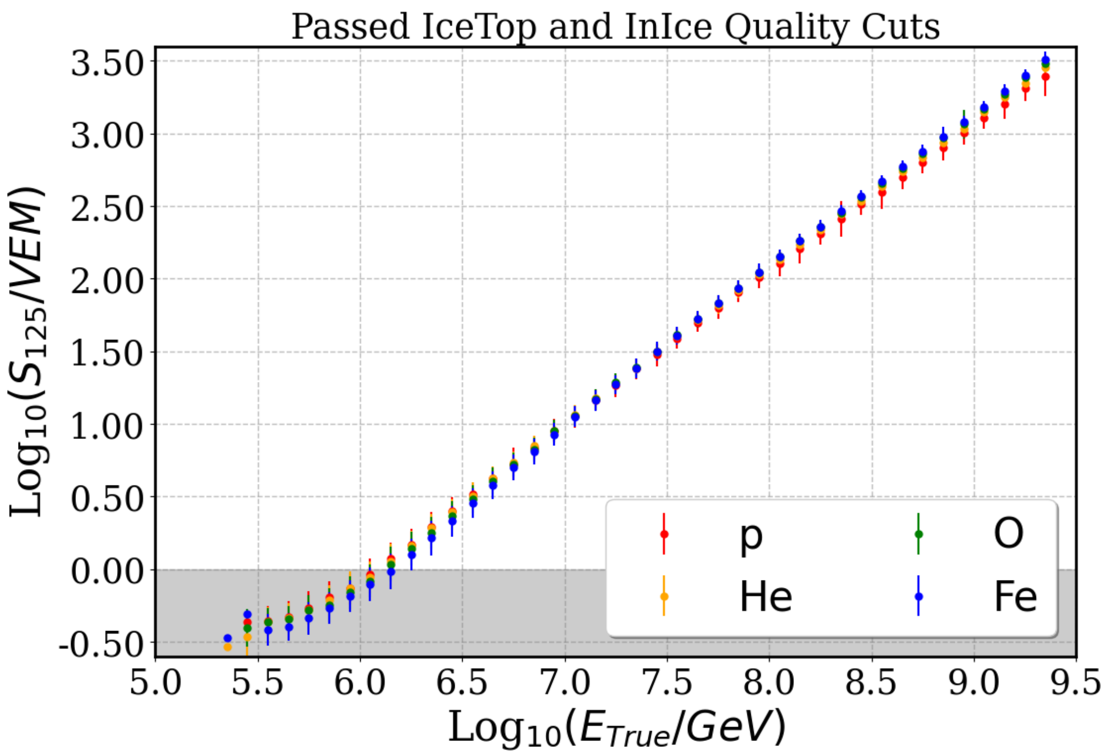


Figure 5.6: $\text{Log}_{10}\text{S}_{125\text{m}}$/VEM as a function of true-MC energy, with SIBYLL 2.1 as the underlying hadronic-model. The underlying MC-simulations have passed quality cuts mentioned in Section 5.3.3. Additionally only events with $\text{Log}_{10}\text{S}_{125\text{m}}/\text{VEM} \geqslant 0$ are used.

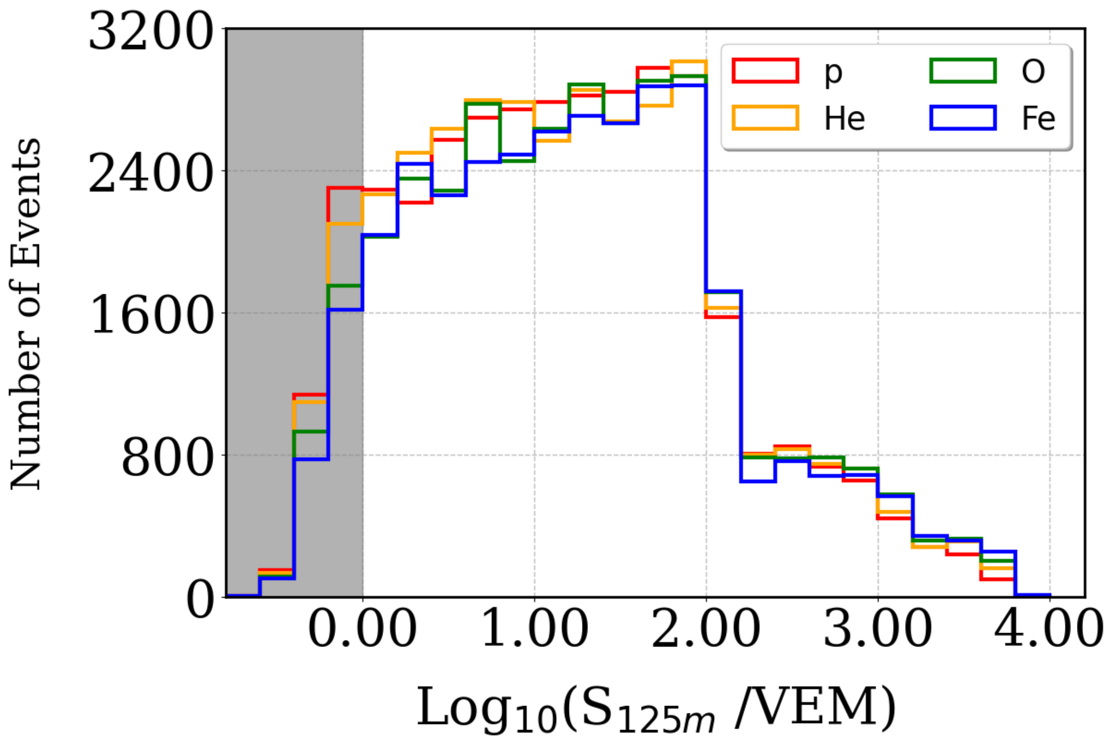


Figure 5.7: $\text{Log}_{10}S_{125\text{m}}$/VEM distribution of MC-simulations (SIBYLL 2.1) which pass quality cuts mentioned in Section 5.3.3. Additionally only events with $\text{Log}_{10}S_{125\text{m}}$/VEM $\geqslant 0$ are used.

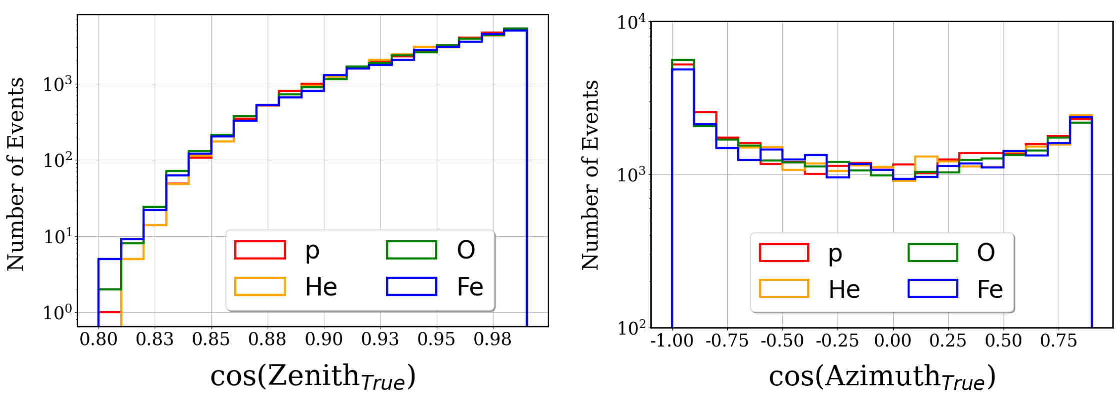


Figure 5.8: Zenith and Azimuth distribution of MC-simulations (SIBYLL 2.1) which pass quality cuts mentioned in Section 5.3.3.

6 GRAPH NEURAL NETWORK - A PRIMER

Machine Learning (ML) is the field of understanding and building algorithms that give machines the ability to learn without being explicitly programmed for the task at hand [419]. A good representation of the input data is generally very crucial to increase the efficacy of ML methods [420]. An ineffective or redundant representation can be inimical to realize the optimal solution. As well-put by LeCun *et al.* in [421], ML for a long time "required careful engineering and considerable domain expertise to design a feature extractor that transformed the raw data" to further learn meaningful representations and patterns from observations. Deep Learning (DL) is a subfield of ML, which focuses on automated representation-learning by composing multiple hierarchical non-linear modules. One of the most basic DL designs is the Multi-Layer Perceptron (MLP), which entails mapping a set of input features onto the target variable(s), while minimizing a pre-defined loss function. This is done by optimizing the weights of connections between the layers in a MLP. However, working with MLP often has a negative impact on model accuracy and increases training time due to the fact that as input feature dimensions and sizes rise, so do the number of trainable parameters. Additionally, we miss out on the opportunity to profit from the inherent spatial, temporal, or other patterns in the input training data. By relying on intrinsic data structures and designing algorithms that can be directly applied to such data structures, we can reduce the cost of having human experts create efficient data representations and achieve greater accuracy. This observation was crucial for the birth [**726791**] and success [**szegedy2015going**, **9451544**, 422, 423] of Convolutional Neural Network (CNN) for image recognition tasks. Similar observations and explorations helped capitalize the power of machine learning in the field of Speech Recognition [424, 425], Natural Language Processing (NLP) [426, 427], and many other areas[1].

A data structure that is prevalent in a lot of natural and artificial systems is graphs. In its simplest form, a graph ($\mathcal{G}$) is defined by a node-set $\mathcal{V}$ (representing entities) and edge-set $\mathcal{E}$ (representing dependencies), i.e. $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Depending on the complexity of a system, the nodes and edges can also be associated with additional information (called attributes). Example of few graph mapped systems (GMSs) are:

- **Natural**
  - **Cosmic-Web:** $\mathcal{V}$ = Galaxies ; $\mathcal{E}$ = Gravitational Bounds
  - **Protein–Protein Interaction Networks:** $\mathcal{V}$ = Proteins ; $\mathcal{E}$ = Interactions
  - **Pandemic Spread:** $\mathcal{V}$ = Census Tract ; $\mathcal{E}$ = Human interactions

1 Convolutional Neural Network (CNN)s take in account key properties of images like the strong correlation with neighboring pixels and invariances (under translation, rotation, scaling). Assimilating sequence and context dependencies in Neural Network (NN) architectures was important for the success of speech recognition and NLP tasks.

- **Artificial**
  - **Social Interaction:** $\mathcal{V}$ = Communities/Individuals ; $\mathcal{E}$ = Overlap/Interaction
  - **Flight-Route Network:** $\mathcal{V}$ = Airports ; $\mathcal{E}$ = Flight Connection between airports
  - **Power-Grid Network:** $\mathcal{V}$ = Power-Generators and Substations ; $\mathcal{E}$ = Transmission lines

As is evident from the examples, a wide variety of systems can be mapped as a graph[2] ($\mathcal{G}$). We can now use these representations to develop fast and effective algorithms by utilizing the advantages of statistical techniques developed in network science [428] and combining them with techniques from ML. This text is also one such exploratory work (and an extended summary of [285]), where we use Graph Neural Networks and our knowledge of EAS-physics to understand CR-composition at ICNO (for more details refer to Chapter 8). This chapter will provide a compact description of Graphs, mathematical ways to represent them, and ongoing research areas to understand the expressivity of Graph Neural Networks.

6.1 GRAPHS - A TOUR D'HORIZON

As aforementioned, a graph ($\mathcal{G}$) is defined by a node-set $\mathcal{V}$ (representing entities) and edge-set $\mathcal{E}$ (representing dependencies) i.e. $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The existence of an edge between two nodes $u, v \in \mathcal{V}$ is denoted as $(u, v) \in \mathcal{E}$. The nodes are sometimes also referred to as neighbors. There are a few properties manifested by most of the GMS, namely:

- The relative placement of nodes is generally irrelevant[3]. The node-connectivity information is adequate to describe the graph structure. Figure 6.1 illustrates this by two different delineations of the identical underlying graph.

Figure 6.1: Different delineations for identical underlying connectivity.

- The information embodied by the graph is invariant with the order in which nodes are indexed. Thus any operation on graphs should satisfy permutational invariance.

2 We make an effort to adhere to the ML community's terminology by referring to the data representation as a "Graph". Data-mining and network-science communities use the term "network" for the same representation.

3 Relative placement and node-labeling order can be crucial if a system has a fixed structure (e.g. spatial, temporal, etc.).

Based on the information flow among the neighboring nodes, the GMSs can generally be categorized into one of the following classes:

- **Undirected :** The information flow is symmetric among neighboring nodes. The flow happens in a bidirectional manner without an inherent preferred direction.
- **Directed :** The information flow is asymmetric. The direction determines the information flow direction.

The categories mentioned here are neither exhaustive nor the only way to categorize graphs. Another categorization can be [429]: 1. Homogeneous/Heterogeneous 2. Static/Dynamic.

6.1.1 *Graph Representation*

For mathematical and computational analysis, information-flow direction and connectivity of a graph-edges is crucial. The study of information propagation can be used to understand the dynamics and details of the GMSs. This can also be used to optimize the information flow direction. This work presents a unique way for edge-connectivity, as a method to introduce inductive bias (based on CR-composition). This is discussed in detail in Chapter 6. A mathematical framework to express graph connectivity is crucial to perform quantitative and qualitative analysis. This can be done by using graph theory, where the graph connectivity is represented by matrices. The most common ways to do such are:

- **Adjacency Matrix ($\mathcal{A}$):**
  It is square matrix representation where the row (and columns) correspond to nodes in a graph. By convention, the presence of an edge between the nodes (i,j) is represented by $\mathcal{A}_{ij} = 1$, and 0 in case no connection exits between the two[4]. The presence of non-zero diagonal elements represents the presence of self-loops. In the case of an undirected graph, the adjacency matrix is symmetric. The adjacency matrix hence can be used to define the neighborhood of each node in a graph. The adjacency matrix for the graphs in Figure 6.1 is given by:

$$\mathcal{A} = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

- **Degree Matrix ($\mathcal{D}$):**
  It is a diagonal matrix that is used to represent the number of neighbors (degree)

4 The sum of row/column values represents the degree (number of neighbors) of a node. For graphs with weighted edges, the adjacency matrix elements can be different from 0 & 1.

belonging to each node in a graph. Hence, except the diagonal entries $\mathcal{D}_{ii}$, all other values are zero[5]. The degree matrix for the graphs in Figure 6.1 is:

$$\mathcal{D} = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}.$$

The adjacency ($\mathcal{A}$) and degree-matrix ($\mathcal{D}$) can be used to evaluate Laplacian matrix $\mathcal{L}$ $(= \mathcal{D} - \mathcal{A})$, which can be used to find clusters in a graph. [430] provides an overview of multiple other properties, including node centrality, clustering coefficient, which can be used to provide additional insights into graph connectivity and the significance of individual nodes.

6.1.2 *Graph Attributes*

We frequently have richer datasets with more details about the entities and relationships involved. As an illustration, in the Cosmic Web system previously stated, the nodes (Galaxies) can incorporate supplementary information such as mass, age, and velocity, while the edges (Gravitational Bounds) may have additional details regarding the gravitational strength. Figure 6.2 illustrates an example graph where each node has two attributes (color just for visualization) and edges with custom weights.

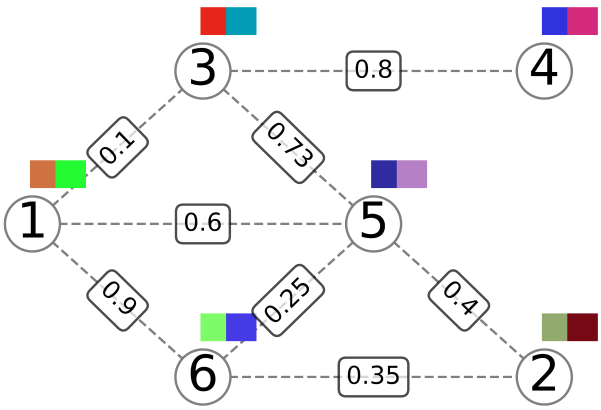


Figure 6.2: An example graph with node-attributes and weighted edges.

5 In a system with predefined edges, a degree matrix can help identify the relative importance of nodes. A node with more connections (higher degree) can generally be considered as a central and prominent node in a graph e.g. an airport ($\equiv$ node) with more flight connections ($\equiv$ edges) is probably an international ($\equiv$ important) airport.

6.2 GRAPH-BASED LEARNING

The success of CNNs (and other simpler architectures) is generally limited to tasks with grid-structured data-structures[6]. Graphs extend this ability to forecast and recognize patterns in data that are difficult to express in Euclidean space. As is evident from the discussion at the beginning of this chapter and Section 6.1.1, Graphs provide a mathematical framework to represent complex systems with variable geometries, relations, or interactions. This flexibility has allowed the application of Graph Theory based methods[7] in a variety of application areas [433–437]. The remarkable success of deep learning (DL) methods in handling grid-structured data has motivated the emergence of a new field: deep learning on graphs. This field leverages the principles of Graph Theory as a solid foundation. It is called by different names like Graph Representation Learning [438] or Geometric Deep Learning [439]. We will try to use an umbrella term of Graph Neural Network (GNN) as a method encompassing different ways of learning with graph data, which use back-propagation [440] as their backbone algorithm[8]. A reason for rapid growth in the field of GNNs is the release of standardized benchmark datasets like Open Graph Benchmark [442][9]. In addition to this availability of modular software libraries like PyTorch Geometric [444], Deep Graph Library [445], TensorFlow-GNN [446] have allowed that the new developments in the field of GNNs can be tested and improved upon, by using them on diverse problems spanning various research domains. Recent understanding of GNNs has helped understand that the majority of other DL architectures can be seen as a special case of GNNs [438, 447–450]. To demonstrate the importance of understanding the operative mechanism of GNNs, it is critical to first learn the variety of problem areas where they have previously been applied and demonstrated superiority over other DL methods:

- **Node Classification**: It is the task of predicting labels ($y_u \forall\ u \in \mathcal{V}$), given the label and neighborhood information about a subset of nodes. For example, Protein folding [451], Protein-function classification [452] ; document classification [453] (for more details see [454, 455]).

- **Link Prediction**: It is the task of predicting neighborhood information (e.g. edge presence/absence, interaction strength, etc.), given the label and neighborhood information about a subset of nodes. For example, Recommender Systems [456] ; Polypharmacy side-effects [457] (for more details see [458, 459]).

- **Cluster Detection**: It is the task of predicting community structures, given the label and neighborhood information for the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. For example, Disease-Pathways [460], Fraud Detection [461] (for more details see [462, 463]).

- **Classification and Regression**: It is the task of learning across a dataset of numerous graphs to predict a graph-level representation, given the label and neighbor-

6 CNNs are also generally limited to rectangular-grid structured data only. This is because it is easier to represent grid-structured data as a matrix which can then be used for mathematical and computational operations. Some works [431, 432] have tried to adapt standard CNNs to non-rectangular data-structures.

7 Read Chapter 2-4 of [430] for a brief review of some of the methods.

8 Recently, an algorithm to train ML-models called Forward-Forward algorithm has been proposed [441]. It replaces the standard forward-backward pass (which uses back-propagation), with two forward passes and removes the need to store gradients.

9 ImageNet [443], a standardized dataset was key to the development of many novel algorithms and architectures in the field of computer vision.

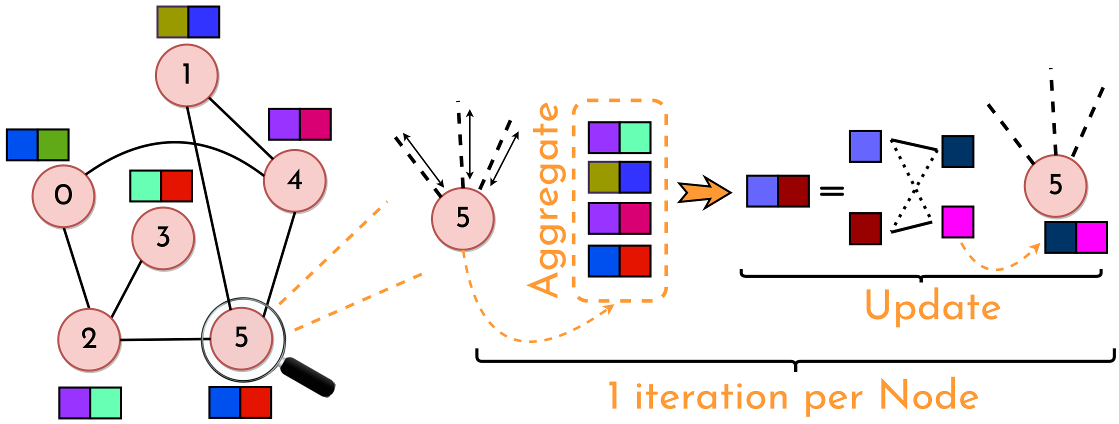


Figure 6.3: Illustration of Message Passing operation. It primarily consists of two operations, namely **AGGREGATE** and **UPDATE**. The operation is repeated at each node (updating the node and/or edge attributes) and this builds up a single layer of a GNN. For details read Section 6.3.1.

hood information for a dataset of graphs. For example, Molecule's Toxicity Prediction [464] ; Malicious software Prediction [465]; Antibiotic Discovery [466] (for more details see [467, 468]). This publication also works on a Graph-regression problem. For details refer to Chapter 6.

6.3 GRAPH NEURAL NETWORK

As is the case for most of the DL methods, GNNs are a method to learn hidden representations/embeddings of input data. Similar to other DL methods, the data-driven representation should depend on the information captured by the structure and the feature information of the data. However, a major design requirement for GNNs is the flexibility to work with variable sizes of input data, in addition to working with non-euclidean/irregularly structured data. The only major requirement among any kind of GNN operation is permutational invariance i. e. it should be independent of the arbitrary ordering of the nodes. The following text will present one method to perform ML on graphs.

6.3.1 *Message Passing*

Message Passing is among the most general signal/information aggregation and sharing mechanisms among nodes of a graph[10]. The message massing operation allows to generate and learn hidden node embeddings which can be used to perform node-level or graph-level predictions. The message-passing operation primarily consists of two operations, namely **AGGREGATE** and **UPDATE**[11]. These are depicted in Figure 6.3 and are utilized in a sequential manner for the following:

10 The other two types namely spectral and attention mechanism can be seen as special cases of it. However, there is no strong agreement in the community about this yet.

11 The scheme followed here is a variation of the many explanations possible to detail the message-passing mechanism.

- **AGGREGATE**: It is an arbitrary permutationally-invariant mathematical operation, which aggregates the information (attributes) at a graph-node from its neighbors[12]. The attributes from the parent-node can also be used. Few examples of such operation can be normalized mean, symmetric normalization [453], etc. This step allows for capturing local structural and relational information in a graph.

- **UPDATE**: This step uses the aggregated information using a differentiable function (generally a NN), to get the updated node representation/embedding. The updated representation can then be used as an input for the next iteration/layer in a GNN. After another iteration, a parent-node will hence have information from a node that is one hop away from it[13].

The above steps can be mathematically expressed as:

$$h_u^{(k+1)} = \mathbf{UPDATE}^{(k)} \left(h_u^{(k)}, \mathbf{AGGREGATE}^{\ (k)} \left(\{h_v^{(k)}, \forall v \in \mathcal{N}(u)\} \right) \right) \tag{6.1}$$

Here $h_u^{(k)}$ is the hidden embedding of node $u \in \mathcal{V}$ at the iteration step k (or k-th layer in the network). $\mathcal{N}(u)$ gives the node's neighborhood. The **UPDATE** and **AGGREGATE** operations are generally permutationally invariant differentiable functions (generally NNs). Hence, during the message passing operation, each node iteratively learns information from nodes with multiple hops away and updates its node attributes. The updated embedding/representations capture structural as well as attribute information of the graph[14]. The node embedding can then be used to make a node-level prediction. In order to make a graph-level prediction the node-embeddings can be globally pooled using a permutational-invariant[15] operation to be then used for a graph-level prediction. In addition to providing flexibility to work with non-euclidean data, recent works [470, 471] have also shown improvements in standard computer-vision tasks (where CNNs used to excel) by using message passing GNNs. Section 8.1.4, Section 8.1.5 and Section 8.1.6 gives detail of specific GNN message-passing operations used in this work. Even though GNNs have shown promise in extending DL to unstructured, non-euclidean / non-uniform datasets, it also suffers from a few issues. Few of the most important limits which are also very crucial in the design choices made in the final GNN-architecture (shown in Figure 8.1) for this work will be discussed in the following text.

6.3.2 *Limits of Graph Neural Networks*

Standard NNs were known to suffer from a multitude of problems like vanishing and exploding gradients, over-fitting, limited data, bias-variance trade-off, adversarial attacks, and label-noise. Solutions have generally been found to allay these issues [472–

12 This step can also take into consideration the presence of any edge-attributes.

13 At the first iteration a parent node collects information from immediate neighbors. However, the neighboring node also gets updated by its neighbors, which might be one hop away from the parent node. Hence, when the iteration happens again, the parent node which now collects information from its updated neighbor will also collect information from a node one hop away.

14 [430] presents a detailed overview of further details and variations of the standard message passing procedure.

15 Studies [469] are testing the possibilities to gain on representation-power by loosening this criterion.

477]. In addition to the still existent remnants of these issues, GNNs are also marred with a variety of other issues. The following text will briefly discuss a few of them[16].

6.3.2.1 *Expressive Power*

MLPs are known for their ability to approximate any measurable function to any desired degree of accuracy, earning them the moniker "universal approximators" [478]. However, understanding and quantifying the expressive power of GNNs is a field of active research. A detailed theoretical understanding of the expressive power of GNNs can be crucial to overcome the current representation limits and establish go-to techniques for different settings of graph-data[17]. One way this can be done is by making connections from GNNs to the problem of Graph-Isomorphism, from the field of Graph Theory. Graph-Isomorphism identifies the existence of an edge-preserving bijection between the vertices/nodes of two graphs. Two graphs, $\mathcal{G}$ and $\mathcal{H}$ are called isomorphic if there exists a mapping f from vertices of $\mathcal{G}$ to those of $\mathcal{H}$, such that if two vertices are adjacent in $\mathcal{G}$ they are also adjacent in $\mathcal{H}$. Identifying two graphs as isomorphic is very closely connected to the research problem explored in this work i.e. cosmic-ray composition. As will become evident from later discussions (Chapter 8) this work intends to identify if a particular graph belongs to one category or another, which can also be performed indirectly by comparing graphs to each other and marking if they are same/similar.

The time complexity of identifying two graphs as isomorphic is currently not known (either solvable in polynomial time [479] or if it is NP-complete [480]). One of the first seminal works to propose an efficient estimate of graph-isomorphism is the Weisfeiler-Lehman (WL) Test [481] (Short Summary - Read Section C.1), by Boris Weisfeiler and Andrey Lehman. The message-passing mechanism of GNNs has strong similarity with the WL-Test [482]. Comparisons with WL-Test allowed us to put theoretical limits on the maximum attainable expressiveness of any GNN architecture. This new-found connection of Graph Theory with GNNs also allowed the birth of architecture like Graph Isomorphism Network (GIN) [482], which is among the most expressive GNN architectures and is as powerful as WL Test. Recent works have extended the WL-Test to higher-order (working on tuples at nodes instead of individual nodes) and was the motivation behind the development of theoretically more powerful architectures like [483–485]. However, sadly on practical datasets, these architectures still under-performs than the simpler ones [486]. This leaves open the door for further improving our understanding of expressivity and making connections between expressiveness and generalization. A recent work [487] moves beyond the WL-Test with still provably powerful graph neural networks. This was obtained by construction of local subgraphs in a graph[18]. This thesis doesn't use subgraphs to test for improvement in the cosmic-ray composition analysis. This was primarily due to computational and time constraints. A future study might be helpful in improving the quality of this work even further.

16 Since the field is still in its infancy, besides the standard plights, the general training of a GNN is generally slower as well inefficient (involves more trial-and-error for finding viable methods) than other DL architectures.

17 GNN-architectures still need an element of data-specific design decisions

18 For a detailed summary of the work read "Beyond Weisfeiler-Lehman: using substructures for provably expressive graph neural networks" and "Beyond Weisfeiler-Lehman: approximate isomorphisms and metric embeddings" by Michael Bronstein on Medium.com

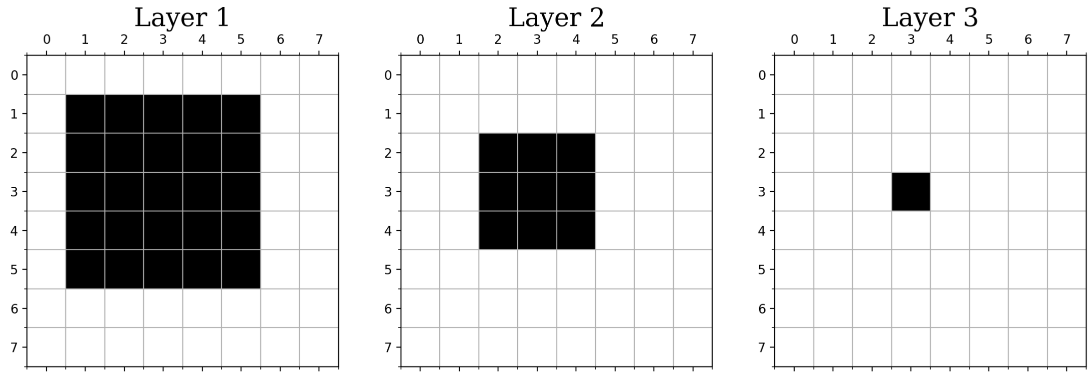


Figure 6.4: An illustration of receptive field among three CNN layers with a 3*3 kernel. The black pixel in layer-3 has a receptive field of the black area in layer-1. Similarly, each black pixel in layer-2 will have a 3*3 receptive area (in the black region) in layer-1.

6.3.2.2 *Receptive Field*

As DL methods become more ubiquitous it becomes increasingly important to introduce interpretability into the trained DL-models. For CNNs, researchers have tried feature visualization [488, 489], model dissection [490, 491], generalization capabilities [492, 493] and by putting theoretical guarantees [494]. Another approach to introducing interpretability is by studying the effect of parts of inputs on the produced output/embedding. This is generally termed as receptive field (or field of view). Figure 6.4 depicts the receptive field among three-layers in a CNN with kernel-size 3*3. The figure depicts the receptive field of a pixel in layer-3 by black-colored boxes. The pixel in the third layer captured information (during convolution/training) from a 3*3 part in layer-2, where each pixel also had information from a 3*3 region in layer-1. Hence the black pixel in layer-3 has accumulated information from the black area in layer-1/input and will be affected by variations in it. On increase in the network depth, the receptive field grows wider. It should be beneficial to increase the receptive field of each pixel in a hidden layer, in order to capture long-range dependencies. This can be obtained by either increasing the number of layers or kernel size. The receptive field for CNNs grows polynomially [495]. However, an increase in layers or kernel size increases the computational complexity as well as pushes hardware-capabilities[19]. Similar to CNNs, we can also define receptive-field for message passing GNNs. In simple-message passing GNNs, the receptive field is built by the nodes which contribute to the updated node representation. An example of receptive-field growth for a CR-simulation event mapped as a graph is depicted in Figure 6.5. The node-neighborhood is defined by kNN-neighbors with k=56, i. e. each DOM(/node) is connected to 56 nearest DOMs. After one iteration/layer a randomly selected DOM in the event will have information from its 56 immediate neighbors. After the second iteration, the DOM will have information from DOMs located one-hop away and so on, leading to an increase in the receptive field with an increase in the number of layers[20]. Another way to increase the receptive field is by increasing the number of neighbors for each node (equivalent to increasing kernel size

19 Dilated-convolutions [496] are one possible solution to prevent this.
20 Drawing receptive field for real-world networks, like the one shown in Figure 8.1, is much more difficult. However, a future study in this direction can enlighten us with new details in EAS-physics.

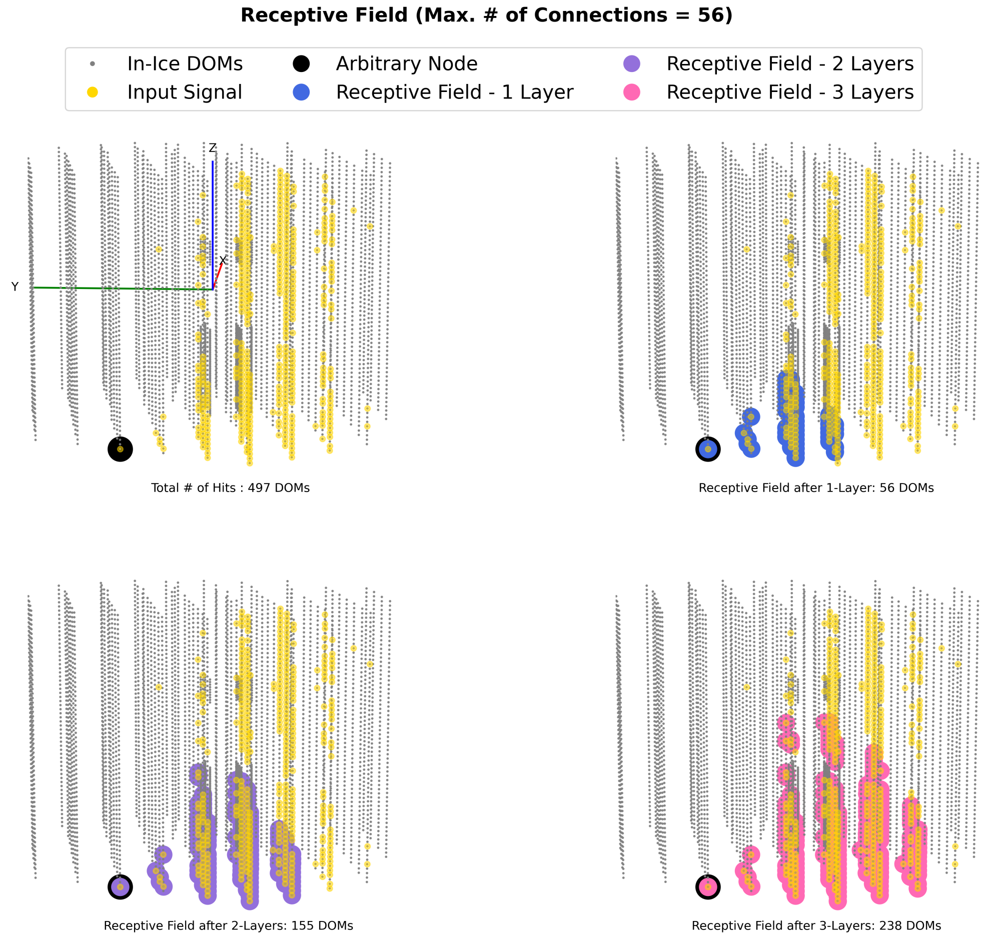


Figure 6.5: An illustration of receptive field for a simulation-event at IC mapped as a graph, for a simple message passing GNN. The neighborhood is defined by kNN-neighbors with k=56, i. e. each DOM(/node) is connected to 56 nearest DOMs. With an increase in the number of layers an arbitrarily selected DOM (marked-black) in the event aggregates information from DOMs located further away from it.

in CNNs). The receptive field for GNNs grows exponentially [497]. Similar to CNNs an effort to increase the receptive field also increases the computational complexity and pushes hardware capabilities. In addition to these, in an effort to increase the receptive field, the prediction performance of a GNN is marred by two more issues, namely over-smoothing (Section 6.3.2.3) and over-squashing (Section 6.3.2.4). Moreover, in a GMS lacking predefined edges, any artificially introduced edge definition can significantly affect the receptive field, leading to substantial performance losses and the introduction of structural biases.

6.3.2.3 *Over-smoothing*

Over-smoothing is a phenomenon seen in deep GNNs [453, 498], where the individual node-embeddings become increasingly similar to each other, leading to loss of any local-level differences in the input-graph. Over-smoothing can be linked to the message-passing mechanism. The aggregate step in message-passing is effectively an

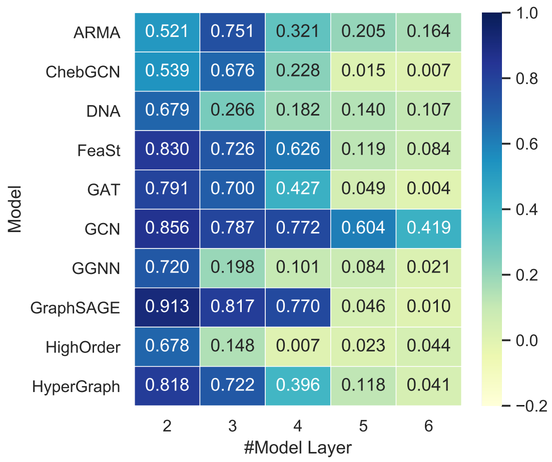


Figure 6.6: Measure of over-smoothing for multiple GNN-methods for *CiteSeer*-dataset [500] (a citation network), using mean average distance (between node embeddings) as the measure. A smaller number indicates that node embeddings look similar to each other. As can be seen, adding more layers increases smoothing. Plot from [499].

averaging/smoothing operation. With an increase in the number of GNN-layers the effective smoothing happens multiple times, washing out differences between neighboring nodes. This can be detrimental[21] in GMSs, where the information is embedded at local as well as global scales. This effect pushes the field of GNNs to have fewer layers[22]. Figure 6.6 (from [499]) depicts a measure of smoothing by using mean average distance (between node embeddings), on *CiteSeer* dataset [500]. A lower mean average distance means that the node attributes are more similar to each other. As can be seen from the figure, with an increase in the number of layers, the nodes start looking similar to each other. [499] also reports performance drops because of this. Graph-level tasks also depend on hidden node-embeddings, and hence for graph-classification tasks too we expect a performance drop on increasing the number of GNN-layers[23].

Chapter 7 will discuss, that from our previous knowledge, we know that there are small as well as large-scale details in the GMS for ICNO (detail in Section 6.5) which are sensitive to CR-composition. Hence, for this work, any loss at low level can be detrimental to performance. Extensive tests for this work favored three layers in the final architecture (shown in Figure 8.1). Skip-connections are possible solutions to get increased performance without increasing the number of layers [505]. Section 6.4 and Section 8.1.7 will detail how such ideas are adapted for this work.

---

21 It also creates heightened sensitivity to the selection of hyper-parameters.

22 This is in total contrast to CNNs, where increasing the number of layers generally helps improve the performance of the model.

23 To prevent over-smoothing a regularization term can be added to the loss during training. However, this is generally time-consuming and scales quadratically with the number of nodes. [501–504] present few viable regularization techniques to train deep GNNs. [505, 506] present viable solutions, based on architecture choices. However, the performance loss with increased depth still remains significant [503]. [507] presents how we can achieve enhanced performance, with increased depth, for GMS for geometric point-cloud data. This can be adapted in a future improvement of this work.

6.3.2.4 *Over-Squashing*

Section 6.3.2.3 introduced why despite the expectation that deeper-GNNs[24] (because of the increased receptive field), in reality, we cannot have very-deep GNNs. As we increase the number of layers and/or neighbors (another method to increase receptive field) GNNs face the problem of over-squashing (or bottleneck) of node-information. This is primarily because as we increase the receptive field each node collects information from exponentially growing [508] neighbors, into a vector of fixed size (node-attributes). This pushes the network to only learn local information [508]. Decreasing the number of layers (as already required by discussions in Section 6.3.2.3) is one possible solution to allay the problem. Decreasing the total number of node neighbors is another-possible solution. Figure 6.5 depicts that even simple message-passing algorithms tend to collect signal from almost half of the footprint for CR simulations at ICNO after three layers. In order to reduce the possible performance reduction because of surplus connections, in the final architecture for this work the maximum number of DOM/node connections is limited to 21 (number of node-connections in the final work is given by Equation 8.1)[25].

6.4 BRIDGING IDEAS WITH CNNS

In the following text, we will discuss two prominent works in the field of CNNs that have helped improve model-performance over a variety of datasets. DenseNet [509] and InceptionNet [510], have allowed to train ML-models efficiently by minimizing the vanishing-gradient problem[26] and learn structures in the input at different scales respectively. These ideas will be adapted to GNNs for our use case. The subsections will give a brief overview into the two architectures.

6.4.1 *DenseNet*

The advent of efficient computing systems has allowed researchers to improve upon the simple initial success achieved with CNNs for digit recognition, about 35 years ago [513, 514]. The success can be attributed mainly to the usage of deep architectures (more hidden layers). However, with an increase in depth of CNNs, we have to overcome the problem of vanishing gradients. With an increase in the number of layers, the gradients (partial-derivative of error with respect to the learnable weights), necessary to update the learnable weights during back-propagation [440], tend to approach zero. This makes the training of the network harder. Several works [**he2016deep**, 515–517] proposed to solve the issue, depend on creating shortcut connections between the early and later layers (or a variation of it). DenseNet, introduced in [509], is among the most prominent of such works. In order to ensure maximal information flow between the layers, it proposed connecting all preceding layers in a network with all the upcoming layers in a network. This is shown in Figure 6.7. The features among

24 if computational-complexity and hardware-constraints are ignored.
25 More neighbors per node beyond 21 were also tested (within computational limits).
26 NNs train by changing the values of learnable-weights during back-propagation. With an increase in the number of layers, the gradients (partial-derivative of error with respect to the learnable weights) tend to vanish. This effectively prevents the weights from changing and hence prevents the model from learning.

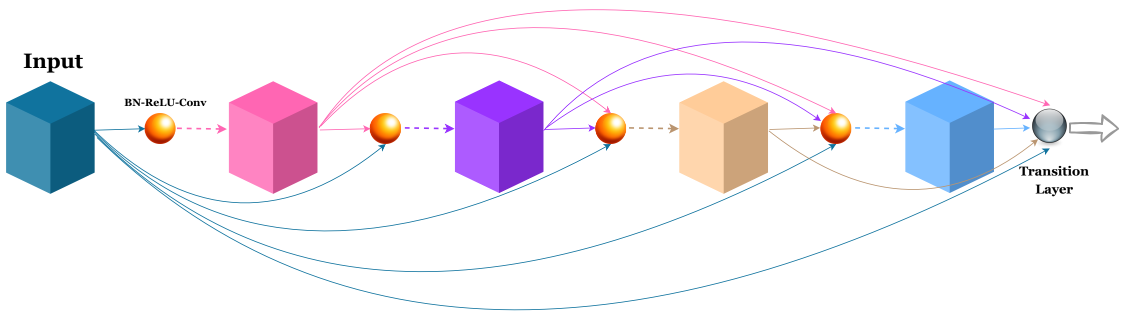


Figure 6.7: An illustration of a 5-layer block in a Densely Connected Convolutional Network. Each layer is connected to every other layer in a forward-fashion way. BN, ReLU, and Conv stand for Batch normalization [511], rectified linear units [512], and Convolution respectively.

different layers are concatenated (instead of averaged) to minimize information loss. During back-propagation, the multiple skip connections help prevent vanishing gradients. The idea will be adapted for GNNs to estimate CR-composition using IT and IC in Section 8.1.7. This helps improve the expressiveness of the final model, since the skip-connections allow the network to behave like an ensemble of shallower networks of variable length [518]. Similar ideas have already been introduced for GNNs [505, 519]. However, this work will present a unique customized version of the DenseNet idea.

6.4.2 *InceptionNet*

Introduced during ILSVRC (ImageNet Large Scale Visual Recognition Challenge)-2014, InceptionNet / GoogLeNet[27] allows learning small as well as large-scale structures in the input (with far fewer learnable parameters than its counterparts). Figure 6.8 presents one module of the InceptionNet. The final model published in the mentioned publication uses multiple such modules, stacked on top of each other. As can be seen in the figure, the module processes the same input in parallel across multiple channels. These are then concatenated, and serve as input for the next layer. The parallel processing of the same input data with variable kernel sizes allows the model to learn information at multiple scales together, while preventing the significant depth increase that would be needed otherwise to obtain similar accuracy. In addition, this allows the network to be wider as well as deeper at the same time, and hence benefit from learning unique representations in both [521]. The idea of processing data in parallel, with different kernel-size, is adapted for the use case of CR-composition estimate and will be presented in Section 8.1.7.

6.5 GRAPH NEURAL NETWORKS AT ICECUBE OBSERVATORY

As mentioned earlier in Section 3.2, IceCube-Gen2 is the planned upgrade to the existing IceCube Observatory at the South Pole. The introduction of the new components, an extended in-ice optical array (see Figure 3.14), a low-energy core (see Figure 3.14),

27 The idea is a brainchild of work done in [520].

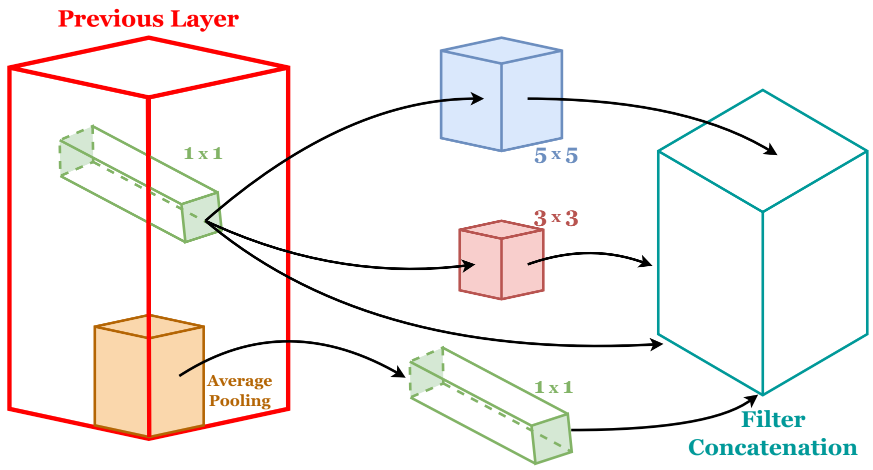


Figure 6.8: An illustration of a module used in InceptionNet [510], with dimension reduction. In the final InceptionNet-architecture, several of such modules are stacked on top of each other. The concatenated output from one layer served as the input for the next layer. Average Pooling can also be replaced by Max pooling. Read Section 6.4.2 for details. The kernel sizes reported here are the same as the original work [510].

a surface air shower array (see Figure 3.15), and an extended radio detector array (see [205]); will result in a modification of the detector geometry, leading to a more irregular configuration. Moreover, our DOMs are also becoming more complex (see Figure 3.13). These extensions are a sharp shift from the current array and DOM-types (see Figure 3.3). Because of the almost-hexagonal geometry and single-PMT DOMs, Convolutional Neural Network (CNN)s are the go-to method for most kind of reconstruction methods[28] at ICNO[29]. However, as we move towards the next generation of the instrument working with CNNs will be much more difficult. This is primarily because of the following reasons:

- The shift to an irregular geometry will make a shift to an orthogonal geometry non-trivial. Similarly a shift to multi-PMT DOMs will require a significant change in the CNN architecture[30].

28 In addition to the likelihood-based methods. Some of these have been discussed in Chapter 5.

29 CNNs need an orthogonal pixelated input. The hexagonal detector array can easily be mapped into an orthogonal geometry [198], with single-PMT DOMs serving as the pixels of the input. See Figure C.1, for the transformation from hexagonal to orthogonal geometry at IceCube.

30 The surrounding pixels (PMTs/DOMs) of a certain pixel (PMT/DOM) build up the neighborhood of the pixel in CNN. During the learning process of any CNN-based architecture, the neighboring pixels generally have an equal importance. This is okay for the current CNN-based architectures in IceCube since the inter-string separation (and inter-PMT / inter-DOM distances) are almost fixed among different combinations of neighbors. However, the separation between different combinations of string (or PMT/DOM) neighbors will generally not be fixed in IceCube-Gen2. Also, IceCube-Gen2 will use multi-PMT DOMs, as shown in Figure 3.13. For IceCube-Gen2, if individual PMTs are mapped as pixels of the CNN, then the neighborhood of PMTs in a single-DOM will be on equal weight-age as the PMTs in the next DOM/string. This is contrary to the real spatial separation and might be degrading to the model performance. To overcome the neighborhood-issue, if DOMs are interpreted as pixels of CNN, then there will be significant information loss (since we will need to aggregate information, from multiple PMTs in a DOM, into a DOM-level information).

- Because of the different detector-geometry IT [522] and DC [198, 523, 524] are generally treated separately in most CNN-based architectures at IceCube (see Figure C.1). This might not be the most optimal usage of the detector array.

- CNN-based architectures generally rely on aggregated pixel-level (DOM-level) information. However, even in IceCube, we have the capability to measure full pulses at the individual pixels (DOMs). Hence, there is a significant information loss. Benefiting from such information might be crucial when we perform low-energy analysis with a focus on using the planned low-energy core or plan on improving the current CNN-based work done at IceCube.

- CNN need a fixed-input. Most of the triggered events in IceCube don't lighten up the full detector, hence we have a very sparse input. This will be even more exaggerated as we move to IceCube-Gen2. CNNs generally have trouble working with sparse data [525]. Furthermore, the presence of additional zeros in the input (empty DOMs), results in a significant computational overhead for real-time applications[31].

In the simplest case, any MC-simulation or real-data at ICNO can also be mapped as a graph. This can be done by interpreting the active-DOMs (can also include inactive DOMs) in an event as nodes of the graph, and the edges can be constructed by using the locality of the DOMs. As can be expected from the above discussions, GNNs are capable of allaying all the above issues with CNNs at ICNO and IceCube-Gen2. GNNs don't need a fixed (shape and size) geometry of input-data and can hence easily integrate IT, DC, and IC together for the use case of ICNO, by mapping them as a point-cloud (of variable size). GNNs also have the capability of working with aggregated as well as pulse-level information. Only using the triggered stations/DOMs help overcome the sparsity problem and reduces computational overhead. GNNs have already shown significant promise in early implementations at ICNO [282, 526, 527]. Chapter 8 will discuss the use case of GNNs for CR-composition estimate at ICNO in more detail. The architecture uses the footprint of coincident air showers measured at IT, DC, and IC. In addition to the footprint, it also uses multiple air-shower observables (the composition-sensitive observables are discussed in Chapter 7) and is trained as a regression model.

31 There are two schools of thought on this. Some might say that the absence of information (DOM-hits) is also information. Hence, we shouldn't worry about the computational overhead for improved performance. The other school of thought is that the choice of DOM-location was almost arbitrarily chosen by us. Hence, the information captured by the triggered DOMs should in principle be enough. I personally belong to the second school of thought. The recent success of GNN-based architectures at ICNO [282, 526], where information from only triggered DOM is used, also seems to favor the latter.

7 CHARACTERIZING COMPOSITION: USING TEV MUONS

...Tho' much is taken, much abides; and tho'
We are not now that strength which in old days
Moved earth and heaven, that which we are, we are;
One equal temper of heroic hearts,
Made weak by time and fate, but strong in will
To strive, to seek, to find, and not to yield.

— Ulysses: Alfred Lord Tennyson [528]

The number of muons produced in an EAS is vital to deduce CR-composition. The depth of shower-maximum (generally referred to as X_{max}) is another EAS observable with similar separation capabilities. X_{max} also has smaller systematic uncertainties, than N_{μ} for mass estimation [103]. At ICNO, we currently lack the capabilities to directly measure X_{max}[1]. However, at IT we can measure the GeV muon-component in EAS[2]. Similarly, IC signal-footprint gives an estimate of the TeV muon content in EAS[3]. This work will focus on using the TeV muon content in EAS to design multiple CR-composition parameters, which probe different directions in EAS-physics. The choice to focus on TeV muons was also made because we already know that the data-MC inconsistencies for TeV muons are much less severe than those for GeV muons [97] (generally referred to as the *Muon Puzzle*). Moreover, it is evident from Elbert-formula (Equation 2.2) and Figure 2.6 that TeV muon multiplicity is a good proxy for primary-type. Hence, it is sensible to look for EAS observables that depend on the TeV muon-multiplicity[4]. The in-ice array of ICNO allows an ideal test case to perform such studies since primarily HE-TeV muons are capable to penetrate approximately 1500 m of Antarctic ice. The following text will probe multiple EAS observables which depend on TeV muon-multiplicity, point of first-interaction of the CR-primary in the atmosphere, and particle-multiplicity in EAS. In addition to proving crucial for CR-composition esti-

1 Future extensions of ICNO will consist of radio-antennas (discussed in Section 3.2). This will provide capabilities to directly measure X_{max} at ICNO. Efforts [529, 530] have already been made to estimate X_{max} at ICNO.

2 The EM-component of EAS is responsible for the majority of the charge deposit at IT. The contribution from GeV muons starts dominating (the contribution from EM-component) at large distances from shower axis [181].

3 Ongoing work [9] is trying to study TeV muon multiplicity at IC in order to constrain hadronic-models. A collective study of this work with [181] can be used to check inconsistencies and constrain hadronic models in the future [390].

4 ICNO currents lack the capability to directly measure muon-number. Other shower observables are hence used as a proxy for it.

mate, this also gives us the possibility to probe the internal consistencies in hadronic interaction models to describe the multiple EAS observables.

7.1 DIFFERENTIAL ENERGY LOSS : $\frac{dE}{dX_{1500m}}$

The HE-TeV muons traversing the in-ice medium lose energy via the process of continuous ionization losses along with large stochastic losses (as can be expected from Figure 2.7). A fraction of the associated Cherenkov light can be collected by one or multiple DOMs. The amount of energy-loss should be proportional to the muon's energy and muon multiplicity (as expected from Equation 2.3). In order to obtain the energy losses a segmented reconstruction along a track is performed using ICoII specific toolkit termed as MILLIPEDE. The details of the segmented reconstruction are presented in Section B.1. The reconstructed longitudinal distribution of IC energy-losses are then fit using the following function (details in [410]):

$$E_{\mu-\text{bundle}}(X) = \left(\frac{\kappa A}{\cos\theta}\right) \cdot e^{-bX} \cdot \gamma_\mu \cdot \left(\frac{E_0}{A}\right)^{\gamma_\mu - 1} \cdot \left[\left(-\frac{E_0}{A}\right)^{-\gamma_\mu}\left(\frac{a}{\gamma_\mu} - \frac{b}{1-\gamma_\mu}\cdot\frac{E_0}{A}\right) + E_{min}^{-\gamma_\mu}\left(\frac{a}{\gamma_\mu} - \frac{b}{1-\gamma_\mu}\cdot E_{min}\right)\right] \tag{7.1}$$

where $E_{min} = a/b(e^{bX} - 1)$, with a = 0.23881 GeV/m, b = $3.2852{\cdot}10^{-4}$/m represent the ionization energy loss and the radiative energy loss constant from the average energy loss for muons. $\kappa = 14.5$ GeV and $\gamma_\mu = 1.757$. Because of degneracy in E_0 and A, for energy-loss behavior the fit uses a constant value of A = 16 (i.e. oxygen)[5]. The MILLIPEDE-unfolded energy deposits (obtained using Equation B.1) is shown by the solid-line in Figure 7.1. The fit, using Equation 7.1, is shown by the dashed line. It was reported in [110] that maximal composition separation is obtained for fit values at slant-depth equivalent to approximately the top of IC. A slant-depth of 1500 m, which is approximately the top of in-ice array for most of the CR events, was chosen to obtain a composition sensitive parameter. Henceforth, it will be referred using $\frac{dE}{dX_{1500m}}/\frac{GeV}{m}$ or $\text{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$. Since, in general Fe-initiated CR-EAS have a greater muon-multiplicity than p-initiated EASs (as can be seen in Figure 2.5), we should expect composition-sensitivity from $\text{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$. Figure 7.2 shows that for EASs, with muon-multiplicity > 10, the parameter is also a very good proxy to estimate the in-ice muon-number.

The mass sensitivity of $\text{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ as a function of $\text{Log}_{10}(S_{125}/\text{VEM})$ [6] is shown in Figure 7.3. 10% of 2012's real-data (burnsample), spread throughout the year, is also used to validate and make comparisons. Few important physics inputs which go into Figure 7.3 are:

- ***Top Panel***: SIBYLL 2.1 [358] hadronic-model is chosen to compare the expectations from simulations with those of burnsample. This is primarily because currently, it provides the most compatible results among different measurements at

5 Hence, only E_0 is the free-parameter. Because of fixing the value of $A = 16$, E_0 can't be interpreted as the primary energy.

6 As shown in Figure 5.6, $\text{Log}_{10}(S_{125}/\text{VEM})$ is a very good energy-proxy. Hence, it can be used to make comparisons between MC simulations and burnsample.

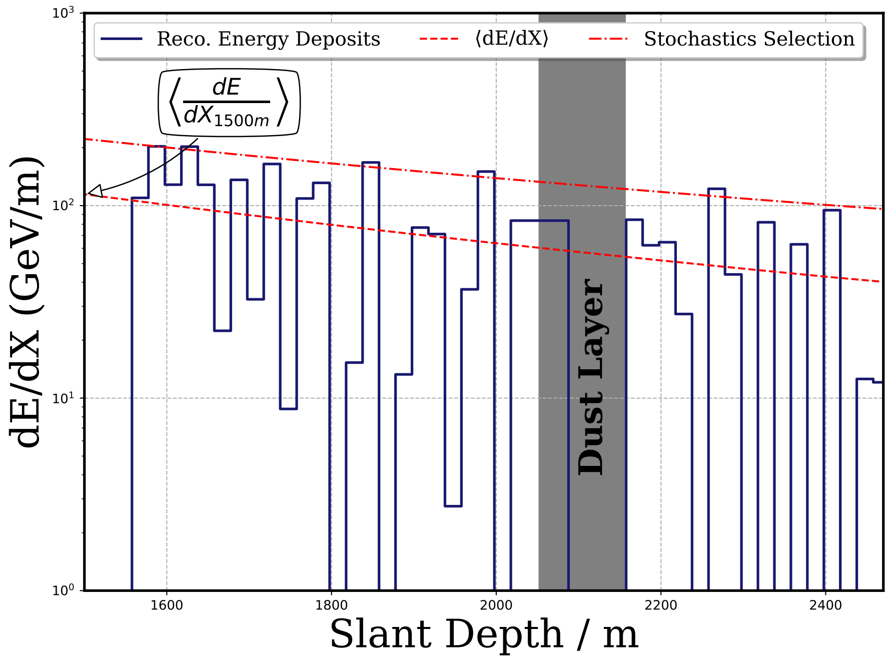


Figure 7.1: Reconstructed energy-deposits for MC-simulation of a proton with $\text{Log}_{10}(E_{True}/\text{GeV}) = 7.83$, as a function of slant depth. It is obtained by MILLIPEDE. The average energy deposit (obtained by Equation 7.1) is depicted by the dashed line. Dashed-dotted line denotes the selection performed to select high-energy stochastic deposits (details in Section 7.2.)

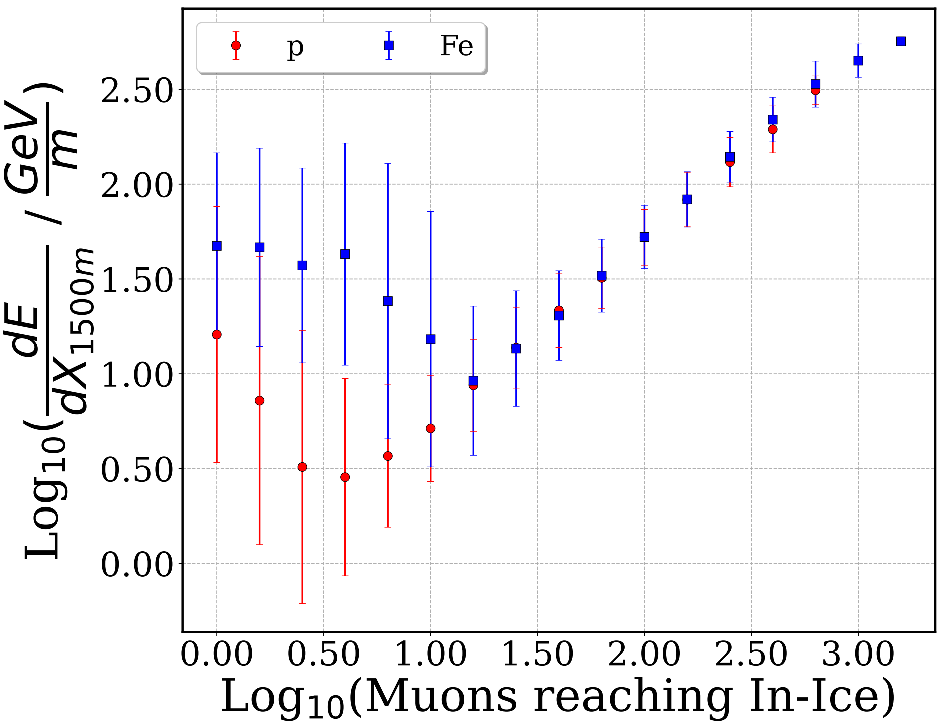


Figure 7.2: Average energy loss at a slant-depth of 1500 m as a function of in-ice muon-multiplicity. The underlying events are IceTop and IceCube (IT-IC) coincident events with $5 < \text{Log}_{10}(E_{True}/\text{GeV}) \leqslant 8$ and passing the IceTop and IceCube (IT-IC) quality cuts mentioned in Section 5.3.3.

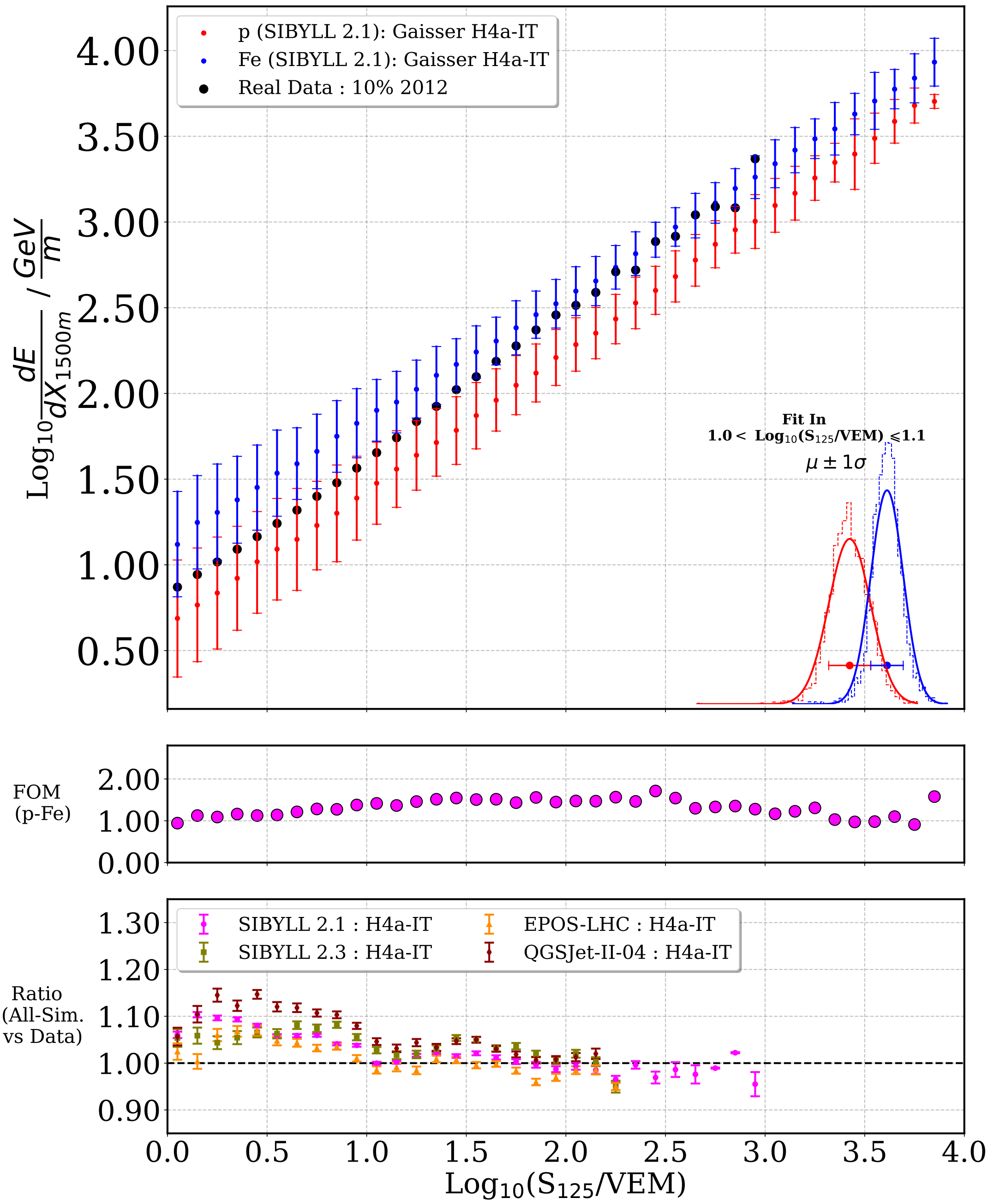


Figure 7.3: ***Top Panel:*** Composition sensitivity of $\mathrm{Log}_{10}\left(\frac{\mathrm{dE}}{\mathrm{dX}_{1500\mathrm{m}}}/\frac{\mathrm{GeV}}{\mathrm{m}}\right)$ as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ for SIBYLL 2.1 (weighted to H4a-IT). Inset Plot shows the fit in a bin. ***Middle Panel:*** FOM for p-Fe separation. ***Bottom Panel:*** Data-MC overlap for different hadronic models and H4a-IT flux-model.
Read Section 7.1, for more details of underlying physics inputs that go into different panels and for physics interpretations. The underlying events are IT-IC coincident events and pass the IT-IC quality cuts mentioned in Section 5.3.3.

ICNO [9, 181, 410]. The simulations have been weighted using a variation of the 5-component[7] H4a-flux model [90][8]. Currently, in ICoII we only have simulations for 4-components[9]. The weights for O are obtained by summing the weights for the Nitrogen and Aluminum groups. This flux served as the apriori estimate for the IT-only flux [58]. Hereon it will be referred to as Variation of Gaisser's H4a flux using only four components (H4a-IT). The inset plots[10] also uses the same hadronic model and flux-model.

- ***Middle Panel***: FOM stands for Figure of Merit (Definition in Section A.3). It is a measure of separation between two one-dimensional distributions. The higher the value, the easier it is to perform composition discrimination between primaries. To prevent overcrowding, FOM for only p-Fe, with SIBYLL 2.1 as the hadronic model and H4a-IT as the flux model is shown.

- ***Bottom Panel***: Four hadronic-models (weighted with H4a-IT) were used to compare simulation overlap with burnsample. This is done by calculating the ratio between simulation and data[11]. For SIBYLL 2.1, all 4-components (p, He, O and Fe) are used. However for other models, because of a lack of simulations (as shown in Table 5.2), only p and Fe have been used. The distributions only go till $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM}) \approx 3$ for SIBYLL 2.1 because of absence of events with higher $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$(or equivalently energy) in the burnsample. For other hadronic models it goes till $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM}) \approx 2.3$ because of lack of HE MC-simulations. A comparison is done owing to the existing scope for refinement in our understanding of the true hadronic model.

Important observations from Figure 7.3 are:

1. ***Top Panel***: p and Fe show a good-separation in $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{\mathrm{GeV}}{\mathrm{m}}\right)$ for the full energy-range[12]. The separation improves slightly with an increase in energy. The plot also shows mean-expectation from burnsample, using black-circles[13]. The data is well contained between expectations from simulations. This is crucial for it to be a useful parameter. The data also shows a transition from p-like composition to a more Fe-like composition with an increase in energy.

2. ***Middle Panel***: For composition-analysis, the bigger the FOM, the better it is. As can be seen from the middle panel of Figure 7.3 (and can be expected from the top panel), the FOM increases from about 1 to slightly less than 2 in the middle and then a decrease[14].

7 p, He, CNO, Mg-Si and Mn-Fe.
8 In this model, the extra-galactic component is assumed to be all protons.
9 p, He, O, Fe
10 For visualization purposes, the fit in the inset plot does oversampling between the fit range. The oversampling is not-used in the main plot.
11 The mean and standard deviation from the Gaussian-fit to the distribution is used, except for Section 7.2.
12 He and O are not shown to reduce overcrowding. The underlying events are also weighted to H4a-IT spectrum.
13 Error bars not shown to prevent overcrowding.
14 A larger overlap is expected at lower-energies because of greater shower-to-shower fluctuations [531]. The overlap at high energies emerges because of reconstruction issues involved in the reconstruction of $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{\mathrm{GeV}}{\mathrm{m}}\right)$ (discussed in [410]).

3. ***Bottom Panel***: Simulation-vs-Data comparison for any composition-sensitive parameter ensures that expectations from the MC-simulations are compatible with the observation of burnsample. Here, all 4 hadronic models for which simulations were present are shown (p+He+O+Fe for SIBYLL 2.1 and p+Fe for rest). We observe that for $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM}) > 1$, we see a good overlap i.e. Ratio = 1 for all hadronic models. However, below this we see a discrepancy for all hadronic models. The discrepancy is of the order of 5% - 15%.

As already shown in Figure 7.2, $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ is a very good proxy for HE TeV muon multiplicity, which is known to be a vital estimator for CR-composition. Hence, the separation capabilities of $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ can primarily be attributed to the muon-content of an EAS. $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ shows a transition from lighter to heavier-composition with an increase in energy (top panel in Figure 7.3). The trend is consistent with the expectation by different flux models [90–92, 532, 533] and measurements from other experiments [534–537] (in the energy-range relevant here). The bottom-panel of Figure 7.3 depicts the data-MC overlap of $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$, using H4a-IT as the flux-model. The parameter is generally consistent for $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM}) > 1$, for the choice of all hadronic models. However, at lower $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ a discrepancy of the order of 5 - 15% is observed, for all hadronic-models. The discrepancy can emerge because of three reasons. The first possibility is that H4a-IT might not be the true flux model. Previous work at ICNO [9] for TeV muon-multiplicity has already indicated that the data for the choice of SIBYLL 2.1 as the hadronic-model is more compatible with other flux models (like GSF [92], GST-3 [91]) than H4a[15]. [9] also indicated that for the choice of H4a as the flux-model EPOS-LHC is more compatible with real data than SIBYLL 2.1 and QGSJet-II-04. This can also be seen here, where the ratio (in bottom panel of Figure 7.3) is closer to 1 for EPOS-LHC generally (than other hadronic-models). Chapter 9 will discuss this possibility in more detail. Another possibility might be that none of the hadronic-models capture TeV muon-multiplicity correctly in the corresponding energy range. This might have connections to the known muon-puzzle (discrepancy between the number of simulated and observed muon content in EAS) for GeV muons. Chapter 9 will also discuss the discrepancy in more detail. Finally another plausible reason might be that our knowledge of ICNO detector might not be perfect. The previous [20] CR-composition analysis at ICNO had large systematic uncertainties. The systematic uncertainties are not treated in this work and are planned to be included in a future publication. Hence, it remains to be seen if the discrepancy is within the detector-systematics.

7.2 TOTAL STOCHASTIC ENERGY

As discussed in Section 2.2.1, bremsstrahlung energy-losses dominate the energy-loss landscape for HE TeV-muons. It is hence imperative to look for EAS observables that help us understand the stochastic losses from such deposits and hence hopefully obtain a composition-sensitive parameter. As mentioned earlier, at a fixed energy we expect higher muon-multiplicity for Fe-initiated EASs. This leads to the expectation of a greater overall deposit of Cherenkov light in IC by Fe-initiated EASs than p-induced

15 The work used H3a [90]. However, for low-energies, the behavior is very similar.

EASs of the same primary energy. However, the muons in a p-initiated EAS are more likely to be extremely high-energy muons[16]. Hence multiple studies were performed to get a stochastic-counting-based composition-sensitive parameter [20, 410, 538]. These studies effectively shifter the reconstructed energy fit (obtained using Equation 7.1) of the MILLIPEDE reconstructed energy deposits. This is given by:

$$\frac{dE_{\mu-\mathrm{bundle}}}{dX}(X_i) > a \cdot \left(\frac{dE_{\mu-\mathrm{bundle}}}{dX}(X_i)\right)^b_{\mathrm{reco}} \tag{7.2}$$

where a = 5 and b = 0.8 [17]. This is denoted using the "Stochastics Selection" curve in Figure 7.1. In order to obtain an uncorrelated composition-sensitive observable, the previous studies relied on counting the number of bins that passed the "stochastics selection" cut. However, the observable provided very minimal composition sensitivity [20, 410]. Owing to it, this work makes a departure from bin counting. Instead, the total energy deposited in the bins where energy deposit passes the "stochastic selection" is evaluated. This is termed as Total Stochastic Energy (TSE). Since TSE tries to use high-energy stochastic bremsstrahlung-deposits and since Fe-initiated EAS are muon-richer we should see more such deposits from Fe than p.

The mass sensitivity of TSE as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ is shown in Figure 7.4. The burnsample, spread throughout the year, is also used to validate and make comparisons. The physics inputs which go into the plots are the same as the one detailed for Figure 7.3 and are mentioned in Section 7.1. Important things to notice about Figure 7.4 are:

1. ***Top Panel***: Similar to the previous parameter, p and Fe show a good separation for the full energy range. The separation improves significantly with an increase in energy. This might be an accumulated effect of greater shower-to-shower fluctuations, as well as the bias of the reconstruction algorithm towards ionization losses (more prominent at lower energies), both of which are prominent at low-energies[18]. Similar to the previous parameter, the data is well contained between expectations from simulations and shows a transition from p-like composition to a more Fe-like composition with an increase in energy.

2. ***Middle Panel***: Although slightly worse than the FOM for $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{\mathrm{GeV}}{m}\right)$, there is a significant improvement in FOM for TSE over the bin-counting method mentioned earlier, and discussed in [20, 410]. It is however important to notice that the reduced FOM is also because of a significant contribution of events where TSE is zero. This can be seen in the inset plot of the Top panel in Figure 7.4. Because of a pileup, the width of the distribution increases, causing a drop in the FOM-value.

3. ***Bottom Panel***: We generally see a good overlap (between burnsample and MC-simulations) for all-hadronic models.

16 Simply put: There are less muons in a p-initiated EAS to get the fraction of energy which goes into the muon component of an air-shower. Hence, each muon on average can receive more energy (in comparison to if it was a Fe-initiated EAS at the same energy).

17 This was termed as the *standard-selection* in the studies mentioned earlier. A *strong selection* with a = 7 and b = 0.9 was also used for those studies.

18 discussed in [410]

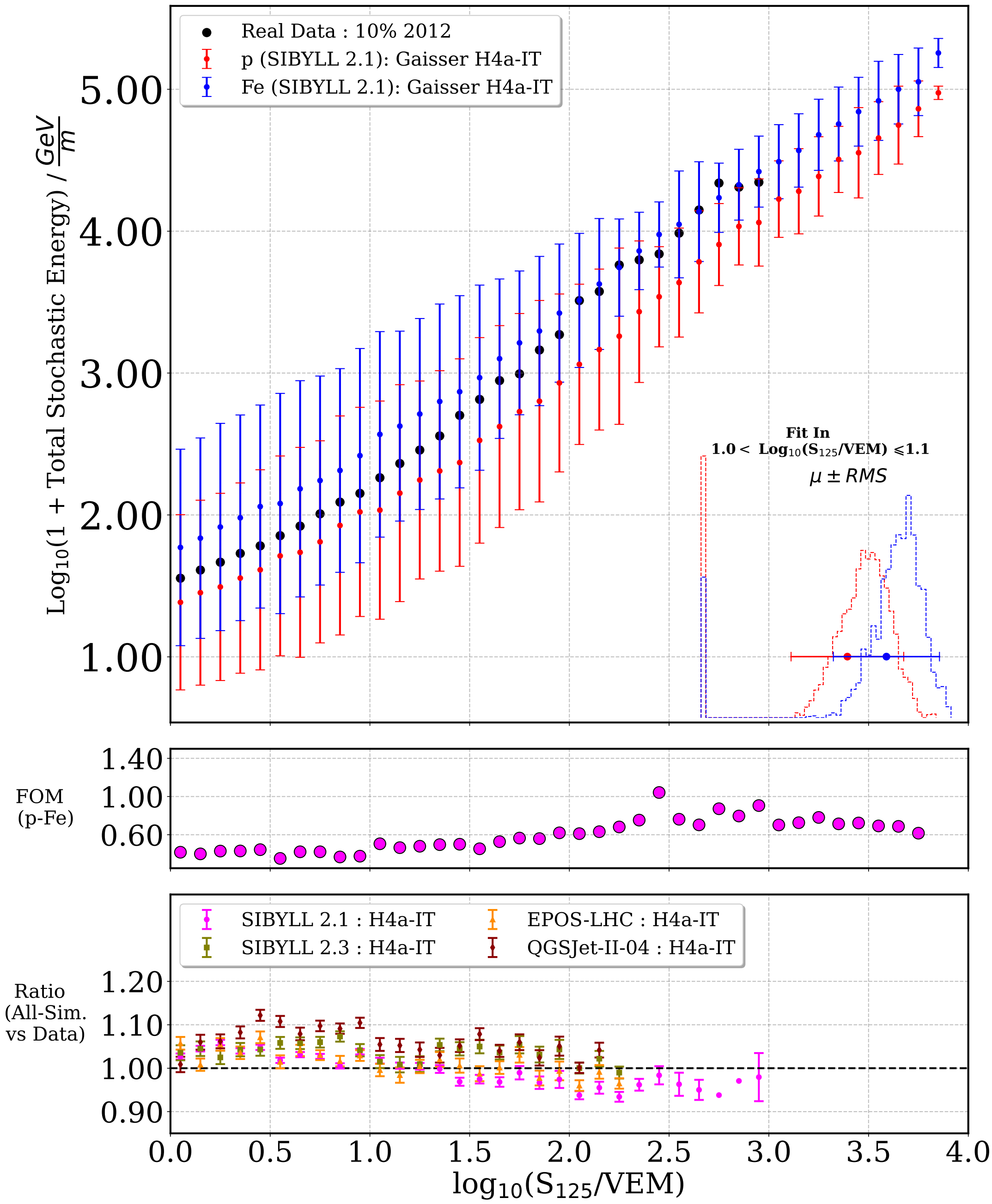


Figure 7.4: ***Top Panel:*** Composition sensitivity of Total Stochastic Energy (TSE) as a function of $\text{Log}_{10}(S_{125}/\text{VEM})$ for SIBYLL 2.1 (weighted to H4a-IT). Inset Plot shows the histogram in a bin and the corresponding RMS estimates. ***Middle Panel:*** FOM for p-Fe separation. ***Bottom Panel:*** Data-MC overlap for different hadronic models and H4a-IT flux-model.
Read Section 7.1, for more details of underlying physics inputs that go into different panels and Section 7.2 for physics interpretations. The underlying events are IT-IC coincident events and pass the IT-IC quality cuts mentioned in Section 5.3.3.

For fixed energy, Fe-initiated EAS have a higher muon-multiplicity than p-initiated EAS. Hence, the corresponding TeV muons by Fe-showers have a greater chance of having local stochastic deposits in IC[19]. However, p-initiated showers can have more extreme deposits because of the energy-budget concentrated among fewer muons[20] (than a Fe shower of the same primary energy). Even though there is a degeneracy predicted for muon-multiplicity (by Elbert formula - given by Equation 2.2), between primary-type and muon-energy (and hence deposit in IC); this test showed the potential of predicting CR-composition by focussing on local stochastic deposits by TeV muons. In the inset-plot of Figure 7.4 (top-panel), we see an underflow bin for both p and Fe. These are events that passed the IT-IC quality cuts mentioned in Section 5.3.3, but had no slant-bin which pass the energy-cut established by Equation 7.2. Hence, these events have total stochastic energy of zero[21]. This might be responsible for reduced composition sensitivity (shown by middle-panel of Figure 7.4) of this feature, in comparison to $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ (shown by Figure 7.3). A fine-tuning of the cut established by Equation 7.2 can hence potentially help improve the composition sensitivity of this parameter in the future. Similar to $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$, with an increase in energy, TSE also shows a transition from lighter to heavier composition (top-panel in Figure 7.4). The data-MC discrepancy (bottom-panel in Figure 7.4) can be explained by the same reasons as mentioned for $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ (towards the end of Section 7.1).

7.3 RATIO PARAMETER

KASCADE-Grande[22] has demonstrated that the ratio of muon-number to total charged-particle number is a good composition-sensitive parameter [534]. Since, IC and IT measures energy deposits instead of their densities, a direct counting of particle-number for different shower components is currently not feasible at ICNO. Reconstructed proxy parameters are employed to test for the usability of particle-type ratios for mass discrimination. It has already been shown in Figure 7.2 that $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ is a very good-proxy for muon-number (especially for muon-multiplicity > 10). $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$, which is an energy-proxy, is used as a total particle-number proxy. It is very evident that no new information is being added here. This parameter is just a test case for the potential of combining analysis like [9, 181] with this analysis, in the future. The Ratio Parameter is given by:

$$\text{Ratio Parameter} = \frac{\mathrm{Log}_{10}(dE/dX_{1500m})}{\mathrm{Log}_{10}(E(S_{125m}))} \tag{7.3}$$

The mass sensitivity of Ratio Parameter as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ is shown in Figure 7.5. Similar to before, burnsample is also used to validate and make comparisons. The physics inputs which go into the plots are the same as the one detailed for Figure 7.3 and are mentioned in Section 7.1. Important things to notice about Figure 7.5 are:

19 These are mainly bremsstrahlung deposits, as can be expected from Figure 2.7.

20 Figure 2.7 shows that with energy loss for TeV muons almost shows a linear increase with muon-energy, for TeV energies. Hence, if muons are more energetic they can deposit more extreme energies.

21 These are still usable events. TSE value of zero is an indication that our cut might be too strong.

22 It was an EAS experiment located at the Karlsruhe Institute of Technology (average atmospheric depth $1022\ \mathrm{g/cm^2}$), Germany.

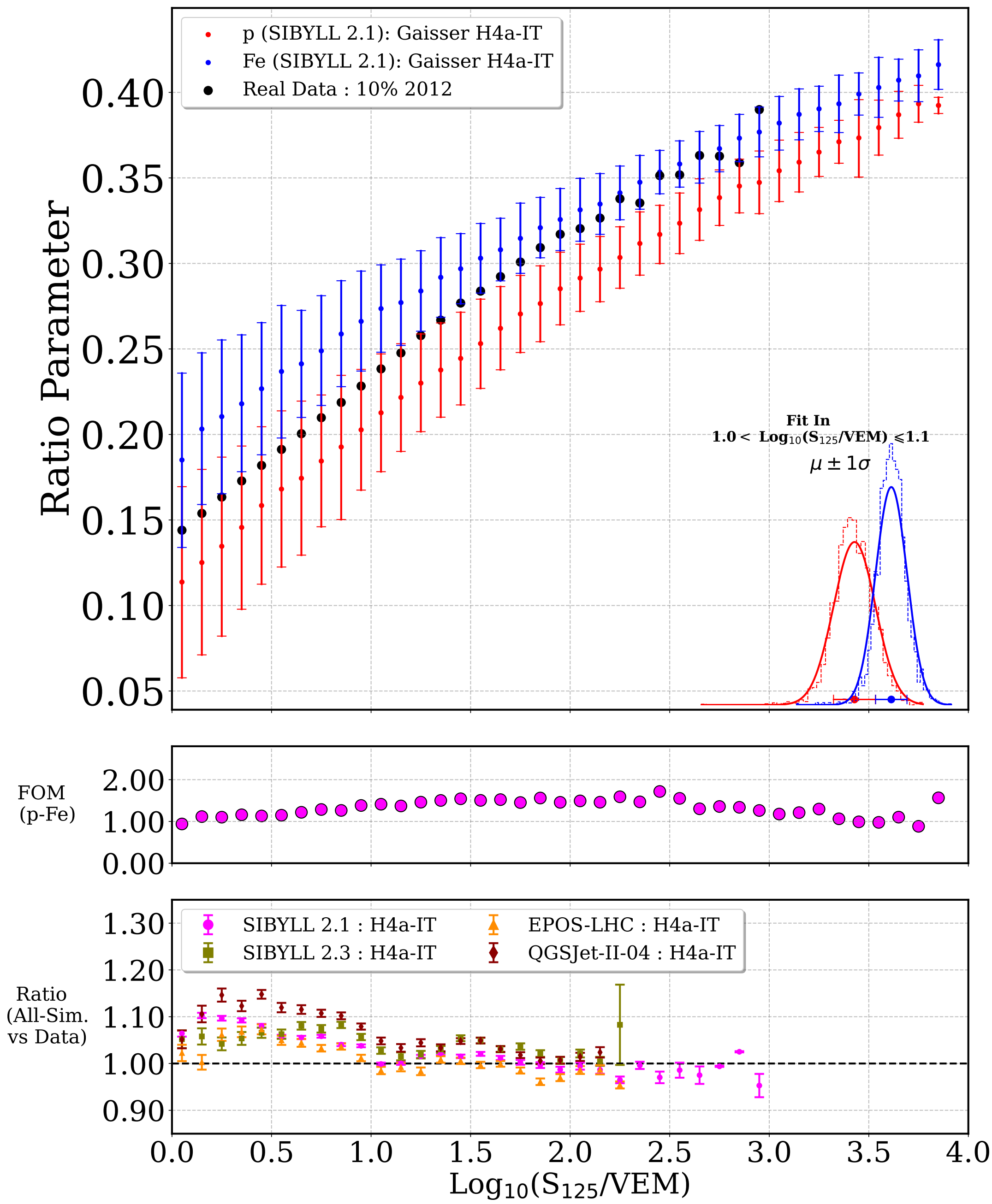


Figure 7.5: ***Top Panel:*** Composition sensitivity of **Ratio Parameter** (Section 7.3) as a function of $\text{Log}_{10}(S_{125}/\text{VEM})$ for SIBYLL 2.1 (weighted to H4a-IT). Inset Plot shows the fit in a bin. ***Middle Panel:*** FOM for p-Fe separation. ***Bottom Panel:*** Data-MC overlap for different hadronic models and H4a-IT flux-model.
Read Section 7.1, for more details of underlying physics inputs that goes into different panels and Section 7.3 for the physics interpretations. The underlying events are IT-IC coincident events and pass the IT-IC quality cuts mentioned in Section 5.3.3.

1. ***Top Panel***: It is expected from Equation 7.3 that we should see the composition-sensitivity very similar to the one we see from $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ (Figure 7.3). This is expected since $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ is designed to have minimal composition sensitivity (see Figure 5.6). Hence, the composition sensitivity is driven primarily by the numerator. As expected, the data is also well contained between expectations from simulations, and with an increase in energy shows a transition from p-like composition to a more Fe-like composition.

2. ***Middle Panel***: The FOM values are also comparable to those obtained earlier for Figure 7.3.

3. ***Bottom Panel***: As expected, the data-MC overlap also looks very similar to Figure 7.3. Similarly, a discrepancy below $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM}) < 1$ can be observed for all hadronic models.

As mentioned earlier, the parameter is a test case for combining (in the future) this work with muon-multiplicity studies at ICNO and is primarily driven by $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$. Hence, very similar to Section 7.1, with an increase in energy we see the transition from lighter to heavier composition. The FOM and data-MC discrepancy behavior also remain the same.

7.4 MEAN RADII AND MEAN CHARGE

The composition-sensitive parameters discussed earlier primarily depend only on the energy deposit by the TeV muons. The information on the lateral spread of TeV muons was generally disregarded in those. However, we expect to see the difference in the lateral spread of muons, between a p and Fe initiated EAS. Fe interacts earlier in the atmosphere than p. Moreover Fe-initiated EASs have a higher muon-multiplicity with lower average muon energy. These phenomena permit muons with larger transversal momenta to be positioned farther away from the shower axis for Fe-initiated EASs, in comparison to p initiated EASs. Hence, we should expect wider muon-bundles for Fe than p. ICNO is unique as it provides us the opportunity to test this hypothesis using an almost pure signal from TeV muons observed in IC. To test this the following steps were followed:

- **Step 1:** Select a pulse and track which will be used to calculate the distance of DOM from the track. Pulses cleaned around track given by IT and the track from IT was used[23].

- **Step 2:** Calculate the perpendicular distance of IC DOM from the track and read charge-deposits at the DOM[24].

- **Step 3:** Select the maximum allowed radii (perpendicular distance from the track) of a DOM to be considered for the analysis. Choosing a very small radius will lead to a very small passing fraction. Choosing large radii on the contrary can

23 **For IceCubers:** *Pulse Map* = CoincLaputopCleanedPulses; *Track* = From Laputop. Other pulsemap maps and tracks were also tested.

24 This is done using ICNO specific Data Derived Differential Deposition Reconstruction (DDDDR) module (details in Section B.2).

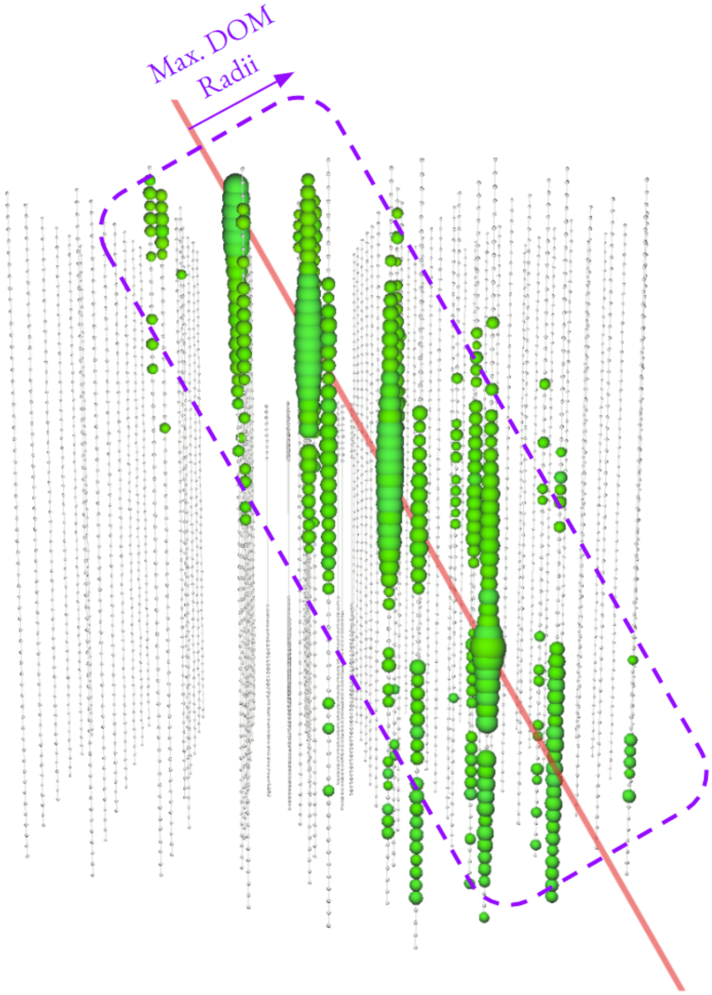


Figure 7.6: A representation of in-ice charge deposit. Here the purple dashed boundary depicts a cylinder of predefined radii, with its axis aligned along the track (denoted by a red line here). The DOMs within the cylinder are used to calculate composition-sensitive EAS-observables discussed in Section 7.4.

diminish the composition sensitivity of the observable[25]. This is depicted using Figure 7.6. The efficiency of selecting a particular DOM radius is given in Table 7.1.

It is clear from Table 7.1, that a maximum DOM radii of 100 m (and greater) doesn't remove any events[26]. In the following, composition-sensitive comparison (and burnsample-MC overlap) for the choice of multiple radii ($r_{Max.} \geqslant 200$) m will be shown. Multiple parameters were tested to look for composition-sensitive observables. As mentioned before, we should expect a difference in *Mean Radii* between p and Fe initiated EASs. Hence, the first parameter tested is *Mean Radii*.

25 This can happen by increasing dependence on late photons. There are known issues in the ability of MC-simulation to replicate late-photons.

26 For a smaller maximum-radii, the events are thrown out because of an absence of any DOM within the maximum-radii cut.

MAXIMUM DOM RADII	PROTON	IRON
20 m	0.5157	0.5137
50 m	0.9581	0.9568
80 m	0.99	0.99
$\geqslant$100 m	1.0	1.0

Table 7.1: Passing Fraction (p and Fe) for different choices of maximum DOM radii.

7.4.1 *Mean Radii*

Mean Radii is evaluated for every event and is defined as:

$$\text{Mean Radii} = \langle r_i \rangle_{r_i < r_{Max.}} \tag{7.4}$$

where r_i is the perpendicular distance of the i^{th} DOM from the track and the average is taken over all DOMs within a distance of $r_{Max.}$ meters. The following will show composition-sensitivity for multiple choices of $r_{Max.}$ (= 200 m, 400 m, and 600 m)[27]. Similar to before, burnsample is also used to validate and make comparisons. The physics inputs which go into the plots are the same as the one detailed for Figure 7.3 and are mentioned in Section 7.1. Important things to notice about Figure 7.7 are:

1. ***Top Panel***: With increase in $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ (or energy), the *Mean Radii* of the charge-deposit from TeV muon increases. This is an indication that muon-bundles become wider with an increase in energy, for both p and Fe. On average, in all energy-bins we see that *Mean Radii* is greater for Fe-initiated EAS, than for p. This is expected since Fe interacts earlier in the atmosphere and Fe-EASs have higher muon-multiplicity, leading to the expectation of muons with larger transverse momentum located far from shower-axis. It shows a good composition sensitivity over a large part of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ (till $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM}) \approx 2.5$ i. e. $\mathrm{Log}_{10}(E/\mathrm{GeV}) \approx 8.5$). As can be seen from Figure A.2 or Figure A.3, with an increase in the maximum DOM radii to 400 m/600 m (from 200 m), we see a good separation throughout the energy range[28]. Hence, the probable cause for lower composition sensitivity, at higher energies for the case of $r_{Max.} = 200$ m is because the composition sensitivity is mostly driven by deposits from muons located far off from the shower axis. For the final analysis (Chapter 8 and Chapter 9) only *Mean Radii* with a maximum-radii cut of 200 m is used[29]. As before, the plot also shows mean-expectation from burnsample, using black dots. For all the cases, it is generally well contained between expectations from simulations[30]. The data also shows a transition from p-like composition to a more Fe-like composition with an increase in energy.

27 Radii between 100 m and 200 m didn't provide any additional composition sensitivity, in comparison to 200 m.
28 In future, an energy-based radii-cut can be tested to look for potentially improved composition sensitivity.
29 A future work can hence include *Mean Radii* for different maximum-radii cuts to potentially improve this work and understand the data-MC overlap better. The variation in *Mean Radii* among the three cuts also has the potential to help us improve our knowledge of the systematics of ICNO.
30 Low statistics are probably the reasons for jumps in the last few bins in the burnsample.

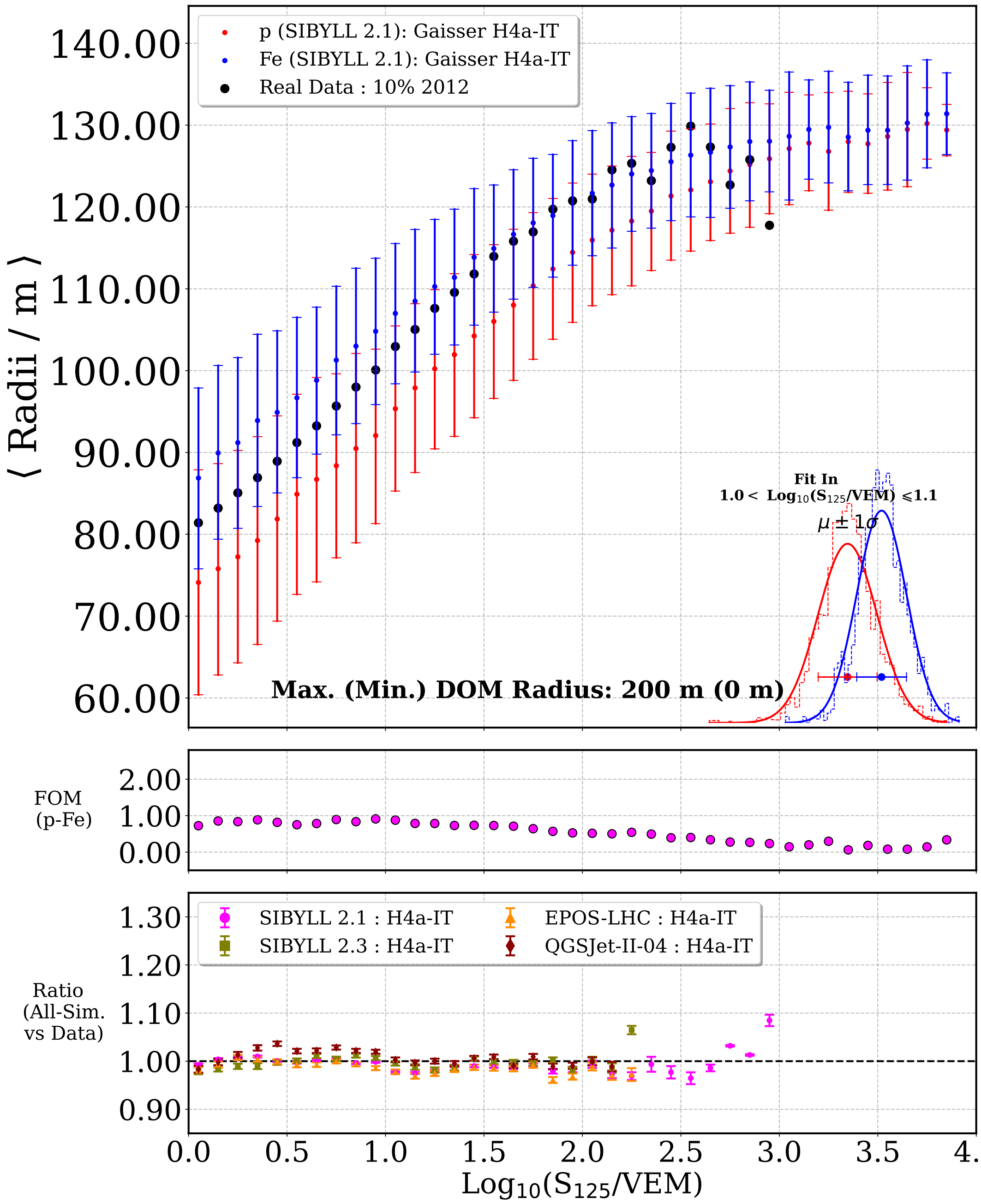


Figure 7.7: ***Top Panel:*** Composition sensitivity of *Mean Radii* (Section 7.4.1) as a function of $\text{Log}_{10}(S_{125}/\text{VEM})$ for SIBYLL 2.1 (weighted to H4a-IT), with maximum DOM distance cut of 200 m. Inset Plot shows the approximate fit in a bin. ***Middle Panel:*** FOM for p-Fe separation. ***Bottom Panel:*** Data-MC overlap for different hadronic models and H4a-IT as the flux-model.
Read Section 7.1, for more details of underlying physics inputs that go into different panels and Section 7.4.1 for the physics interpretations. The underlying events are IT-IC coincident events and pass the IT-IC quality cuts mentioned in Section 5.3.3.

2. ***Middle Panel***: As mentioned before, *Mean Radii* shows a good composition sensitivity till $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM}) \approx 2.5$. This can also be seen using the FOM plot. The FOM drops significantly after this. However, with an increase in $r_{Max.}$ (Figure A.2 and Figure A.3), we see a significantly improved FOM at higher energies.

3. ***Bottom Panel***: As mentioned before data-MC comparison for any composition-sensitive parameter ensures that expectations from the MC simulations are compatible with the observation of burnsample. As can be seen from Figure 7.7, *Mean Radii* (with maximum DOM distance of 200 m) has a very good Data-MC overlap over the entire range observed. With an increase in radii (Figure A.2 and Figure A.3), we see an almost constant bias throughout the full energy range (of the order 5-12%), for the choice of H4a-IT as the flux-model.

The *Mean Radii* parameter has shown promise in performing CR-composition analysis using a shower-observable which is not directly connected to energy-deposits by muons, but has dependence on the point of first-interaction (of the CR-primary) in the atmosphere[31]. This opens up the possibility of checking internal consistencies among hadronic models to explain multiple shower-observables and test their compatibility with real data. A quick-glance on the data-MC overlap of *Mean Radii* (bottom-panel in Figure 7.7) shows a significant improvement (particularly at lower-energies) than $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{\mathrm{GeV}}{m}\right)$ (bottom-panel of Figure 7.3). Section 7.5 and Chapter 9, will discuss this in more detail. This parameter can also be used to update low-energy spectrum studies at ICNO [166]. Similar to other shower-observables, with an increase in energy, the parameter also shows the transition from lighter to heavier composition. The effect of including more data and including detector-systematics remains to be seen.

7.4.2 *Mean Charge*

Mean Charge is evaluated for every event and is defined as:

$$\text{Mean Charge} = \langle \mathrm{Log}_{10}(c_i) \rangle_{r_i < r_{Max.}} \tag{7.5}$$

where c_i and r_i is the charge and the perpendicular distance of the i^{th} DOM from the track and the average is taken over all DOMs within a distance of $r_{Max.}$ meters. Similar to Section 7.4.1, the following will show composition-sensitivity for $r_{Max.}$ (= 200 m, 400 m and 600 m). The burnsample and the physics inputs which go into the plots remain the same as the one detailed for Figure 7.3 and are mentioned in Section 7.1. Figure 7.8 presents composition sensitivity of *Mean Charge* for the case when $r_{Max.} = 200$ m. Important things to notice about Figure 7.8 are:

1. ***Top Panel***: For p as well as Fe primary, the *Mean Charge* deposit increases with increase in $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ (or energy). This is expected since from Figure 2.5

31 The width should also have a dependency on the muon-multiplicity. However, it is expected to be not the leading factor. A detailed study of the contribution of first interaction and muon-multiplicity to the width of muon-bundle can provide new insights into EAS-physics. This was not pursued for this work since the CORSIKA-simulations used for this work were generated without the E-HISTORY option [539]. E-HISTORY allows tracking of the precursor particles (mother, grandmother particle, etc.) responsible for the observation of a particle at the detector level. However, record-keeping of history is memory-intensive. Future studies can probe the correlations with new limited datasets.

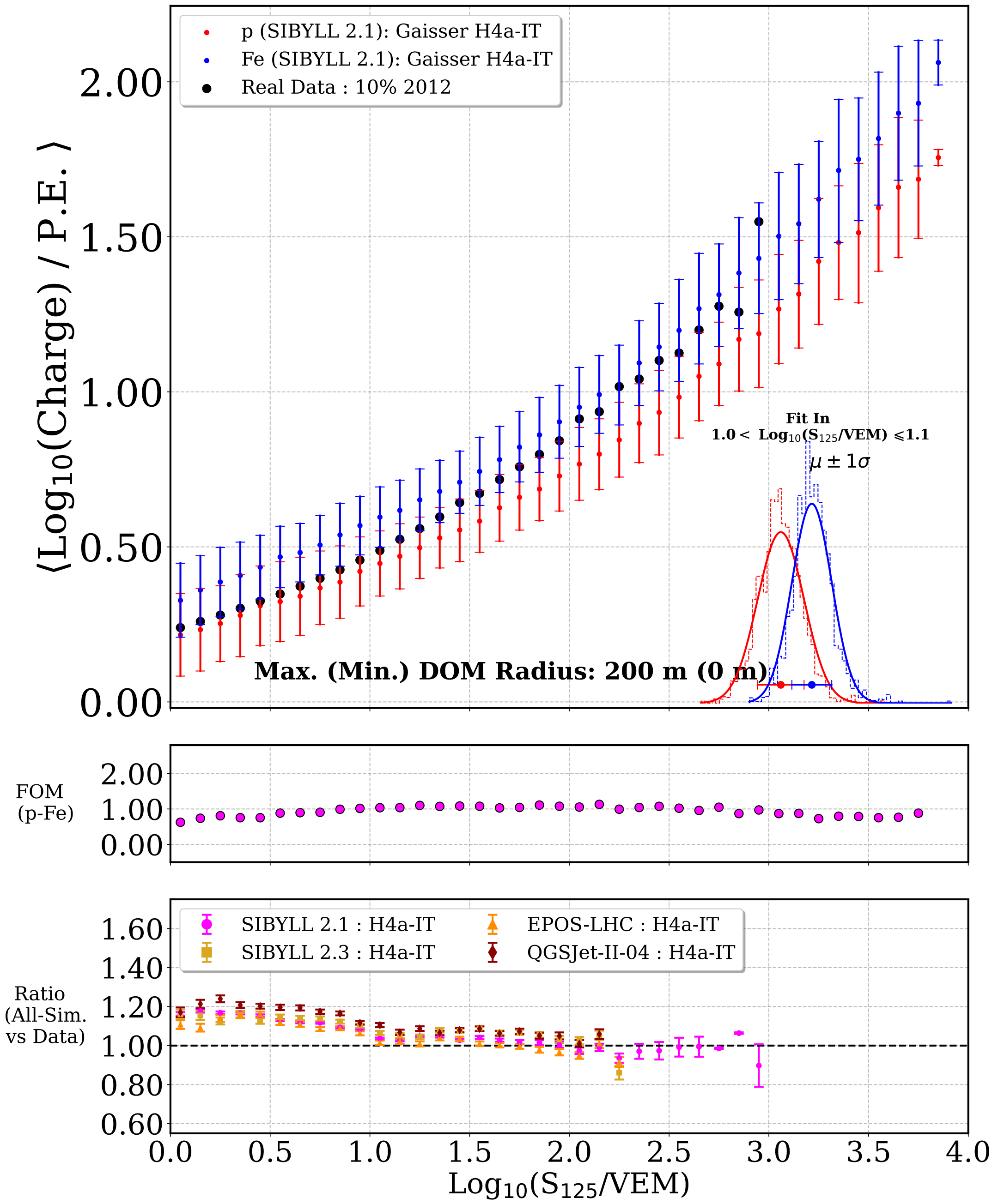


Figure 7.8: ***Top Panel:*** Composition sensitivity of *Mean Charge*(Section 7.4) as a function of $\text{Log}_{10}(S_{125}/\text{VEM})$ for SIBYLL 2.1 (weighted to H4a-IT), with maximum DOM distance cut of 200 m. Inset Plot shows the fit in a bin. ***Middle Panel:*** FOM for p-Fe separation. ***Bottom Panel:*** Data-MC overlap for different hadronic models and H4a-IT flux-model.
Read Section 7.1, for more details of underlying physics inputs that goes into different panels and Section 7.3 for the physics interpretations. The underlying events are IT-IC coincident events and pass the IT-IC quality cuts mentioned in Section 5.3.3.

we know that the muon-multiplicity increases with an increase in energy. Hence, we should also see a corresponding increase in TeV-muons. This is responsible for the trend in charge-deposit seen in IC. In all bins, a greater charge deposit is seen by Fe than for p. This is also expected, since Fe-showers generally have a higher muon-multiplicity than p-showers (Figure 2.5). *Mean Charge* also shows a good composition sensitivity over the full $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ range (full energy range). With an increase in maximum radii ($r_{Max.} = 400$ m $\Rightarrow$ Figure A.4 or $r_{Max.} = 600$ m $\Rightarrow$ Figure A.5), we don't see any visible change in separation, in comparison to the case of $r_{Max.} = 200$ m. The data also shows a transition from p-like composition to a more Fe-like composition with an increase in energy.

2. ***Middle Panel***: The FOM almost remains constant with energy as well as with a change in $r_{max.}$.

3. ***Bottom Panel***: For the case of $r_{Max.} = 200$ m, we see a good burnsample-MC overlap for $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM}) \gtrapprox 1$. At lower energies, we see a discrepancy of the order of 15-20%.

The *Mean Charge* parameter is primarily dependent on muon-multiplicity. The parameter has close connections with $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$. In contrast to depending on muon-energy reconstruction from MILLIPEDE[32], *Mean Charge* is simpler to evaluate and comparable in separation capabilities as $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$. A future improvement of this parameter can help bring the separation capabilities fully at-par with $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ and beyond. The data-MC seen below $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM}) \lessapprox 1$, can be explained by the three-possibilities as discussed for $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ (towards the end of Section 7.1). These are namely: a) H4a-IT might not be the true flux-model b) None of the hadronic-models capture the TeV muon-multiplicity well c) Our knowledge of ICNO-systematics might need improvement. It remains to be seen that testing with more data or including detector systematics will help resolve the discrepancy. There is also an overall increase in discrepancy (almost a constant shift) for ($r_{Max.} = 400$ m $\Rightarrow$Figure A.4 or $r_{Max.} = 600$ m $\Rightarrow$Figure A.5). Since the data is well contained between expectation from simulations (*Top Panel*) and we see a good FOM (*Middle Panel*), *Mean Charge* could still be a good composition parameter. The discrepancies are discussed further in Section 7.5 and Chapter 9.

7.5 COMPOSITION CONSISTENCY AMONG OBSERVABLES

The previous sections discussed multiple CR composition-sensitive observables. These probe multiple directions in air-shower physics, like energy-deposits by TeV muons (continuous and stochastic), lateral spread[33]. Since the true CR-composition for an event should be fixed, it is crucial that composition expectations from multiple composition-sensitive observables should be similar in an energy bin. Any discrepancy among

32 $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ is obtained from unfolding charge-deposits observed at DOMs to get energy-deposits along slant-length in IC and a further fit to those deposits (discussed in Section 7.1 and Section B.1).

33 Ratio Parameter (Section 7.3) is planned to be a particle-multiplicity based parameter in the future. For now, muon energy deposit and their lateral spread are the two main directions being explored in this work.

shower observables can emerge because of the one or an accumulated effect of multiple of the following reasons:

- **Hadronic Models**: Our knowledge of hadronic interaction models (discussed in Section 5.1.2) still has scope for improvement and their predictions are subject to large theoretical uncertainties [342, 343]. Figure 5.1 and Figure 5.2 present the difference in expectation of muon-multiplicity and lateral-distributions[34] among the hadronic models. As can be seen from Figure 5.1 there are significant differences in muon expectation among the hadronic models. Since the true-hadronic model is not known, it is crucial to consider the differences between them. For e.g. If SIBYLL 2.3c was the true hadronic model, then interpreting the data with SIBYLL 2.1 will wrongly interpret the muon excess as heavier composition (Fe-like). A similar statement can be made about shower width by looking at expectations from pseudorapidity distributions presented in Figure 5.2.

- **Flux Models**: Based on the choice of flux-model (Details in Chapter 2) the contribution and/or the interpretation of the shower-observables in MC and/or data can change. As already stated in the discussion towards the end of Section 7.1, for the choice of SIBYLL 2.1 as the hadronic model a prior study [9] at ICNO seems to favor flux models like GSF [92], GST-3 [91]) than H3a. In contrast to this other study at ICNO [20] seems to favor H3a and H4a as the most compatible model[35]. Hence, in addition to the hadronic model, extreme care needs to be taken when interpreting real-world data using a flux model. This thesis intends to predict cosmic-ray composition. Hence, this works indirectly tries to give another expectation of a viable flux model. Chapter 9 will detail the consistency of the flux-parametrization thus observed among multiple shower-observables and among hadronic models. The flux models effectively only changes the fractional contributions from the individual elements in an energy bin. Hence, for the time being, if the data is contained between simulations (p and Fe are the two extremes), it should be safe to proceed. The results (fractional contributions) from this work can be then tested to check the consistency of the result among shower-observables.

- **Detector Systematic:** An incorrect knowledge of detector can shift the expectation for an EAS-observable in MC-simulations. This can lead to wrong physics conclusions on real data. The standard to study such effects at ICNO is to simulate datasets with the conservative estimates of the detector systematics and then see the effect on the final results [20]. Because of time constraints, such studies are left for future publication.

In order to test compatibility among the multiple shower-observables discussed before, the comparison between $\frac{dE}{dX_{1500\ m}}(\mathrm{GeV/m})$-*Mean Radii*, and *Mean Charge*-*Mean Radii* will be performed[36]. Similar to before, it is crucial to discuss what goes into plots like Figure 7.9, before discussing specific cases. Some important physics inputs which go into such plots are:

34 The plot presents pseudorapidity distributions.
35 Since the work was marred with large systematic uncertainties, other models were also compatible within systematic limits.
36 The choice was made since these are the representative examples of shower-observables depending on muon-multiplicity and EASs lateral-extent.

1. The plot is divided into six blocks, based on $\text{Log}_{10}(S_{125}/\text{VEM})$-value. The blocks are $\text{Log}_{10}(S_{125}/\text{VEM}) = [0, 0.5)$, $[1.0, 1.5)$, $[1.0, 1.5)$, $[1.5, 2.0)$, $[2.0, 2.5)$ and $[2.5, 4)$. The choice of binning was done to prevent information loss due to excessive averaging and to avoid insufficient event or simulation representation within each bin.

2. In each block, the associations are:
   - **Color:** The color associations are, Red = proton, Blue = Iron, Purple = H4a-IT weighted expectation. For SIBYLL 2.1, all 4-components (p, He, O, and Fe) are used. For other hadronic models only p and Fe are used. Black color is used to denote burnsample.
   - **Line Style:** The line style shows association with choice of hadronic interaction model and is given by, —— = SIBYLL 2.1; – · – = EPOS-LHC; – – = SIBYLL 2.3c and · · · · = QGSJET-II-04.

3. The ellipses denote the confidence region for the parameters. For each combination of the parameters and hadronic model, the central (smaller) ellipse denotes the 1σ confidence interval. Whereas the outer one denotes the 2σ confidence interval. The simulations have been weighted using H4a-IT as the flux model. The color and line associations remain the same as described earlier.

4. The side-histograms are projections of the observable values along that axis. The MC-histograms have been weighted using H4a-IT as the flux model.

5. For *Mean Radii* and *Mean Charge*, the maximum (and minimum) DOM radii (details in Section 7.4) is mentioned on the top of each plot.

We should see an overlap between data and simulation ellipses with the same composition expectation from different shower observables - if a hadronic model replicates the interaction in the atmosphere well, if H4a-IT is the most compatible flux-model with data, and if there are no inconsistencies between hadronic interaction models. The following text will discuss if this is the case.

7.5.1 $\frac{dE}{dX_{1500m}}$ *vs Mean Radii*

Figure 7.3 and Figure 7.7 present the expectation for various primary types and the composition sensitivity of the earlier discussed $\text{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{\text{GeV}}{\text{m}}\right)$ and *Mean Radii* as a function of $\text{Log}_{10}(S_{125}/\text{VEM})$ respectively. Figure 7.9 borrows information from those figures and plots the confidence intervals for *Mean Radii* and $\text{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{\text{GeV}}{\text{m}}\right)$ in different $\text{Log}_{10}(S_{125}/\text{VEM})$-bins. Important things to notice about Figure 7.9 are:

1. As can be seen from the side-histograms in different blocks, as we move towards higher $\text{Log}_{10}(S_{125}/\text{VEM})$-bin or equivalently primary-energy, we have fewer simulations (colored histograms) as well as burnsample (black histogram). This is expected since the number of simulated EASs at higher energies are far fewer than at lower energies. Similarly, with increase in energy we expect to see a power-law drop in flux of data.

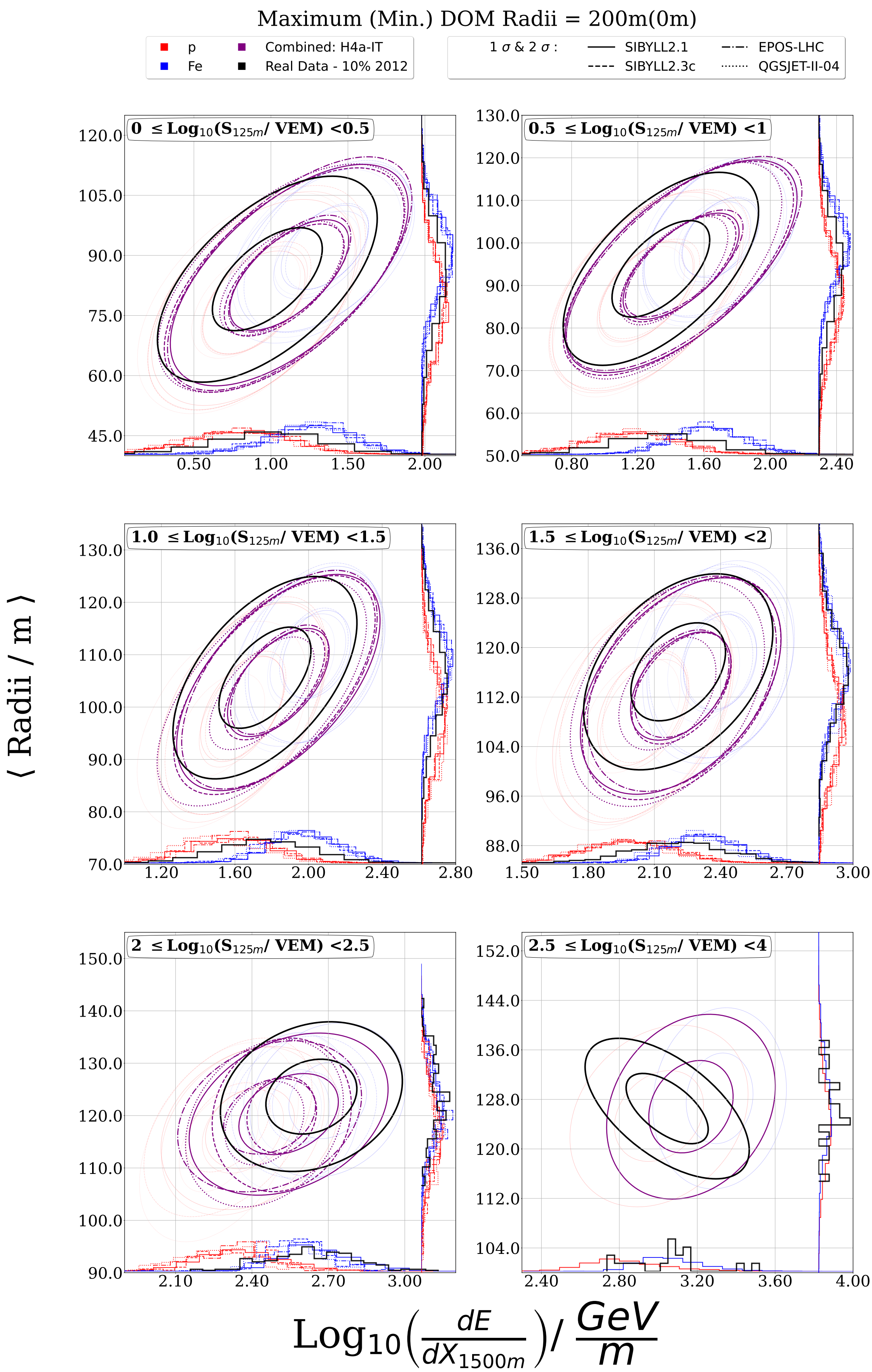


Figure 7.9: Confidence-region for $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ and *Mean Radii* and Data-MC overlap for different hadronic interaction models.
Read Section 7.5, for more details of underlying physics inputs that goes into different blocks and Section 7.5.1 for physics interpretations. The underlying events are IT-IC coincident events and pass the IT-IC quality cuts mentioned in Section 5.3.3.

2. The burnsample seems to well-bounded within p and Fe (as can also be expected from Figure 7.3 and Figure 7.7). However, the expectation from H4a-IT model shows a greater shift (from burnsample) in $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ than *Mean Radii*, for the choice of all hadronic-models. The discrepancy in expectation of the true-composition of the burnsample, among shower-observables, can also be seen from the side-histograms in almost all $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$-bins. As the discrepancy exists for all hadronic models, one possible reason for this discrepancy can be that all the hadronic models describe the lateral-spread of TeV muons better than their multiplicity[37]. Another possible reason for the discrepancy can be that H4a-IT is not the flux model most compatible with real data. Chapter 9 will quantify the discrepancy (and compatibility) of various shower observables for the choice of various flux-model and compare the compatibility of results from this work and H4a-IT, with burnsample. Another possible reason for the discrepancy, among shower-observables, might be that the detector-systematics knowledge has scope for improvement and the discrepancy among shower-observables might be covered (and resolved) under detector-systematics. As mentioned before the detector-systematics study for this work is planned for a future publication. Hence, it remains to be seen that if including detector systematics will help resolve the discrepancy among shower observables.

3. In general the expectation from different hadronic models match well with each other, with the greatest fluctuations seen for QGSJET-II-04. The slight differences can be expected from underlying differences in the models (see Figure 5.1 and Figure 5.2).

As a summary, the comparison for composition expectation between *Mean Radii* and $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ suggest that either all the hadronic-models describe TeV muon-lateral spread better than their multiplicity, and/or H4a-IT is not the flux-model most compatible with burnsample at ICNO. Chapter 9 will describe the compatibility with burnsample for the various choices of hadronic models and flux models. Meanwhile, the possibility of composition expectation, between shower observables, becoming compatible with the inclusion of detector-systematics remains.

7.5.2 *Mean Charge vs Mean Radii*

The consistency check among *Mean Radii* and *Mean Charge* is presented in Figure 7.10. The plot borrows the information shown in Figure 7.7 and Figure 7.8. Similar to the previous subsection, the confidence intervals (for *Mean Radii* as a function of *Mean Charge*) are shown in different $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$-bins (more info in Section 7.5). Important things to notice about Figure 7.10 are:

1. Similar to the earlier comparison, as we move higher in energy we see fewer simulations and events in burnsample.

2. Similar to before, the burnsample is well bounded between p and Fe (expected from Figure 7.7 and Figure 7.8). Similar to $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ we see a greater

37 Reminder: Mean Radii is a measure of lateral-spread of TeV muons. $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ is a proxy for TeV muons.

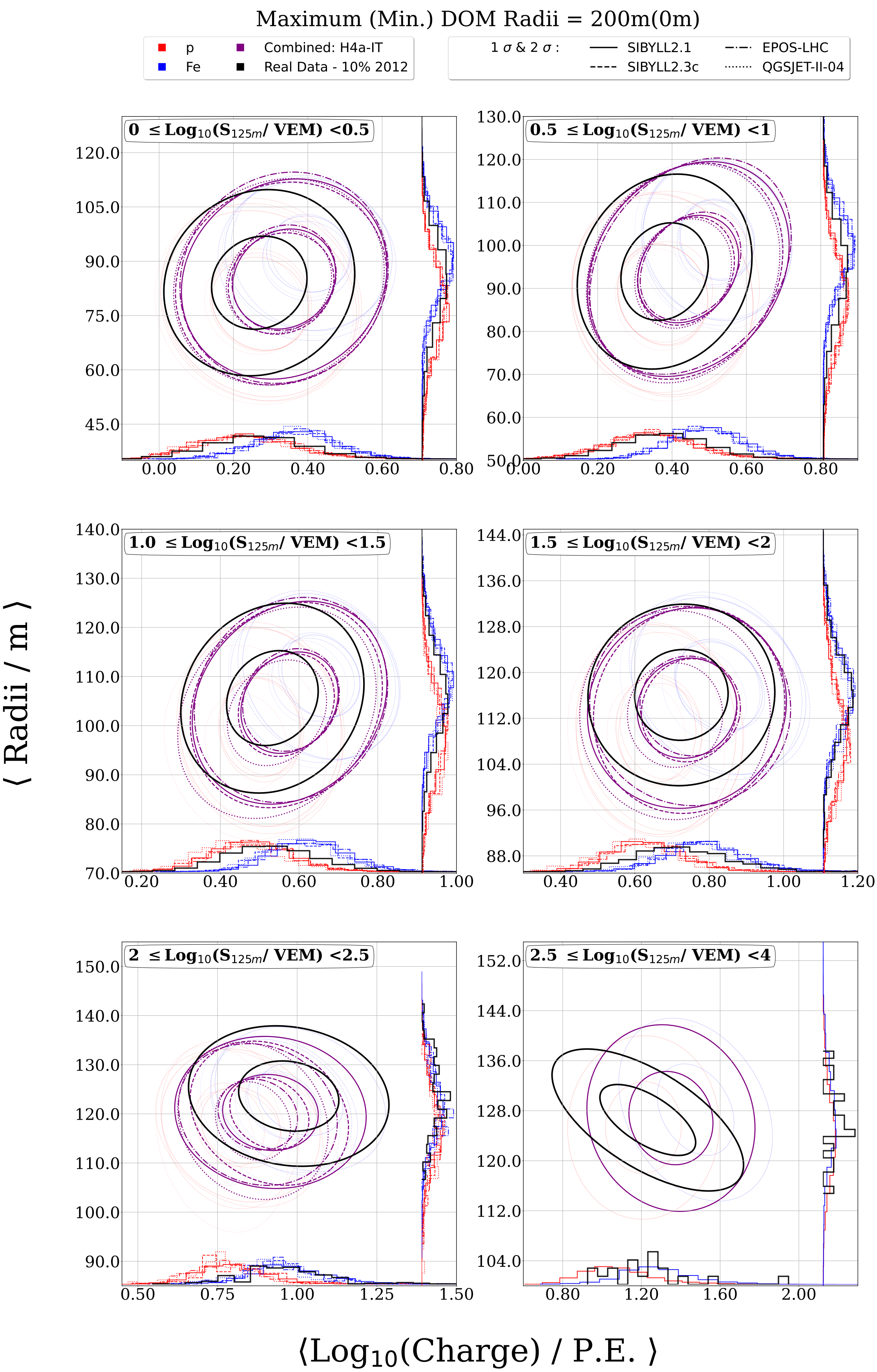


Figure 7.10: Confidence-region for *Mean Radii* as a function of Mean Charge and Data-MC overlap for different hadronic interaction models.
Read Section 7.5, for more details of underlying physics inputs that go into different blocks and Section 7.5.2 for physics interpretations. The underlying events are IT-IC coincident events and pass the IT-IC quality cuts mentioned in Section 5.3.3.

discrepancy with burnsample for *Mean Charge*, and hence the suggestion for a slightly different composition (from side-histograms) among the different shower-observables. This can be expected if all the hadronic models describe lateral-spread better than the multiplicity. Similar to before the possibility of H4a-IT not being the most-compatible flux model, and the affect of detector systematics on the predictions remains.

3. Similar to the earlier comparison, the expectation from different hadronic models generally match well with each other, with the greatest fluctuations seen for QGSJET-II-04. The fluctuations can be expected because of the underlying differences in hadronic models.

As a summary, the comparison for composition expectation between *Mean Radii* and *Mean Charge* also shows tensio in overlap seen earlier. This is expected, since *Mean Charge* and $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ are both muon-multiplicity based parameters. And hence discrepancy seen earlier in *Mean Radii*-$\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ is also visible here. As mentioned before, Chapter 9 will describe the compatibility with burnsample for the various choices of hadronic models and flux-models. The prospect of achieving compatibility in composition expectations among shower observables, while accounting for the incorporation of detector-systematics, persists.
This chapter summarized multiple CR-composition-sensitive observables for ICNO. The observables primarily depend on energy deposits by TeV muons. They try to capture muon energy deposits (continuous and stochastic), muon multiplicity, and lateral spread. It is crucial to benefit from combined information captured by these observables and the observed shower footprint, to get the maximal composition sensitivity. The following chapters will discuss the combined potential of using the shower-observable with shower footprint to get the maximal composition sensitivity, in addition to discussing discrepancies among various shower-observables and burnsample.

8 ANALYSIS

"make infinite use of finite means"

— Wilhelm von Humboldt [540] ; Popularised by Noam Chomsky [541]

This work aims to obtain fractional spectra for different primary types in cosmic-rays (CRs), using data measured at IceCube Neutrino Observatory (ICNO). The study of change in primary-type contribution with increasing energy has important astrophysical implications [542–550]. It can provide insights into the origin and acceleration mechanism, responsible for the acceleration of CRs to the highest energies. An abundance of one primary type over another can help understand the dynamics (acceleration-mechanism, age, etc.) of such sources. Moreover, since CRs also carry the imprints of the magnetic field they traverse, a study of CR-composition can also help gain insights into the magnetic fields (galactic and extra-galactic) and the associated transport properties of CRs. The study of CR-composition also provides a unique probe into the study of particle physics models (like hadronic interactions, nuclear fragmentation, and particle-production models), at energies that are generally not accessible at accelerators (and their associated detectors). The primary-type contribution in real data can be obtained by multiple methods. This chapter will introduce a method based on the estimation of the logarithmic mass of each EAS, using a Graph Neural Network (GNN). This is an extension of the analysis done in [20]. Here instead of using few composition-sensitive observables, the full-shower footprint is also used. In addition, newer composition-sensitive observables have also been developed (read Chapter 7 for details). Hence, this work benefits from the combined usage of shower footprint as well as composition-sensitive observables. The logarithmic-mass estimates (for each shower) can then be used to get fractional contributions for primary types, using a template fitting method. This will be discussed in Section 8.4.

8.1 MASS-ESTIMATE USING GRAPH NEURAL NETWORK

GNNs provide the flexibility to work with GMSs having irregularly-geometry[1]. Section 6.5 introduced the necessity of using GNNs at ICNO. CR-composition estimate provides a good test case to check the potential of using GNNs at ICNO, and potentially using it for IceCube-Gen2 in the future. In order to realise this the observed footprint (at IT and IC) is mapped as a graph, with tanks (IT) and DOMs (IC) serving as the nodes of the graph (details in Section 8.1.1). After edge-construction (details in Section 8.1.2), the event-mapped-as-graph can serve as input for any custom GNN architecture, to predict composition. Figure 8.1 shows the final-form of GNN-based architecture used in this work. The subsequent subsections will detail the reasoning behind the various

1 ICNO consists of multiple-detector components, IT (Section 3.1.5), IC and DC (Section 3.1.6) with varying geometries.

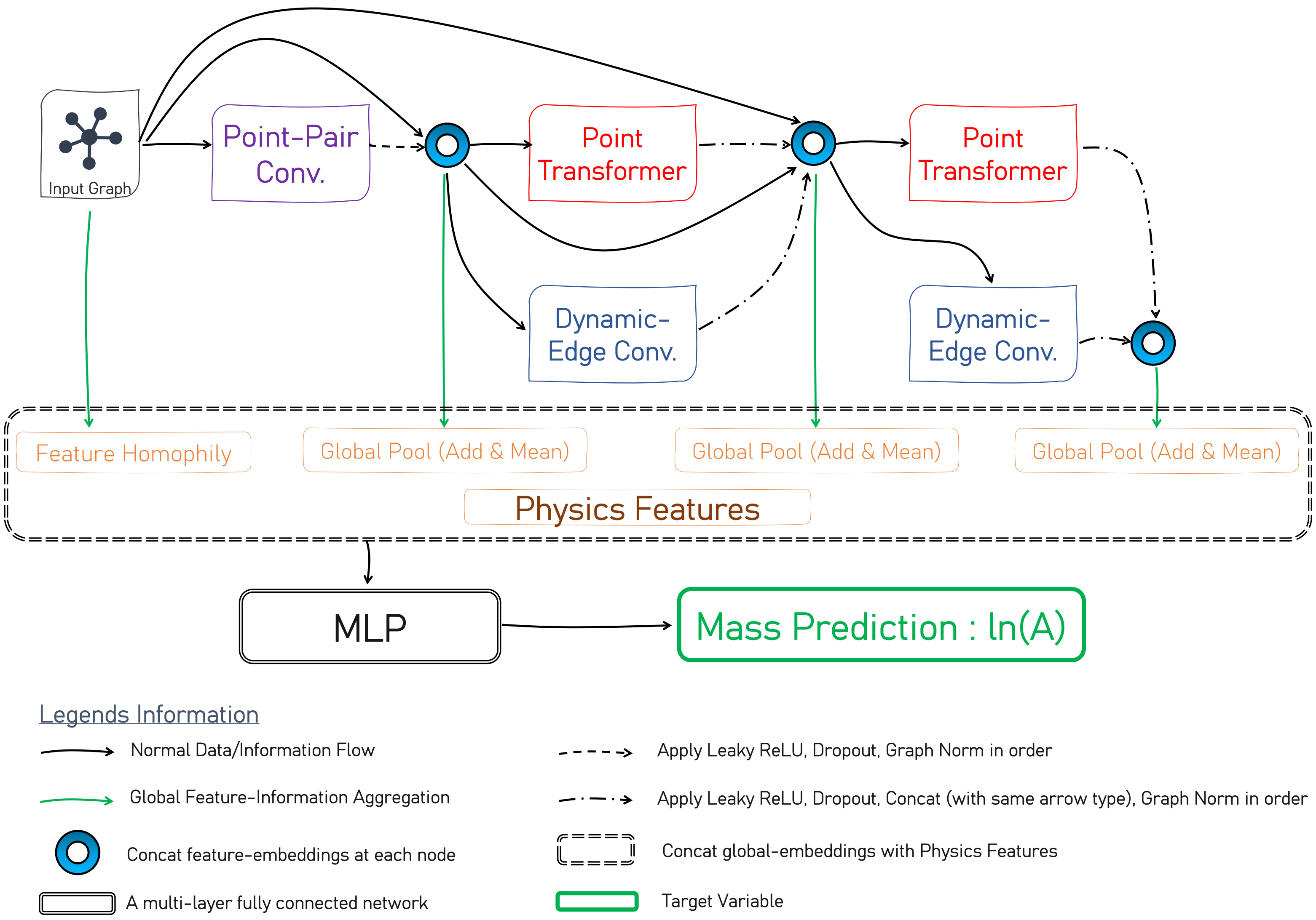


Figure 8.1: GNN-based architecture for CR-Composition estimate at ICNO. For details read Section 8.1.

choices made in the construction of input-graph (Section 8.1.1, Section 8.1.2), in the network architecture (Section 8.1.4 - Section 8.1.8) as well as other details (Section 8.1.3, Section 8.1.9).

8.1.1 *Nodes: The Fine Print*

As discussed in Chapter 6, a graph ($\mathcal{G}$) is defined by a node-set $\mathcal{V}$ (representing entities) and edge-set $\mathcal{E}$ (representing dependencies), i. e. $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. In a lot of real-world datasets, the graphs are rich and have associated attributes. These attributes provide enhanced representation capabilities and can also capture domain knowledge. For our use case, the nodes of the graph are tanks (IT) and DOMs (IC). The EAS events mapped as a graph have rich node-attributes. The features associated with each node of the graph (or each DOM/tank in an event) can be categorized into three types, namely:

- **Charge-like:** It is evident from multiple shower-observables discussed in Chapter 7 (and [174, 296, 530]), that the charge-deposit in IC and IT detector has composition sensitivity. Hence, using charge-deposit information seems to be a natural choice. The signal pulse at each tank/DOM can be summarized in multiple ways. The ones used for this work are:
    1. **Total Charge** - For IT, the HLC-pulses obtained after cleaning using a plane-wave approximation of shower-front are used. For IC, the pulses cleaned around track given by IT are used[2]. Each tank/DOM, can observe multiple hit-deposits in an event. The hits observed at each IT-tank and IC-DOM are then used to get the integrated charge and assign the *Total Charge* at the tanks/DOMs.
    2. **Maximum, Mean and Median Charge** - Same pulses (as for *Total Charge*) are used to quantify the maximum, mean and median charge observed at each tank/DOM.
    3. **Charge in first and last deposit** - Same pulses as before are used to get the charge deposit in the first and last hit in a tank/DOM.
- **Space-like:** It is evident from Section 7.4.1 that the spatial extent of an EAS has composition-sensitivity. Prior works [181, 296, 551] at ICNO also see composition sensitivity based on the shower-observables which utilize/depend on the spatial-extent information. Hence, utilizing the spatial information should also help improve composition-sensitivity[3]. The space-like features used in this work are:
    1. **True Coordinates** - It uses the same tanks/DOMs as used for charge-like features. The location of tanks/DOMs in ICNO is read-out.
    2. **COG-centered Coordinates** - It uses the tanks/DOMs as before. However, the origin of the coordinate-system is shifted to the charge-weighted COG (separately for IT and IC). This was done because if a shower's footprint is translated (at another shower core, keeping the shower-axis parallel), then

2 **For IceCubers:** IT: *IceTopLaputopSeededSelectedHLC* ; IC: *CoincLaputopCleanedPulses*

3 The information is also crucial to the edge-construction of the graph (discussed in Section 8.1.2) and to the GNN convolution layers used in this work (discussed in Section 8.1.4 and Section 8.1.5)

the deposit in the gedanken-experiment[4] should look very similar to the one seen before[5]. Hence, it is appropriate to quantify the charge-deposit location in approximate shower coordinates.

- **Time-like:** Studies [552–555] performed at other CR-detectors like *Pierre Auger Observatory* have suggested that the rise-time of the signal observed at detector components is a mass-sensitive parameter. Performing such studies at IC and IT is an ongoing effort. However, connected studies in gamma-hadron separation at IT [183] have come up with shower-observables that use temporal information, and also show composition sensitivity in addition to providing gamma-hadron separation capabilities. Hence, it is pragmatic to use time-like information in addition to charge and space-like features. In order to ensure that two similar events, observed at different times in the detector[6], are assigned similar time-like attributes at the tanks/DOMs, the time is measured with respect to the reconstructed time of shower-core at IT. The time-like features used in this work are:
    1. **Time corresponding to maximum charge deposit** - It uses the same tanks or DOMs as used for charge-like features. The time of maximum charge-deposit (w.r.t. reconstructed-time of shower-core at IT) is assigned for each tank and DOM-hit.
    2. **Time for first and last charge deposit** - Using the same scheme as before, the time of first and last hit is assigned.

8.1.2 *The Beauty of Meaningful-Connections*[7]

The information captured by the locality of a pixel in a grid-shaped image generally has a correlation with it. CNNs excel at extracting information from this correlated input (using local-information-aggregating kernels). Similar to grid-shaped input used in CNNs, graphs also define the locality of a node (equivalent to a pixel) by its edges (used in GNNs). For CNNs the locality is generally predetermined by the grid-shaped input. In contrast to this most of the graph mapped systems (GMSs) don't have a predefined notion of edges[8] and hence the notion of locality. It holds for our use case too[9]. Hence, defining node neighborhood provides us with the opportunity of directing our GNNs to a set of possibly true solutions (predicting true CR composition for our use case) in an efficient manner. The following text will detail how this opportunity is availed.

Inductive Bias[10]

It is well known that NNs are able to model almost any function to arbitrary accuracy, and hence they are sometimes referred to as *Universal Approximators* [558–563]. Even

---

4 a thought experiment
5 under the hypothesis that we have detector components everywhere and ice-properties are similar everywhere
6 *"Time flies over us, but leaves its shadow behind"* - The Marble Faun [556] , Nathaniel Hawthorne
7 *"We are all connected in the great circle of life."* - The Lion King, Mufassa
8 not true for most natural GMSs like proteins, drug-interaction, prey-predator, etc.
9 The following text will discuss that the notion of spatial-locality of neighboring DOMs to a parent DOM is considered and used.
10 Also referred as *Prior*.

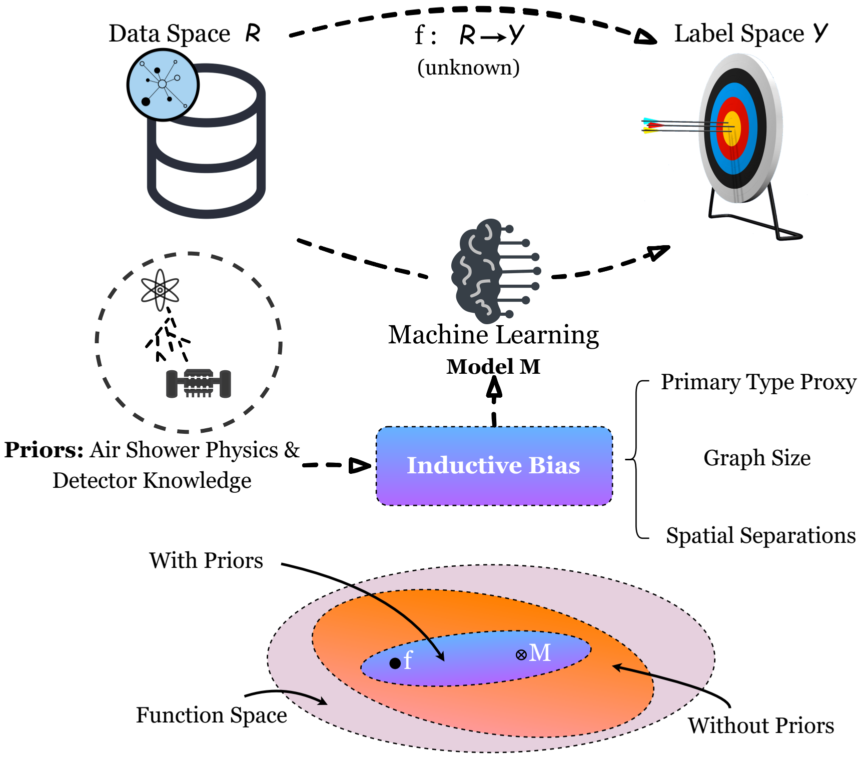


Figure 8.2: NNs are used to learn a mapping (M) from Data Space ($\mathcal{R}$) into the Label-Space ($\mathcal{Y}$), of an unknown function f. NNs are generally able to learn a good set of possible solutions (labeled as "Without Priors") by training, in the available dense "Function Space". Adding tight Inductive biases (priors) help reduce the space of possible solutions even further (labeled as "With Priors"), with greater chances of the learned function being closer to the truth. The inductive-biases mentioned here are for the use case of CR-composition estimate at IT-IC (read Section 8.1.2 for details) Illustration inspiration by [557]

a simple NN can produce a dense class of functions, each with similar accuracy on a dataset. However, in general, we cannot ensure that any given learned function is also generalizable. It is hence essential to reduce the search space of possible solutions for a NN [564]. This is in contrast to how ML is generally practiced today. Generally, end-to-end learning [421, 565, 566] with minimal hand-engineering is preferred in ML. However, the excellent successes achieved (in vision, language processing, decision making etc.) using this approach have only been possible because of the availability of huge volumes of training data and because of access to cheap, fast, and efficient computational resources. This hasn't been the case always[11]. More efficient solutions were preferred over flexible methods/models. In the limited training-data domain, which is generally the case in High-Energy Physics and is also true for this work, we need ways to direct our ML-model to obtain generalizable solutions in an efficient manner[12].

11 Fields like relational reinforcement learning [567] and statistical relational learning [568] have specifically focused on developing generalizable models with minimal dependencies on the training dataset.

12 Similar to biology, we need to jointly depend on *nature* and *nurture*, in order to benefit from the complimentary strength of structure (generalizable model) and flexibility (data-driven model). Inspiration from [438].

This can be done by biasing our training procedure in a way that constrains the model to search in a limited space of learned functions. A bias that allows such is generally referred to as *Inductive Bias*. This is illustrated in Figure 8.2, wherein the search of the true mapping function "f" (from Data Space $\mathcal{R}$ into the Label-Space $\mathcal{Y}$), we put *inductive biases* (Priors) which limit the space of possible solutions for the learned-mapping $\mathcal{M}$. The inductive biases mentioned in the figure are specific to our use case and will be elaborated upon in the following text. Most NNs are already build-up of entities (non-linear activations, regularization-methods, dropout, normalization-methods, optimizer, loss-functions etc.) which prefer one set of solutions over the other. However, GNNs provide the additional benefit of allowing us to use *relational inductive biases*. *Relational inductive biases* refer to the set of *inductive biases* which limits the set of possible interactions between the inputs (node-information here) during the learning procedure, and hence in the process constraining the set of possible solutions. From hereon "relational inductive biases" for graphs will simply be referred to as "inductive biases".

Leveraging Inductive Bias for Composition Analysis

As mentioned in Section 6.1, the building blocks of a graph ($\mathcal{G}$) are its nodes ($\mathcal{V}$) and its edges ($\mathcal{E}$). Normally, the nodes and vertices also have associated attributes (read Section 6.1.2 for details). Section 6.5 and Section 8.1.1, details that for our use the DOMs (IC) and tanks (IT) are interpreted as the nodes of the graph. Associated to each node we have charge and spatio-temporal measurements. Section 8.1.1 mentions the details of how we aggregate these measurements on a node level. There is a limited amount of operations we can perform on these measurements. To approach the real solution we should introduce inductive biases which help in estimating the composition. It is difficult to introduce any kind of inductive bias in the node attributes since these are based on point-level measurement. A visible place left in a graph to introduce inductive biases, is in the process of constructing edges between the nodes. There are also multiple reasons to include inductive biases for our use case. Some of them are:

- Depending on the energy of the primary particle, the size of the shower-footprint changes. With an increase in energy (for the same primary), we should expect more DOM-hits [98]. Since, this work focuses on estimating CR-composition from such graphs (/footprint), our edge construction should account for the size of the footprint (/cardinality of the graph). Showers from the same primary particle should be embedded closer to each other, independent of the size of their footprint at IT-IC.

- In addition to mapping graphs of similar size closer to each other (in the embedding space), we should also ensure that graphs (/footprint) from different primary types are mapped farther to each other[13].

- In the simplest case graph edges have no associated weights to them. Hence, for such cases during the message-aggregation phase of a GNN (read Section 6.3 for details), all the neighbors of a node have equal contribution. For our use case, when the nodes are interpreted as DOMs and tanks, the neighbouring-DOMs (/nodes) to a specific DOM (/node) will mostly be located at a different spatial

13 **Summary** : Our edge-construction should try to decrease intra-class distance and increase inter-class distance, in the embedding space.

distance to the parent DOM and tank (node). It is more likely that two nearby DOMs and tanks have deposits from nearby sections of the Cherenkov-cone (For IT: daughter-particles from nearby-sections in shower-profile), than two far off DOMs/tanks. Hence, when defining neighbors of a particular DOM (/node), it makes sense to weigh the edge between two nearby DOM (/node) more than two far-off DOMs.

By introducing *inductive biases* in the edge construction, we also intend to improve upon the accuracy on composition prediction[14]. In the following, we will discuss how the requirements to introduce *inductive-bias* (for edge construction) are introduced in an iterative manner:

step 1: DECIDING EDGE CONSTRUCTION POLICY - Most biogenic and anthropogenic systems already have a predefined interaction scheme (protein-drug interaction: chemical bonds; prey-predator system: food chain; flight network: flight connections). However, for our use case (signal deposit at DOMs and tanks) we don't have a clear interaction scheme. At best we can say that two-nearby DOMs (tanks) should have correlated deposits. Since our graphs(/ events/ point-clouds[15]) are embedded in euclidean-space, in the simplest case it makes sense that we connect each node (DOM) to its spatial neighbor. This can be done using the k - Nearest Neighbours (kNN) approach. Here "k" is the number of neighbors each node (DOM) has in a graph. The upcoming steps will discuss how "k" is decided for each graph. The kNN connections are made using the spatial location of the node(/DOM)[16]. This edge-construction policy is also common in other works in the field of High-Energy Physics (HEP) and astrophysics [577–579].

step 2: CARDINALITY PREDICATED INDUCTIVE BIAS - As mentioned earlier, for the same primary particle we want the prediction of the model to be independent of the size of the footprint (cardinality of the graph). To understand how this can be achieved it is important to make connections between CNNs and GNNs. CNNs learn by transforming an input into a higher-dimensional embedding-space by iteratively moving a kernel[17] with learnable parameters over the input/intermediate representation. The kernel size in a CNN determines the local neighbourhood (in addition to the receptive field - Read Section 6.3.2.2 for details) of a certain pixel. A bigger kernel means that each pixel in the input/intermediate-representation (in a CNN) can benefit from the information of more neighboring pixels. Similarly, for GNNs if kNN algorithm is used to make edges in a graph, then a greater "k" means more neighbors/connections. Hence, the kernel size in CNNs has equivalence with the value of "k", used in making

14 Section 8.1.6 will detail how the model still has the possibility to move away from the inductive-bias introduced here.

15 Point cloud is a set of data-points embedded in a three-dimensional Euclidean space, represented by the coordinate-locations. The events observed at ICNO can be viewed as point clouds. These point clouds can then serve as a GMS, where the nodes are the point-cloud locations and the connections/edges are based on spatial locality.

16 Making kNN connections in feature-space (spatiotemporal measurements + charge) for the input-graph was also tested. This did not yield any visible gains in the accuracy of predictions. Other methods [569–576] were briefly-explored and could be pursued further.

17 For CNNs, "kernel" is a matrix which convolves (slides-over) over a given input and updates the input in the process, using a mathematical operation (dot-product, maximum etc.) with learnable parameters.

kNN edges/connections, for the use case of GNNs. CNNs generally use a stack of layers (output of one layer serves as input for the next one), with convolution-operation (using kernels with learnable weights) happening at each layer. For the purpose of classification (/regression), after the n^{th} layer the matrix-like (/image-like) embedding is flattened[18] to get a vector of fixed size. This then serves as the input for a MLP[19]. Similarly, GNNs also use stacking of layers (although much fewer than CNNs - Read Section 6.3.2.2 for details of why). To get a global-level representation (for classification / regression) the node-level embeddings are pooled[20] to get a fixed-size vector which then serves as the input for a MLP, with everything being trained in an end-to-end fashion. For a classification (/regression) task it is hence crucial that the convolution part of the CNN (and the graph part of GNN) finds a good representation of the input. A good representation for classification (/regression) should have minimal intra-class distance and maximal inter-class distance.

To understand how we can obtain a good representation, let's take an example of "cat-classification"[21]. If it is easy for a human to classify an image, in general it should also be simple[22] for a CNN to do the same [585]. To replicate the convolution-operation in CNNs, Figure 8.3 presents a set of cat images. The top row shows a set of input images of a cat. The first-cat in the top-row represents the original image. For the other two, the first image in the row is scaled down. The dimensions (number of row and columns) of the input are kept the same. The red square represents a Gaussian-blur kernel [586], which is moved over the input. The kernel-size remains the same between the first two images. For the third image, the kernel size is scaled down (by the same scale as used to scale down the cat). The bottom row presents the output after one convolution operation[23]. As mentioned before, we want to reduce intra-class distance. For cat-classification by eye, this translates to the ease of classifying/recognizing the image of a cat (and its attributes) after convolution, independent of the size of the cat. In the first column of Figure 8.3, it is easier for us to recognize a cat and its attributes (halo, whiskers, heart, wings, and maybe even eyeballs) in the output image of a convolution. If we keep the kernel size the same (second column), it is a bit more difficult to recognize a cat, and all the attributes are almost washed out. However, if we reduce the kernel size, recognizing the cat (and its attributes) becomes easier again. If we only focus on the active pixels in the first and last column of the image, the blurred output seems to be similar. We want to reduce intra-class distance, hence it is more favorable for us to adapt our kernel size according to the size of the cat. Since in CNNs the convolution operation happens repeatedly, it is more likely that the big-cat (column 1 in the figure) and small-cat (columns 2 and 3 in the figure) will be embedded closer in embedding space when the kernel-size was adapted to the size of the cat, than otherwise. This idea can easily be translated to the use case of

18 Here, flattening is the reshaping of a (n, n) matrix into a column vector of shape $(n * n, 1)$.

19 Everything is trained in an end-to-end fashion. Read [580] for a brief overview of how CNNs work.

20 This work uses feature-wise add and mean pooling. However, other methods [465, 581, 582] were also tested.

21 "One cat just leads to another." — Ernest Hemingway [583] (Letter from Finca Vigia, Cuba, to his first wife, Elizabeth Hadley Richardson in 1943.)

22 There are still areas where humans-excel and ML based models fail miserably [584]. There is still some hope left :-P.

23 CNNs learn the weights in the kernel during the learning procedure. However, the subsequent pooling operation performed (to get the output) is the same and is effectively mostly a blurring operation.

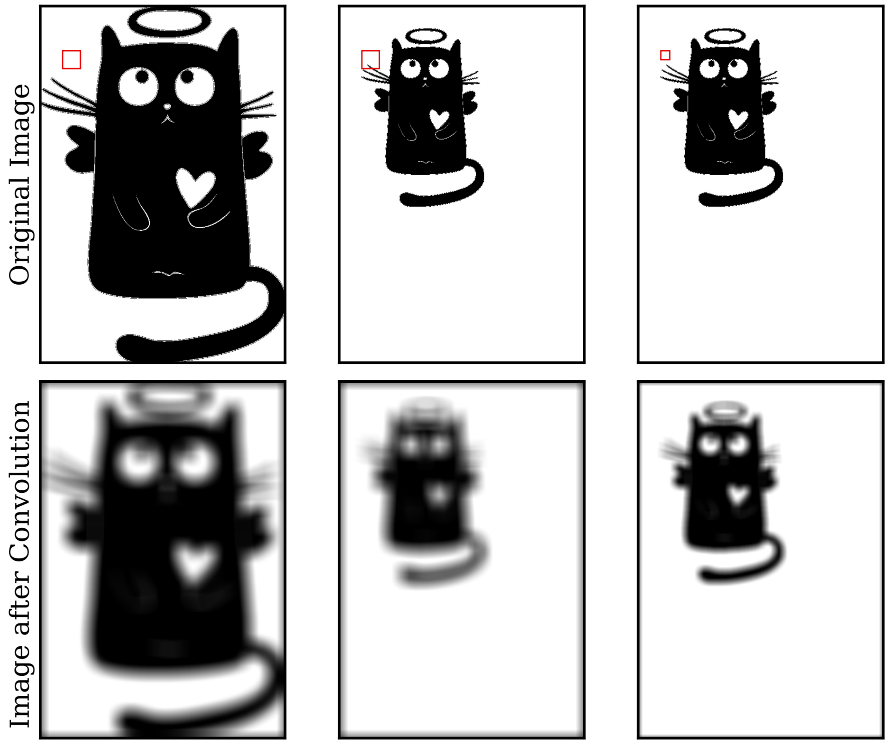


Figure 8.3: An example demonstration presenting kernel-size dependence on the convolution output (bottom row). The kernel (red-square) chosen here is a Gaussian-blur kernel [586]. The kernel-size remains the same for the first two columns and is scaled down for the third one. To obtain a similar level of blurring (for cats of different sizes), it is essential that we also adapt the kernel size accordingly. Read Figure 8.1.2 to understand how the inspiration of adaptive kernel size motivates an important cardinality predicated inductive bias in the edge-construction procedure of the input-graph of the GNN. If you are not a cat person, visit Twitter post made by Paras Koundal on April 12, 2023 (https://shorturl.at/rtvL2), to see the same thing for a Penguin image.

graphs and GNNs. The size of the cat can be thought as the cardinality of a graph[24]. The kernel size (for CNNs), as mentioned earlier, has equivalence with the value of "k" used in making kNN edges, for the use case of graphs. Hence, if we use a GNN based algorithm to perform "cat classification", we should have a higher "k" in the kNN edge-creation method for the bigger cat than the smaller-one. To summarize, for GNNs to reduce intra-class separation, we should make more connections (higher "k") for a bigger graph (higher cardinality) and vice versa. It is now crucial to understand why the idea of "cardinality-predicated kNN" should be useful for the use-case of CR-composition prediction using GNNs.

We know from Chapter 7 that we have composition sensitivity in multiple shower-observables. These shower-observables are driven from different aspects of the charge deposit seen in IC. Similar holds for the footprint seen at IT [174, 296, 530]. Hence, it is crucial that we maximally benefit from all the aspects with minimal information loss for composition prediction, independent of its primary energy. One major change

24 Bigger-cat is equivalent to saying that there are more-active pixels; and pixels can be interpreted as the nodes of a graph.

we see in charge-deposit, with an increase in energy, is the size of the footprint. For a primary of higher energy, on average we expect more DOM-hits [98], than the same primary at lower energies. When the DOM-hits are mapped as nodes of the graph, we should take care that the finer details of the charge deposit are not totally washed out during the message-passing operation of a GNN. Drawing parallels of the footprint-size with the cat-size hence leads to the conclusion that for bigger event (shower with more DOM and tank hits) we should have more connections and vice versa. This is a crucial *graph-cardinality predicated inductive bias*, that has been introduced in this work[25]. **Step 5** will discuss how the model still has the possibility to transition from a cardinality-dependent edge construction to an independent one, and the decision was based on model performance. This allows us to decide our final architecture choice not only on the expectations laid out in this section but also based on the model's accuracy and performance.

step 3: COMPOSITION PREDICATED INDUCTIVE BIAS - The last step helped introduce the essential inductive bias needed to reduce the intra-class distance between different graphs (/footprint) belonging to the same category (/primary type). The next sensible step is the introduction of an inductive bias which increases inter-class distance. For our case, this translates to the introduction of *composition predicated inductive bias*. However, this seems at odds with the task at hand. How can we introduce a *composition predicated inductive bias*, if we are trying to predict the composition? Chapter 7 again comes to our rescue here. We have a significant number of composition-sensitive parameters already. These can help us provide a hint of composition to the model (via the process of edge construction). **Step 5** will present how it is introduced.

step 4: PROXIMITY PREDICATED INDUCTIVE BIAS - Is it likely that signal deposit in two DOMs/tanks with greater spatial proximity will be from closeby part of the Cherenkov-cone (For IT: daughter-particles from nearby sections in shower-profile), than two far off DOMs/tanks. Hence, for every node (/DOM and tank), it is sensible to assign more weight to the edges (as edge-attribute) which are spatially closer (to that node) and vice versa. Other works [577, 578] explore similar methods for edge-weighting for HEP-experiments.

step 5: THE FELLOWSHIP OF THE MODALITIES[26] - The previous steps mentioned various methods to introduce *inductive bias* during the edge construction of air-shower events mapped as graphs (/point-clouds). One additional thing which has to be considered (in addition to incorporating the *inductive biases*), is that we prevent the problem of over-smoothing (read Section 6.3.2.3 for details) and over-squashing (read Section 6.3.2.4 for details). Also, with an increase in value of "k", the computational footprint (storage and computations) increases[27]. Hence, the maximum allowed "k"-value

25 Section 8.1.6 will discuss how we still keep the scope (for the model) to explore possible-solutions beyond the solution-space allowed by the introduction of the inductive bias.
26 Inspired from *The Lord of the Rings: The Fellowship of the Ring* [587].
27 Bigger "k" for a graph means we will have more number of edges (or equivalently more and/or bigger tensors during training).

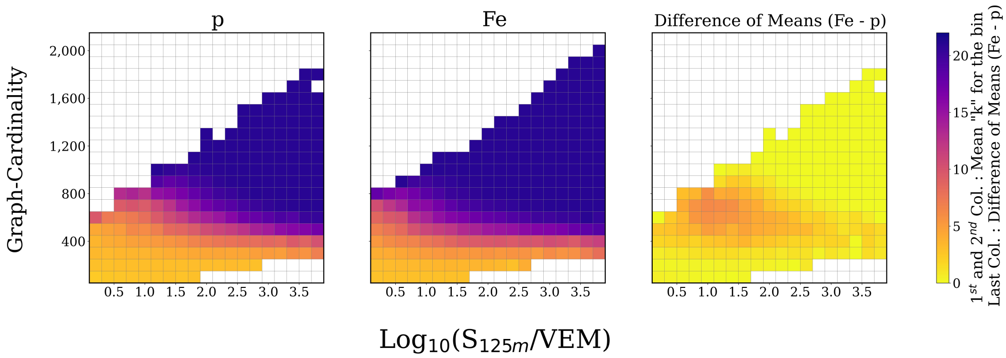


Figure 8.4: Mean-"k" (in a bin), for the training-dataset, as a function of graph-cardinality and $\text{Log}_{10}(S_{125}/\text{VEM})$. The k value for kNN-edge construction for every DOM is obtained from Equation 8.1. The first and second columns illustrate the mean value for p and Fe respectively, for the training dataset. The third column represents the difference in means (between Fe and p). Read Section 8.1.2 for details.

for any event (/graph) cannot be arbitrarily big. The form of "k" used to build kNN-connections for a graph with cardinality $\mathcal{N}$ is given by:

$$k = 1 + \left\lceil \frac{\mathcal{K}_{Max.}}{1 + A \cdot \exp[\mathcal{R} \cdot (B - B \cdot \mathcal{R} \cdot \mathcal{N})]} \right\rceil \tag{8.1}$$

where k is the number of edges for the graph with cardinality $\mathcal{N}$. $\mathcal{K}_{Max.} = 20$ is the maximum number of allowed edges for a node in any event (/graph)[28], in the MC as well as the burnsample. $\mathcal{R}$ is the *Ratio Parameter*[29] (Read Section 7.3 for details.). $A = 20$ and $B = 0.1$ are the hyperparameters of the model. The form of k allows the transition from a model with *cardinality predicated inductive bias* (for $\mathcal{A} \neq 0$) to one without it (for $\mathcal{A} = 0$). A sparse hyper-parameter sweep favored a non-zero value of "k". "$\lceil\ \rceil$" denote the ceiling function [588]. As can be expected from Equation 8.1 (adapted sigmoid squashing function), with an increase in the cardinality of a graph $\mathcal{N}$, the number of edges increase[30]. After the decision of neighbors; using kNN algorithm, for each node, in a graph (/event), the edges connecting neighbors of a node are weighted using their euclidean-separation (with weights summing to 1 for each node)[31].

Figure 8.4 presents the binned mean-"k" in the training dataset, as a function of graph-cardinality and $\text{Log}_{10}(S_{125}/\text{VEM})$ (i. e. the average number of connections, for each-DOM in the event, belonging to events in a particular cardinality-$\text{Log}_{10}(S_{125}/\text{VEM})$ bin). The first and second columns represent the distribution, in the training dataset, for p and Fe respectively. The third column represents the difference of means (between Fe and p), for each cardinality-$\text{Log}_{10}(S_{125}/\text{VEM})$ bin. As can be guessed from Equation 8.1, the difference here is primarily the result of *composition predicated inductive bias* mentioned earlier[32]. Further studies can focus on improving the strength of multiple inductive biases introduced in this work and even introducing more. Figure 8.5

28 $\mathcal{K}_{Max.} = 20$ help prevents over-smoothing and over-squashing problem discussed before.There are 20 connections to the nearby nodes and an additional self-loop for each node exits.
29 Introduces the *composition predicated inductive bias* into the edge-construction, discussed in **step 3**.
30 Introduces the *cardinality predicated inductive bias* into the edge-construction, discussed in **step 2**.
31 Introduces the *proximity predicated inductive bias* into the edge-construction, discussed in **step 4**.
32 Since, for the same cardinality-$\text{Log}_{10}(S_{125}/\text{VEM})$ bin, the only difference between p and Fe, for k-value calculation (given by Equation 8.1) is in the value of $\mathcal{R}$ (composition sensitive *Ratio Parameter* - Section 7.3).

represents an example EAS (simulated), with coincident deposits in IT-IC, with edges between the tanks/DOMs built using policy laid out in Equation 8.1. The graph thus constructed serves as input for the GNN-model (labeled as "Input Graph" in Figure 8.1).

8.1.3 *Homophily*

Homophily of a graph gives a quantitative measure of the probability of nodes sharing the same labels being located in close proximity to one another within the graph[33]. There are multiple ways to measure homophily in a graph. The most prominent among them are "Edge Homophily Ratio" and "Node Homophily Ratio"[34]. *Edge Homophily Ratio* refers to the proportion of edges connecting nodes with identical class labels within a given graph [591]. *Node Homophily Ratio* is the Edge Homophily Ratio normalized across node-neighborhoods [592]. Both should have values between 0 and 1. Graphs with high homophily will have a high edge and node homophily. For the use case of this work, *Node Homophily Ratio* is used and will simply be referred to as *Node-homophily* or *Homophily* hereon. As can be expected from the definition, changing the node-connections of a graph should also change the value of homophily[35].

Most GNN-based architectures are designed to work in a high-homophily regime and fail miserably otherwise. Table 8.1 (adapted from [591]) shows the mean classification accuracy for a synthetic graph dataset (details in [591]), for low-homophily (h=0.1) and high-homophily (h=0.7) regime. It can be seen that when we move from a high-homophily regime to a low-homophily regime, we see a significant drop in classification accuracy for all GNN based architectures[36]. Some recent works [593, 594] have shown that it might still be possible to use the standard GNN architectures in a low-homophily regime. However, it is still not clear if these results hold for any general graph dataset. A workable solution chosen by quantifying the homophily and include the knowledge in the GNN architecture.

For our use case, we have features associated to each node. The individual features can be considered as labels of the node. It is then trivial to calculate *homophily* for features associated with nodes (DOMs) of the graph. In Figure 8.6, Figure 8.7, Figure C.2 and Figure C.3 we see homophily measure for few of them (in-ice deposits only)[37]. Except for Figure C.3, the choice of all hyperparameters is the same as the final-model (which is used to get results in Chapter 9). Important observations from the figures are:

1. Total Charge in general has low-homophily ($\leqslant 0.6$), for both p and Fe. This is depicted in *Top Panel* of Figure 8.6. This is expected from Section 2.2.1.2. The in-ice charge deposit is primarily because of TeV-muons. Stochastic bremsstrahlung-

33 The idea also has very deep connections and is useful in studying social networks [589]. The underlying social graph in most social networks is a high-homophily graph since networking mostly happens between individuals sharing similar interests (*like attracts like*).

34 Read [590] for other ways to measure graph homophily.

35 An example graph will likely have a higher homophily if it was fully connected (every node connected to every other node), than if it was otherwise.

36 The architectures listed in the table are among the most famous GNN architectures. GNNs is a very active field of research and there are way more GNN architectures. PyTorch-Geometric's [444] website (pytorch-geometric.readthedocs.io) gives a list of few other famous GNN architectures.

37 The underlying simulations are from the training-dataset and haven't been weighted to any flux-model.

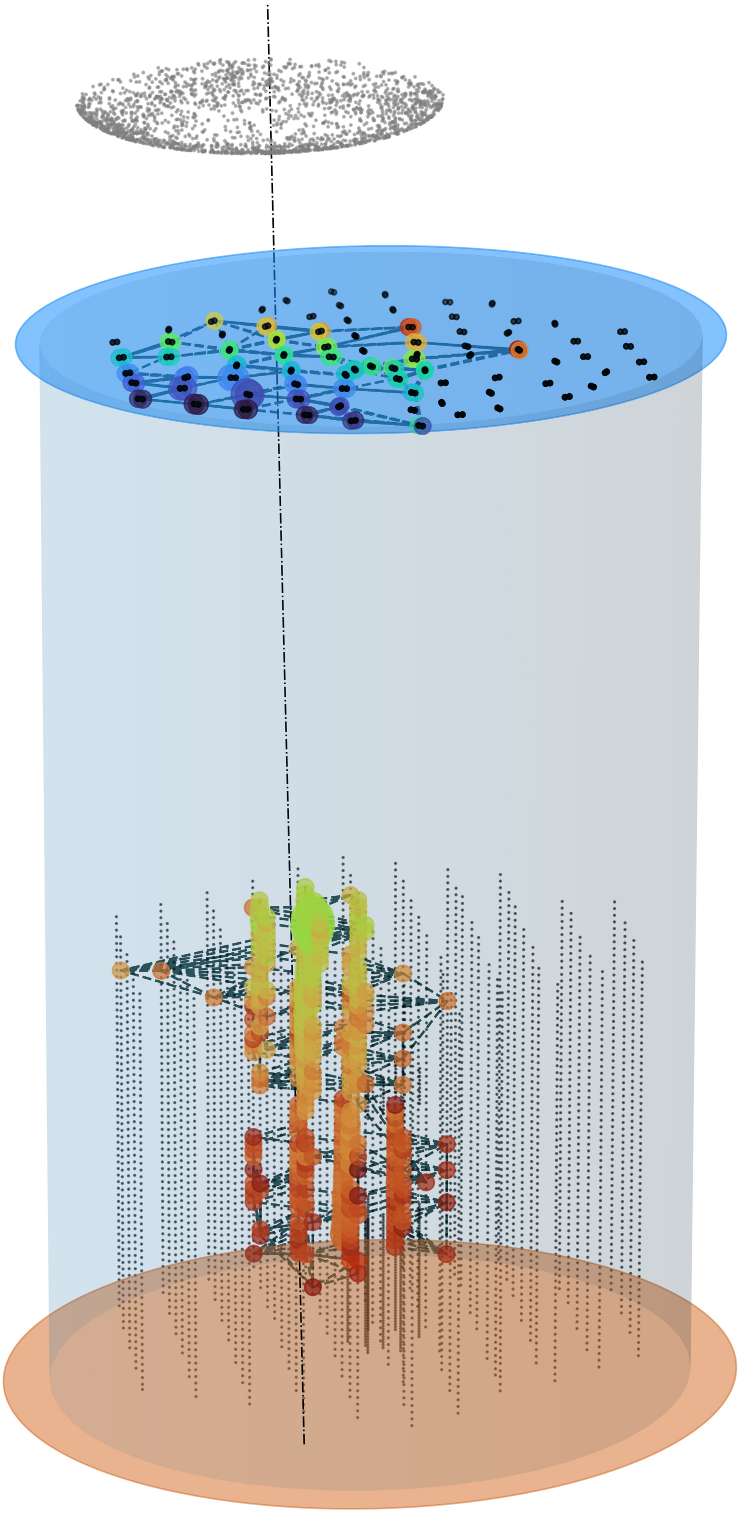


Figure 8.5: An illustration of a simulated EAS, with coincident footprint at IT and IC. The color of the footprint here corresponds to the time of the hit (blue = early, red = late) and the radii of the circles correspond to the signal strength at the station/DOM. The dashed lines, connecting the active tanks and DOMs, are built using the kNN-policy given by Equation 8.1. The event mapped as a graph serves as the input (labelled as "*Input Graph*" in Figure 8.1) for the GNN-architecture (Figure 8.1). The track (slanted dashed-dotted line) and shower front (cloud of grey points on top) are also depicted for visualisation purposes.

GNN ARCHITECTURE	$h = \mathbf{0.1}$	$h = \mathbf{0.7}$
GCN [453]	37.14±4.60	84.52±0.54
GAT [595]	33.11±1.20	84.03±0.97
GCN-Cheby [449]	68.10±1.75	84.92±1.03
GraphSAGE [452]	72.89±2.42	85.06±0.51
MixHop [596]	58.93±2.84	84.43±0.94

Table 8.1: Mean classification accuracy for a low-homophily (h=0.1) and high-homophily (h=0.7) regime, on a synthetic graph dataset (details in [591]), for famous GNN based architectures. The table is adapted from [591].

loss is the dominant energy loss mechanism at these energies. Because of the stochastic nature of these deposits, it will become less likely that two neighboring nodes (DOMs) have similar charge deposits. Hence we should expect lower homophily. From Section 7.2 we know that stochastic deposits increase with an increase in energy. Hence, we should see a drop in homophily as we move towards air showers with higher primary energy. This can also be seen in Figure 8.6 (*Top Plot*), where we see that at higher-energies (bigger $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$-value [38]) maximum events have lower homophily. We see visually similar homophily for the time of maximum charge deposit (Figure 8.6: *Bottom Plot*) and the z-coordinate of the triggered nodes/DOMs (Figure 8.7: Bottom Plot), with the coordinate-origin at the charge-weighted COG[39]. A detailed study of if stochastic deposits indeed are the only underlying reason for lower homophily is needed.

2. Since we don't expect any major azimuthal difference in the footprint of vertical air-showers at IT[40], we should also expect an almost azimuth-symmetric transverse-footprint in IC, for IT-IC coincident showers[41]. Hence, neighboring nodes (DOMs) for our case should have a similar value of x and y coordinate (with the coordinate-origin at the charge-weighted COG). Hence, we should expect higher homophily for charge-weighted x and y coordinates, than z-coordinate. This can be seen in Figure 8.7 (Top Plot) and Figure C.2, in comparison to Figure 8.7 (Bottom Plot). From Figure 7.7, we know that *Mean Radii* for TeV-muons is slightly different for p and Fe-initiated showers. The consequences of it can also be seen here, in the differences of distribution between p and Fe-initiated showers among the coordinates.

3. We generally see a difference in homophily distribution, between p and Fe, for the choice of most features. It is a strong motivation to use homophily as a feature for CR-composition analysis at ICNO.

38 *Reminder:* $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ is an energy-proxy (see Figure 5.6).

39 It is challenging to draw a conclusive inference solely based on the visual similarity of the three graphs, as it remains uncertain whether additional node-attributes do not provide supplementary information to the network. Homophily is a measure of similarity within a node attribute and doesn't capture any information about the correlation among node attributes.

40 For vertical air-showers the footprint looks circular. It shifts to a an ellipse-like footprint with an increase in incident zenith-angle. This work uses EAS with a maximum zenith angle of about 36 degrees (see Figure 5.8 for the zenith and azimuth distributions of the dataset).

41 This is an ongoing area of research. Works like Section 7.4 are a step towards this direction.

4. The gap in homophily ($0.4 <$ Homophily $\leqslant 0.5$) seen in a feature like Total Charge (Figure 8.6 : *Top Plot*) is an artifact of how-many neighbors each node (/DOM) in a graph has and how neighbors are defined. The gap vanishes on an increase in the maximum number of neighbors from 10 to 40. This can be seen in Figure C.3.

8.1.4 *Point-Pair Convolution*

The events at ICNO mapped as graphs/point-clouds with attributes (discussed in Section 8.1.1) and edge-construction (discussed in Section 8.1.2) has CR composition-information embedded at local as well as global-scales (Read Chapter 7). Point-Pair feature convolution (PPFConv) [597, 598] is designed to effectively capture unique 3D local features from point clouds in an efficient manner and can even work with sparse point clouds. It builds upon the PointNet-architecture [599], which is among the most prominent DL architecture for point-clouds (read Section C.4 for a very brief summary of PointNet architecture). PPFConv was specifically designed to find correspondences between two partially overlapped point clouds. Hence, it lays special focus on learning good local discriminative features which can describe 3D local patches from point clouds. The core components in the convolution are:

- **Local Patches**: Local patch for a reference node in a point cloud refers to its neighborhood. It is crucial to learn a good representation of the input data (or embeddings) to get the best possible model performance. PPFConv uses multiple methods to enrich the geometric information, namely:
  - Use the given spatial information of the point cloud (coordinates).
  - Evaluate and use the surface-normal in the local patches [600].
  - Evaluate point-pair features for the local patch [601–603]. This is given by the 4-vector:

$$F(r_i, r_j, n_i, n_j) = \left[\| d \|_j, \angle(n_i, d), \angle(n_j, d), \angle(n_i, n_j)\right] \tag{8.2}$$

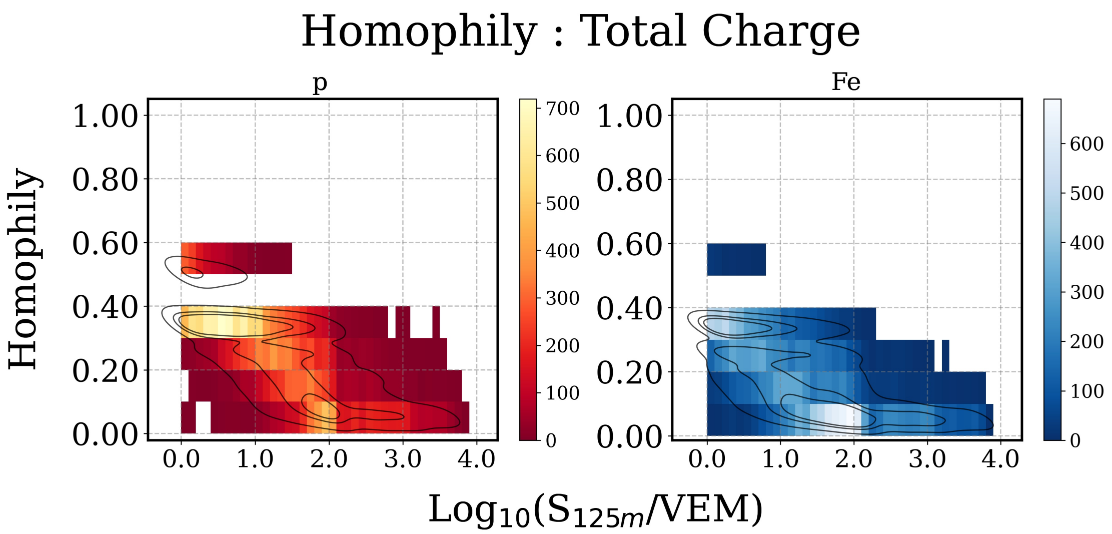
    where r_i and r_j give the spatial location of the two nodes (points/DOMs). n_i, n_j represent their normals and $d = r_j - r_i$. Here, $\angle(a, b)$ represent the angle between the vectors a and b.

- **Network**: The local-patches are treated individually using PointNet[42]. The individual PointNets in the local patches learn local representations. These local representations are then aggregated using a max pooling operation to give a global representation. The global representation is then concatenated to the local features to give global context in the local patches which can then be used further. The resulting representation is utilized as input to MLPs, which help learn local representations that are globally contextualized.

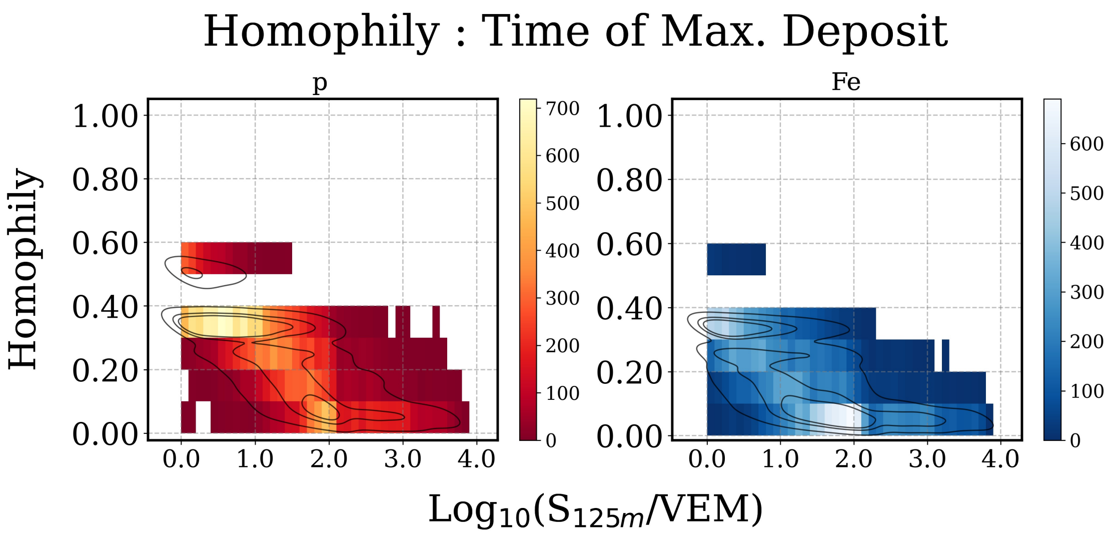
The components can be written together as:

$$X_i' = \gamma\left(\max_{x_j \in X} \phi(X_j, F(r_i, r_j, n_i, n_j))\right) \tag{8.3}$$

42 This is done concurrently to speed up the evaluation.

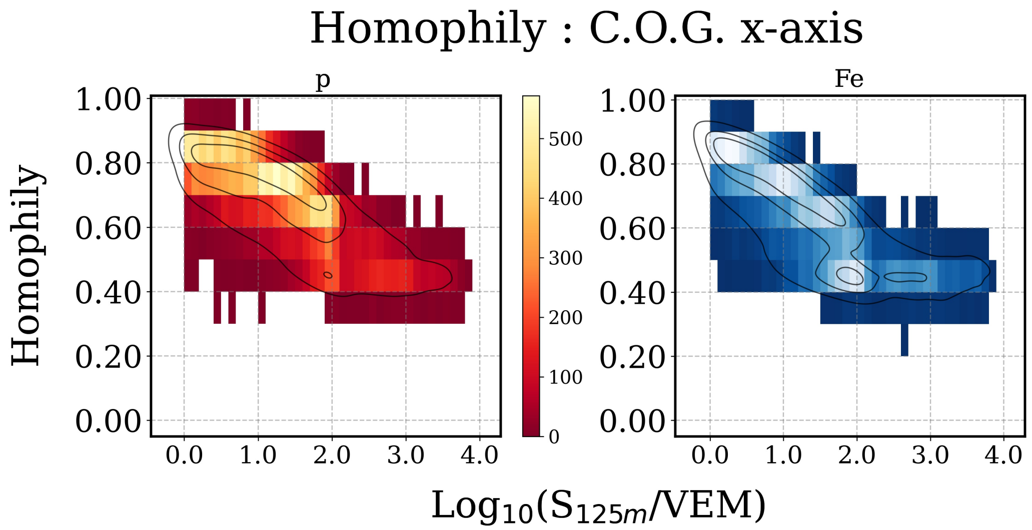


(a) Homophily for Total charge measured at a node (DOM)

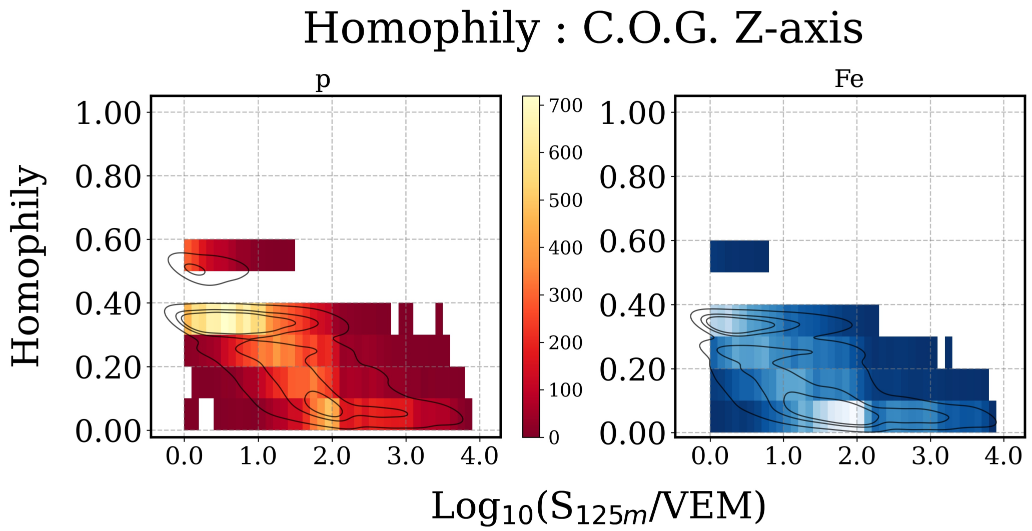


(b) Homophily for Time of maximum charge-deposit at a node (DOM)

Figure 8.6: Homophily for Total Charge (*Top Plot*) and Time of maximum charge-deposit (*Bottom Plot*) as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$, for p and Fe. Only in-ice deposits are used here. The maximum number of connections any node (/DOM) in the dataset can make (with neighboring nodes) is 10. The underlying events are from the training dataset (unweighted). The color represents the counts in a bin. The overlayed black curves represent the 2D Kernel Density Estimations (KDEs). For interpretation read Section 8.1.3.



(a) Homophily for Charge weighted x-coordinate of the node (DOM)



(b) Homophily for Charge weighted z-coordinate of the node

Figure 8.7: Homophily for Charge weighted x and z coordinate (with the coordinate-origin at the charge-weighted COG) as a function of $\text{Log}_{10}(S_{125}/\text{VEM})$, for p and Fe. Only in-ice deposits are used here. The maximum number of connections any node (/DOM) in the dataset can make (with neighboring nodes) is 10. The underlying events are from the training dataset (unweighted). The color represents the counts in a bin. The overlayed black curves represent the 2D KDEs. For interpretation read Section 8.1.3.

where X_α and X'_α represent the previous and updated node embedding (/representation) respectively. γ and ϕ represent a NN or MLP, with learnable parameters. χ defines the local-neighbourhood of i. A similar procedure is repeated at every node-neighborhood combination, with shared weights and gradients.

As mentioned earlier in Section 2.2.1.2, Section 7.1, Section 7.2, muon-deposits in IC are primarily by stochastic deposits. Also depending on the primary type, we see a difference in the deposits. To discriminate between primary types, at fixed energy, it is crucial to understand these local deposits in the global context of the shower deposit[43]. PPFConv is hence ideal to be used for our use case. For IT, we know that as we move far from the shower axis we see a higher fraction of muons (primarily GeV muons), than the electromagnetic fraction. Similar to in-ice deposits it should be beneficial to understand the local deposits at IT in the global context of the shower deposit.

8.1.5 *Point Transformer*

Transformers are a class of deep-learning models that have gained significant success in the field of natural language processing [604–607] and computer vision [608–610] in the last few years. The underlying model under the famous ChatGPT tool [300] by OpenAI is also a transformer-based model[44]. The effectiveness of Transformer based models can be attributed to their underlying "Attention Mechanism"[45] (read Section C.5 for a brief summary of Transformers and Attention and the benefits they provide over RNN.). *Attention Mechanism* improves on standard NN/CNN architectures, by adding contextual information in the training/learning process. During the training process of a model it helps determine which parts of an input are more useful (leads to higher accuracy on the task at hand) than the other[46]. Since during the training process the model learns which parts of an input are more important than the other, attention maps can also be used to understand if the model is indeed learning from expected regions of interest and a higher accuracy is not from any random part of an input [619].

43 We see a degeneracy between energy and stochastic deposits (Figure 7.4) i.e. we can obtain stochastic deposit same as a Fe-initiated shower for a p-initiated shower at higher energies. Hence, giving the global context of the footprint (as a proxy for shower energy) at local patches during the learning of GNN should be beneficial.

44 Here, I am referring to ChatGPT based on GPT-3.5. GPT stands for *Generative Pretrained Transformer* and is the most prominent among the class of Large Language Models (LLMs) (a language model with generally $\geqslant \mathcal{O}(10^6)$ learnable parameters and trained on large volumes of unlabelled text) [611]. One additional ingredient, other than Transformer, which makes ChatGPT powerful is the fine-tuning done by Reinforcement Learning from Human Feedback (RLHF) [612]. This allows for aligning language models and preventing unhelpful outputs. Google's BERT (Bidirectional Encoder Representations from Transformers) [606] which powers most of the Google Search in English [613] and in 70 other languages [614] is also a Transformer-based model.

45 [604] was the first prominent work using *Attention Mechanism* (primarily because they didn't use RNNs. Read Section C.5 for a brief summary of how *attention-mechanism* based Transformers; introduced in the paper; is better than RNNs.). However, there are a few earlier works too [615, 616].

46 *Attention Mechanism* can be thought of how humans behave when reading a text (or viewing a scene). We focus on keywords (or specific segments) rather than the full input presented to us [617, 618].

This is crucial to generate trust in any deep-learning model when implemented in the real world[47].

Among the various variants of *Attention* [616, 620–622], the one useful for our work is *Self-Attention* (introduced with Transformers [604]). *Self-Attention* allows a model to learn dependencies in an input that are independent of the output. It is performed as a *set-operation*[48], i. e. it is independent of the order of individual elements in an input. It is also independent of the cardinality[49] of the input. This makes it ideal for the use case of graphs. As mentioned in Section 6.1, permutational invariance of the input is a crucial requirement to work with graphs. Also, graphs can be of any size[50]. Hence, to learn contextual information, *self-attention* based *Transformers* are best suited for graph-based learning. "Point Transformers" fulfill the criteria for a *self-attention* based learning on 3D point clouds [623]. *Point Transformer* allows a variety of tasks including semantic scene segmentation, object part segmentation, and object classification. This is ideal for our use case since estimating the composition from the shower footprint is in principle an object-classification task, where our object is the shower footprint mapped as a point cloud, and classification is done for the primary type. *Self-attention* is locally applied in *Point Transformer*. This allows for learning which nodes/points are more important than the others. For our use case, it is equivalent to learning which DOMs should be more important during the training/procedure. Using node level importance, during the training procedure should in principle provide additional benefit to the local patch-level learning done by *Point-Pair Convolution* (Read Section 8.1.4 for details.). [623] details that *Point Transformers* are highly scalable (Can even work with point clouds of cardinality $\approx \mathcal{O}(10^6)$ easily.) in addition to being highly expressive. This also makes them flexible for implementation in IceCube-Gen2, in the future.

A single *Point Transformer* layer can be written as:

$$X'_i = \sum_{x_j \in \chi} \alpha_{ij} \odot \left(\beta(X_j) + \delta_{ij}\right) \tag{8.4}$$

where

$$\alpha_{ij} = \rho\left(\gamma\left(\phi(X_i) - \psi(X_i) + \delta_{ij}\right)\right)$$

and

$$\delta_{ij} = \Omega(r_i - r_j).$$

Here χ, r_α, X_α and X'_α have the same meaning as Equation 8.3. α_{ij} are the attention coefficients and ρ is the normalization function (softmax[51] here). γ maps (using a NN or MLP) the input into embedding-space for the attention-coefficient calculation and δ_{ij} represents the positional embedding. $\odot$ is the Hadamard product. ϕ, ψ and Ω are NNs or MLPs with learnable parameters.

47 It can also be useful for our use case (estimating CR-composition at ICNO). We already know from Chapter 7, that specific parts of the shower footprint are more useful for studying CR-composition than others. Understanding if the model is indeed learning from those can provide new insights for CR-analysis at ICNO, as well as for CR-Physics in general. A study to explore this is planned for the future.

48 A set is a collection of elements of any kind (numbers, alphabets, objects etc.).

49 Reminder : It refers to the size of the input. In the case of a sentence, it can be the number of words. For graphs, cardinality refers to the number of nodes.

50 As mentioned in Section 6.2, the ability of graph-based learning to work with input of variable size makes it more useful than CNNs.

51 $\mathrm{Softmax}(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)}$

8.1.6 *Dynamic-Edge Convolution*

Section 8.1.2 detailed an EAS physics-inspired method to create graph edges. The step introduces the important *inductive-bias* needed to perform an enhanced CR-analysis at ICNO. However, the step also fixes the connectivity among nodes in the graph. "Dynamic-Edge Convolution" is used, in order to also benefit from a fully-learnable edge creation. It was proposed in [624], and builds graph edges dynamically, using k - Nearest Neighbours (kNN), in embedding/feature space. Since the graph is created using embeddings, two nodes (/DOMs and tanks) that were not connected earlier (because of spatial separation between them) can now lie in each other's neighborhood (i.e. can create an edge between the two). This allows the model to learn non-local dependencies/correlations. Since the edges are constructed during run-time this tends to slow down the training process a bit. Hence, instead of using "k" given by Equation 8.1, a fixed value of $k = 5$ was used just for the *Dynamic Edge Convolutions*[52]. This hence also increases the receptive field (read Section 6.3.2.2 for details), while remaining sparse and hence preventing over-smoothing(Read Section 6.3.2.3 for details.) and over-squashing (read Section 6.3.2.4 for details). This convolution also preserves the permutational invariance of the input. *Dynamic-Edge Convolution* has already shown promise for low-energy event classification and reconstruction at ICNO [526].

A single *Dynamic-Edge Convolution* layer variant used in this work can be written as[53]:

$$X_i' = \max_{j:(i,j)\in\mathcal{E}} \gamma(X_j, X_j - X_i) \tag{8.5}$$

where X_α and X_α' have the same meaning as Equation 8.3. The aggregation happens over the kNN in embedding space, where $\mathcal{E}$ represents the edge-set thus formed.

8.1.7 *Bridging Ideas: DenseNet and InceptionNet*

Section 6.4.1 and Section 6.4.2 introduced the idea of DenseNet and InceptionNet respectively. They have proved essential in CNNs to ensure that training ML models don't suffer from vanishing gradient problem (definition in Section 6.4) and the models learn structures at different scales. Both are crucial to train GNN-based models efficiently and to increase the expressiveness of GNN models. *DenseNet* (details in Section 6.4.1) relied on making skip-connections between all preceding layers in a network with all the upcoming layers. Similar idea is also adapted here. One major difference between the *DenseNet* architecture and the adaption in this work is that the concatenation happens at a node level rather than the layer level. The node-level concatenation increases the node's feature-vector size but allows the graph cardinality to remain the same. This also allows the network to learn correlations and explore the receptive-field across layers. The implementation of *DenseNet* idea in the final GNN-architecture is shown by the variety of skip-connections in Figure 8.1. The concatenation operation is depicted using the bluish rings in the figure. *InceptionNet* (details in Section 6.4.2), was useful in CNNs to efficiently learn structures at different scales. This was done by using parallel processing of input (/embedding), using kernels of different sizes. This allows

52 The rest of the GNN architecture shown in Figure 8.1, still use the edges constructed by kNN algorithm with "k" given by the Equation 8.1.

53 The *maximum* aggregation-scheme here can also be replaced with sum or mean aggregation.

CNNs to learn structures at multiple scales in parallel. The idea is adapted to our use case by doing the parallel processing of the Point-Transformer (details in Section 8.1.5) and Dynamic Edge-Convolution (details in Section 8.1.6) layers together. The edge network between the two will generally be different (because of dynamic edge construction by dynamic edge convolution). This opens up the possibility for the network to learn different representations from the same input representation, in parallel. This parallel-processed output, after every layer, is also concatenated with the output from the DenseNet-connections. Hence, the networking among layers is an amalgamation of DenseNet and InceptionNet. To test the potential of such networking on more famous and standard graph datasets like Open Graph Benchmark [442] is planned in the future.

8.1.8 *Global Pooling*

As discussed in Section 6.3, global-pooling operations are used to obtain a comprehensive graph-level representation of a graph. It allows to aggregate information of graphs with varying cardinality into a vector of fixed size and is permutationally invariant to the labeling of graph nodes. This work follows a flat-pooling method[54], where the sum and mean of the vector-embedding is performed (element-wise) [482, 626]. The global pooling is performed after every layer to benefit from the hierarchical representation learned in different layers[55]. These are labeled as "Global Pool" in Figure 8.1.

8.1.9 *Adding Physics Features*

In addition to the footprint-based (/graph-based) learning, it is judicious to benefit from the variety of composition-sensitive EAS-observables introduced in Chapter 7. The composition sensitive parameters used are: $\mathrm{Log}_{10}\left(\frac{\mathrm{dE}}{\mathrm{dX}_{1500\mathrm{m}}}/\frac{\mathrm{GeV}}{\mathrm{m}}\right)$ (Section 7.1), *Total Stochastic Energy* (Section 7.2), *Ratio Parameter* (Section 7.3) and Mean Radii and Mean Charge (Section 7.4) (with a maximum DOM-separation cut of 200 m from the track). In addition to these $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ (energy-proxy) and IT-based zenith and azimuth reconstructions (directional-info) are also used[56]. All these are labeled under the umbrella term "Physics Features" in Figure 8.1.

8.1.10 *Confluence of Representations*

The global feature representations learned using multiple varieties of convolutions described earlier are concatenated with the homophily measure (Section 8.1.3) and multiple EAS observables (Section 8.1.9). The final representation thus obtained benefits from both footprint information (low-level information) as well as reconstructed shower observables (high-level information). These are then used as an input to the MLP, which tries to predict the logarithmic mass of the primary i. e. ln(A) [57].

54 The work follows the nomenclature laid out in [625].
55 This also has inspirations in human behavior. Humans tend to depend on hierarchical abstractions in order to learn overall and fine-grained details about entities [627, 628].
56 **For IceCubers:** Laputop based reconstructions are used.
57 Everything is trained in an end-to-end fashion.

8.1.11 *Loss Function and other details*

Loss functions are used to quantify the performance/accuracy of a model. They are also essential for the back-propagation procedure needed to update the learnable weights during training (by taking partial-derivative of error with respect to the learnable weights). Various loss functions were tested. The final choice was *Smooth-L1 Loss*. It is given by:

$$\mathrm{Smooth-L1\ Loss} = \mathrm{mean}\left(\begin{cases} 0.5 \cdot (x_n - y_n)^2/\beta & \text{if } |\mathcal{Y}_n - y_n| < \beta \\ |\mathcal{Y}_n - y_n| - 0.5 \cdot \beta & \text{otherwise} \end{cases}\right), \tag{8.6}$$

where $\mathcal{Y}_n$ and y_n correspond to the true and predicted value of the n^{th} element in the training-batch. β was chosen to have a value of 0.3.

In addition, to the choice of architecture and loss function, the choice of *learning-rate* and *optimizer* is also a crucial choice to train NN-models. An initial learning-rate of 0.003 with a scheduler (reducing learning-rate once loss has stopped decreasing), with SGD-optimizer [629] (momentum=0.9) was determined to be the best choice.

In order to determine the performance of the model for composition estimation and CR primary-discrimination, *Normalized Overlap Area* (NOA) is used as an intermediate measure. It is the measure of overlapping area between two normalized distributions (read Section A.4 for details.). In our case, the normalized distributions will be the mass-prediction distributions in the test dataset, on various primary types. A lower NOA is better for easier separation between primaries.

8.1.12 *Mass Prediction*

For the purpose of training, the dataset shown in Figure 5.7 was divided into three parts, namely Training-dataset, Validation-dataset, and Test-dataset[58] in the proportion 70:10:20. Each primary-type (p, He, O and Fe) had an almost equal-contribution (i. e. $\approx$ 25%), in each energy-bin. The simulations used to train the model have SIBYLL 2.1 as the hadronic model. No additional weighting was used. The response of the GNN-based architecture (Figure 8.1), as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$, on the test dataset can be seen in Figure 8.8. Different colors (bottom panel of Figure 8.8) correspond to different primary types (p = Red, He = Orange, O = Green, and Fe = Blue). The dashed-dotted horizontal lines (bottom panel) correspond to the true mass (i. e. true ln(A)) of the primary types. The colored distributions are the KDE-fits to the underlying histograms. It is important to note that, for better visualization, the vertical height of a KDE has been scaled separately for each primary type. In an ideal-case, we would want the distributions corresponding to each primary type to only peak at their true mass, while having minimal overlap with other distributions[59]. The Top-Panel indicates the NOA for different primary-type combinations. A lower NOA is better for better discrimination between primary types. The best separation (minimum NOA) is seen for the p-Fe combination. This is expected since they lie on the opposite ends of the ln(A) prediction extremities. With increase in energy ($\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$), the

58 The terminology used is the same as [630], where Test-dataset refers to the unseen dataset and is not used during any process of training.
59 From EAS-physics we do know that because of shower-to-shower fluctuations we do expect an overlap.

overlap generally decreases. This is anticipated since at low energies we expect greater shower-to-shower fluctuations leading to greater overlap between the primaries. The separation power (using NOA) is an improvement over previously published work from the collaboration [20, 282, 631]. In order to obtain a spectrum it is also necessary that we have an energy estimate for each event. The following section will detail a method used to obtain energy estimate for each individual event.

8.2 ENERGY-ESTIMATE USING GRADIENT BOOSTED DECISION TREES

In order to obtain a composition-spectrum (/fraction-spectrum) it is also crucial to estimate the energy of the primary which initiated an EAS. As shown in Figure 5.6, $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ is a good energy-proxy as it is almost linearly correlated to the true-energy. Hence, $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ should in principle be sufficient to give a preliminary energy-estimate. The strength of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ is expected to vary with change in the zenith (to a significantly lesser level with change in azimuth), for air-showers initiated with primaries of the same energy. This is primarily because of the difference in atmosphere and the subsequent snow at IT, which particles have to propagate before depositing signal in IT-tanks. From Figure 5.6, we can also see a very slight composition bias. Hence, instead of only using $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ as the energy-proxy, for the final work and hereon, a Gradient-Boosted Decision Trees (GBDT) [632] was used to estimate the energy[60].

The following air-shower observables are used to train the GBDT:

- $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$: It is considered as a good-energy proxy. This can be seen in Figure 5.6. Read the description of Equation 5.7 for more details.
- Reconstructed Zenith and Azimuth[61]: Provides necessary directional information.
- $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$: Read Section 7.1 for details.
- Total Stochastic Energy : Read Section 7.2 for details.
- Ratio Parameter : Read Section 7.3 for details.
- Mean Radii (maximum DOM-radii = 200 m) : Read Section 7.4.1 for details. It is labeled as R200 in Figure 8.10.

The bias and resolution of energy prediction and the comparison with previous work from ICNO is depicted in Figure 8.9. The *Top-Panel* and *Bottom-Panel* respectively depict the bias and resolution of energy prediction as a function of predicted energy. The solid (and dashed) lines are results from this work, and circles represent previously published results [20]. We see a significant improvement in the energy bias for high-energy $\mathrm{Log}_{10}(\mathrm{Energy}_{\mathrm{Reco.}}/\mathrm{GeV}) \geqslant 8$, and comparable results at low-energies. The energy resolution is similar at low energies, with improvement at higher energies.

60 Multiple techniques were tested (simple linear fit, neural networks, Gradient Boosted Decision Trees etc) to estimate energy. Regression using GBDTs provides the best bias and resolution over the whole energy range and is also competitive with the previously published results [20].

61 **For IceCubers:** Laputop based reconstructions are used.

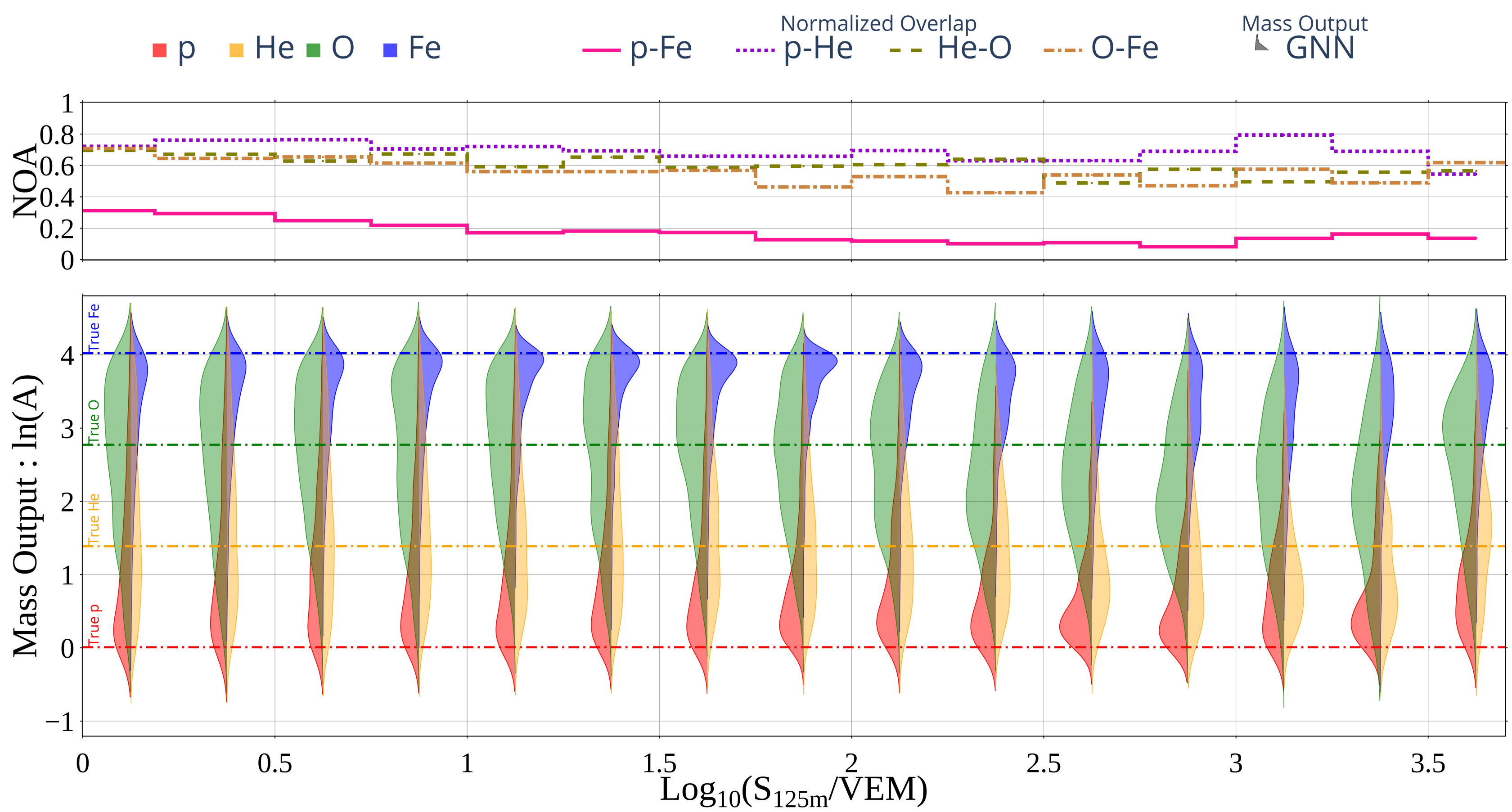


Figure 8.8: Response of GNN-based architecture (Figure 8.1) to various primary-types on the test-dataset, as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$, is shown in the *Bottom Panel*. The NOA (read Section A.4 for details) is also shown in *Top Panel*. The underlying simulations use SIBYLL 2.1 as the hadronic interaction model, and the events are IT-IC coincident events and pass the IT-IC quality cuts detailed in Section 5.3.3. Read Section 8.1.12 for details.

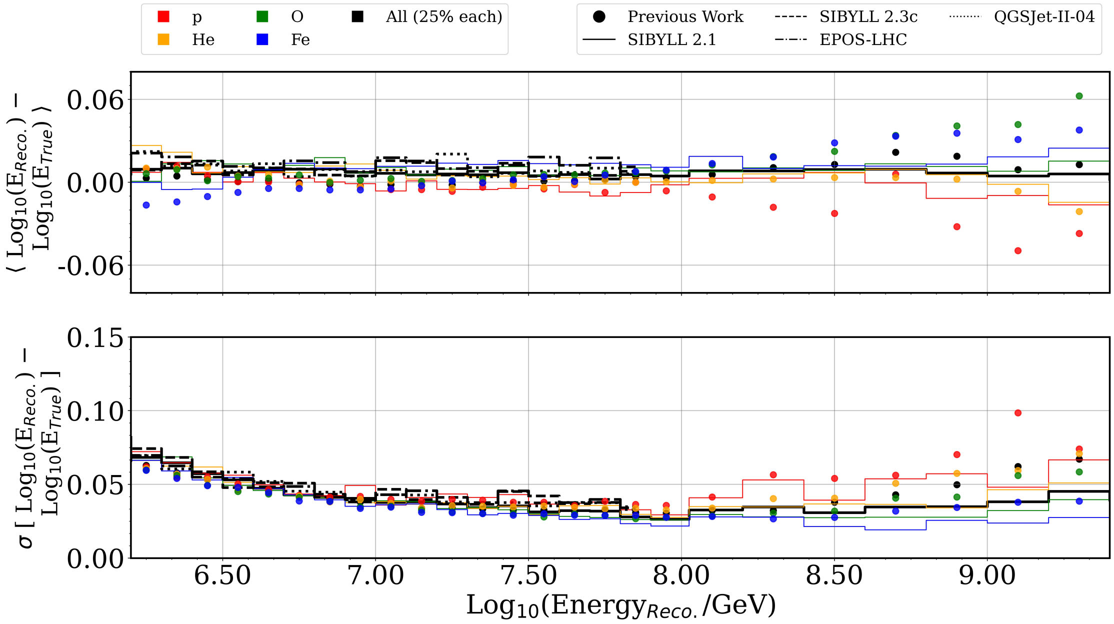


Figure 8.9: The bias (upper-panel) and resolution (lower-panel) of energy-prediction as a function of predicted energy. The individual primaries and an equal fraction for each primary type are shown. The response for different hadronic-model is also shown. The comparison with previously published work [20] is also shown.

Using GBDTs[62] also allows us to estimate the feature importance of the input features. Feature importance indicates the relative importance of the input features and hence indicates the order in which the features contribute to the predictive power of the model. It can help add useful interpretation to the final model predictions [633] (or select features) and improve generalization [634]. Figure 8.10 shows feature-importance of energy-prediction. Feature weights are used to quantify the importance. Here weight denotes the number of times a feature appears in a tree. As expected, $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ has the highest feature-importance. Reconstructed zenith and azimuth have the second and third highest feature importance respectively. This is also expected, since variation in zenith and azimuth are expected to create the greatest difference in reconstructed $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$, for air-showers with the same primary energy. Composition-sensitive parameters, as expected, are much less important for energy prediction.

8.3 MASS PREDICTION ON MEASURED DATA

A mass-estimate (Section 8.1) and an energy-estimate (Section 8.2), equips us with the ability to make predictions on MC as well as burnsample. This is shown in Figure 8.11. The bottom panel shows the KDE-expectations for the MC (p = Red, He = Orange, O = Green, and Fe = Blue) and burnsample (black). The x-axis has now been replaced with energy-prediction from the GBDT (details in Section 8.2). It is important to notice that the vertical height of the KDEs have been scaled separately for each primary type and

62 or any other tree-based method

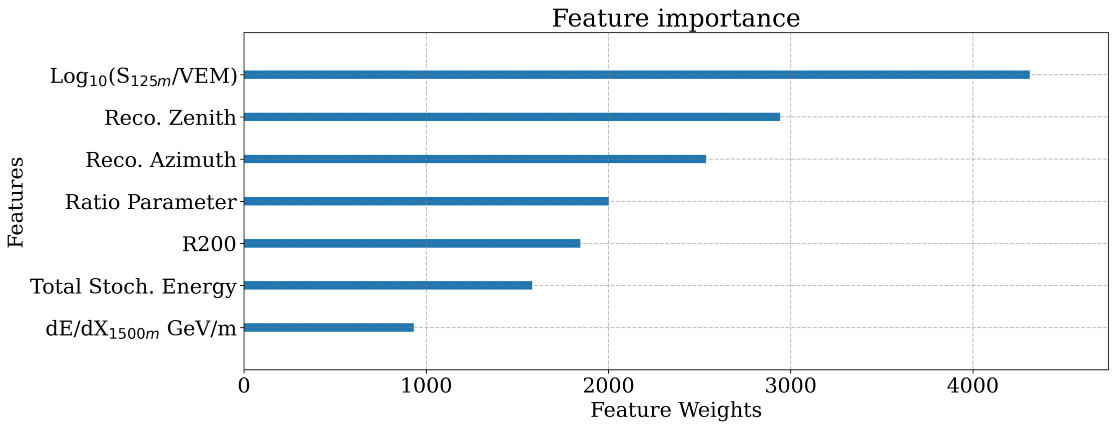


Figure 8.10: Feature importance for energy-prediction, for coincident IT-IC events. Read the text for details of the features.

burnsample, for better visualization. Similar to Figure 8.11, the top panel shows the NOA and similar conclusions (about NOA) as earlier can be drawn.

On comparing the burnsample-distribution with MC-distributions in Figure 8.11, the burnsample seems to indicate a mixed-composition at lower-energies and a Fe-like composition at the higher-end. The next section will introduce a likelihood-based approach to reconstruct the individual fractions of primary-types in the burnsample. The events in the burnsample are statistically limited above reconstructed energy i. e. $\mathrm{Log}_{10}(E_{\mathrm{Reco.}}/\mathrm{GeV})$ of about 8.5.

8.4 TEMPLATE FITTING : METHOD AND TESTING

Figure 8.11 presents the mass-expectation of the burnsample as a function of reconstructed energy. The figure seems to indicate that at high energies, the burnsample is mostly Fe-like. However, drawing any quantitative conclusions become difficult at lower energies. Also, even though the overlap between p and Fe is very less, we see a significant overlap between other primaries. In order to make stronger statements about the underlying distributions in burnsample, a *Template Fitting* method is used. The template fitting method will be used to unfold the burnsample-distribution as a combination of the contributions from the four primaries. Hence, it can also allow us to get the fractional contributions of each primary type. There are several methods and their-variations [635–640] to perform template fitting.

This work uses a similar *Template Fitting* method as the previous relevant publication from ICoIl [20]. It uses templates created from individual primaries (using MC-simulations), using an adaptive KDE method [641][63]. Because of a limited number of events in the test dataset, to get stable templates the energy-bin size was increased to 0.2 (from 0.1 in [20]) till $\mathrm{Log}_{10}(E_{\mathrm{Reco.}}) < 8$ and to 0.3 (from 0.2 in [20]) after $\mathrm{Log}_{10}(E_{\mathrm{Reco.}}) \geqslant 8$. The templates are then used to fit the binned data. The template fitting method can be expressed as finding solutions for:

$$\mathrm{Data} = N_p \cdot \mathrm{KDE}_p + N_{He} \cdot \mathrm{KDE}_{He} + N_O \cdot \mathrm{KDE}_O + N_{Fe} \cdot \mathrm{KDE}_{Fe} \tag{8.7}$$

63 The templates are created using the improved Sheather-Jones (ISJ) algorithm [641], implemented in zfit [642].

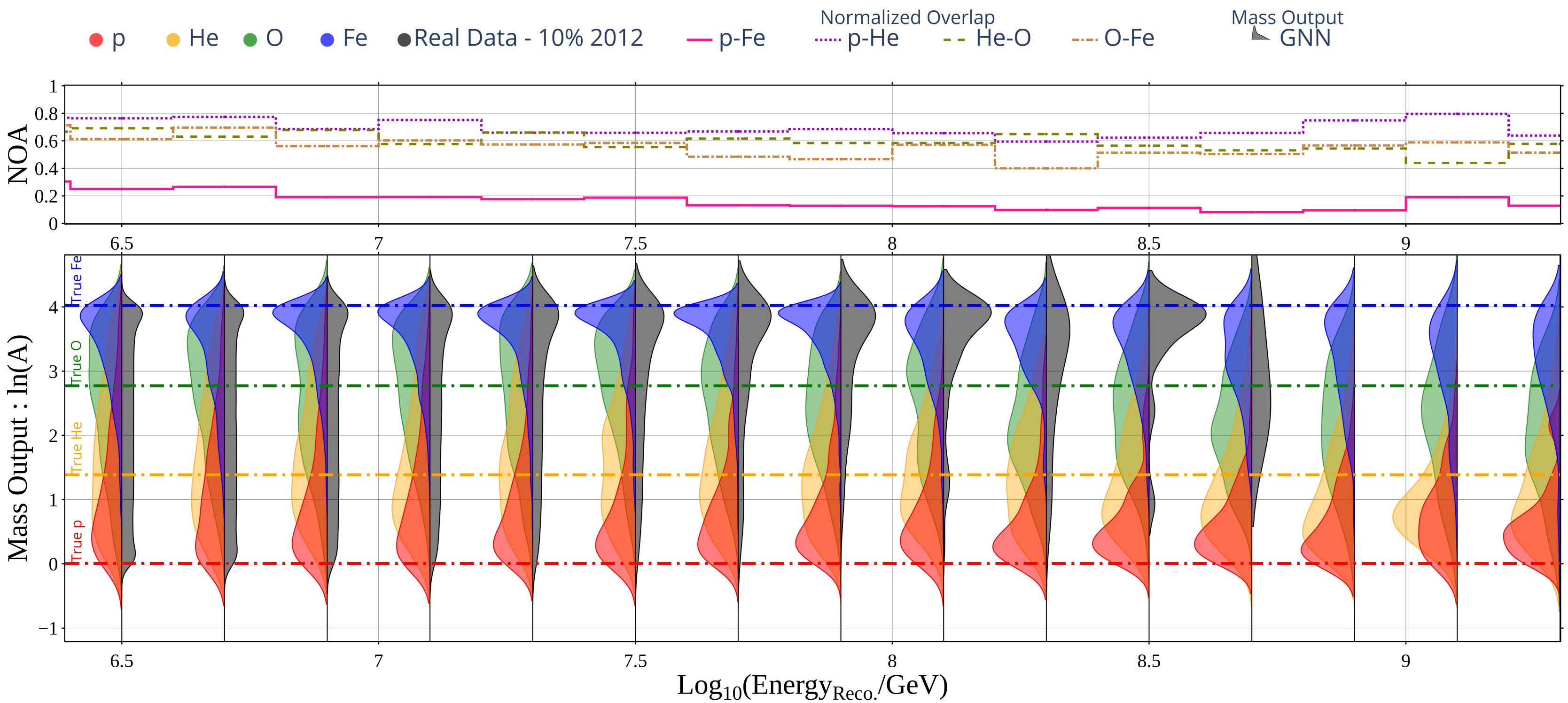


Figure 8.11: Response of GNN-based architecture to MC and burnsample, as a function of reconstructed energy (using Section 8.2). The KDEs for MC and burnsample are shown on opposite sides of each bin (Bottom Panel). The NOA (read Section A.4 for details) is also shown (Top Panel). The underlying MC-simulations use SIBYLL 2.1 hadronic-model, and real-data is 10% of 2012's real-data (burnsample). The events pass the IT-IC quality cuts detailed in Section 5.3.3.

where, $KDE_{p/He/O/Fe}$ are the template-KDEs obtained using the MC-simulations[64], and $N_{p/He/O/Fe}$ are the free-parameters of the extended likelihood-minimization procedure (and represent the reconstructed event-number for each primary-type). The minimization is performed using *Iminuit* [643][65], with Hesse-minimizer [637]. An equal fraction of each primary is given as the seed for the minimization i. e. $N^{initial}_{p/He/O/Fe} = N_{Data}/4$. In addition to that $N_{p/He/O/Fe}$ is limited to lie between 0 and N_{Data}. In addition to providing the best-fit values, *Iminuit* also estimates the error from likelihood profile analysis. This is crucial to provide confidence intervals in our final prediction. The extended likelihood-minimization process provides an estimate of the following:

$$N_p, N_{He}, N_O, N_{Fe} \text{ and Covariance Matrix} = \begin{pmatrix} Var(x_1) & \dots & \dots & Cov(x_1, x_4) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ Cov(x_4, x_1) & \dots & \dots & Var(x_4) \end{pmatrix}$$

Here, $Var(x_i)$ represents the variance of the i^{th}-distribution and $Cov(x_i, x_j)$ is the covariance of the i^{th}-distribution with respect to the j^{th}-distribution. The diagonal elements of the covariance matrix provide an estimate of variances of N_p, N_{He}, N_O, N_{Fe}. In order to test the performance of the method, the mean logarithmic mass (hereon referred to as $\langle ln(A) \rangle$ or simply mean-mass) is calculated. It is given by:

$$\langle ln(A) \rangle = f_p \cdot ln(A_p) + f_{He} \cdot ln(A_{He}) + f_O \cdot ln(A_O) + f_{Fe} \cdot ln(A_{Fe}) \tag{8.8}$$

where $\langle ln(A) \rangle$ represent the mean logarithmic mass, and $f_\alpha = \frac{N_\alpha}{N_p + N_{He} + N_O + N_{Fe}}$ is the fractional-contribution from primary-type α(=p/He/O/Fe) with logarithmic-mass $ln(A_\alpha)$ ($ln(A_p) = 0.007$, $ln(A_{He}) = 1.386$, $ln(A_O) = 2.772$ and $ln(A_{He}) = 4.022$). To calculate error on the $\langle ln(A) \rangle$ (and the fractions), error-propagation using the covariance matrix is performed.

Before, testing the template fitting method on burnsample, it is crucial to ensure that the method is unbiased to a mixture of custom composition fractions. Five different cases were checked[66], namely Pure-p, Pure-He, Pure-O, Pure-Fe, and a periodic variation between p and Fe. This is depicted in Figure 8.12. Top Left = Pure p, Top Right = Pure He, Middle Left = Pure O, Middle Right = Pure Fe and Bottom = Sinusoidal variation between p and Fe. Section A.6 presents the underlying template-fits done for each plot (in energy-bins) in Figure 8.12. Figure 8.12 depicts that the template fitting procedure is able to reconstruct the $\langle ln(A) \rangle$ successfully for a variety of underlying compositions. The figures in Section A.6 depict that in spite of having limited events for KDE template creation, the reconstructed templates are able to fit the test-data very well[67]. Chapter 9 will present the fractions and $\langle ln(A) \rangle$ obtained from the template-fitting procedure on the burnsample and the associated physics-interpretations.

64 For our case, these are the KDEs generated using the test-dataset, binned in reconstructed-energy.
65 It allows likelihood and least-squares fits of parametric models to data.
66 The underlying MC-simulations use SIBYLL 2.1 hadronic-model.
67 More simulations for template creation can help improve the method even further. Creating more MC-simulations for CR-analysis at ICNO is an ongoing effort.

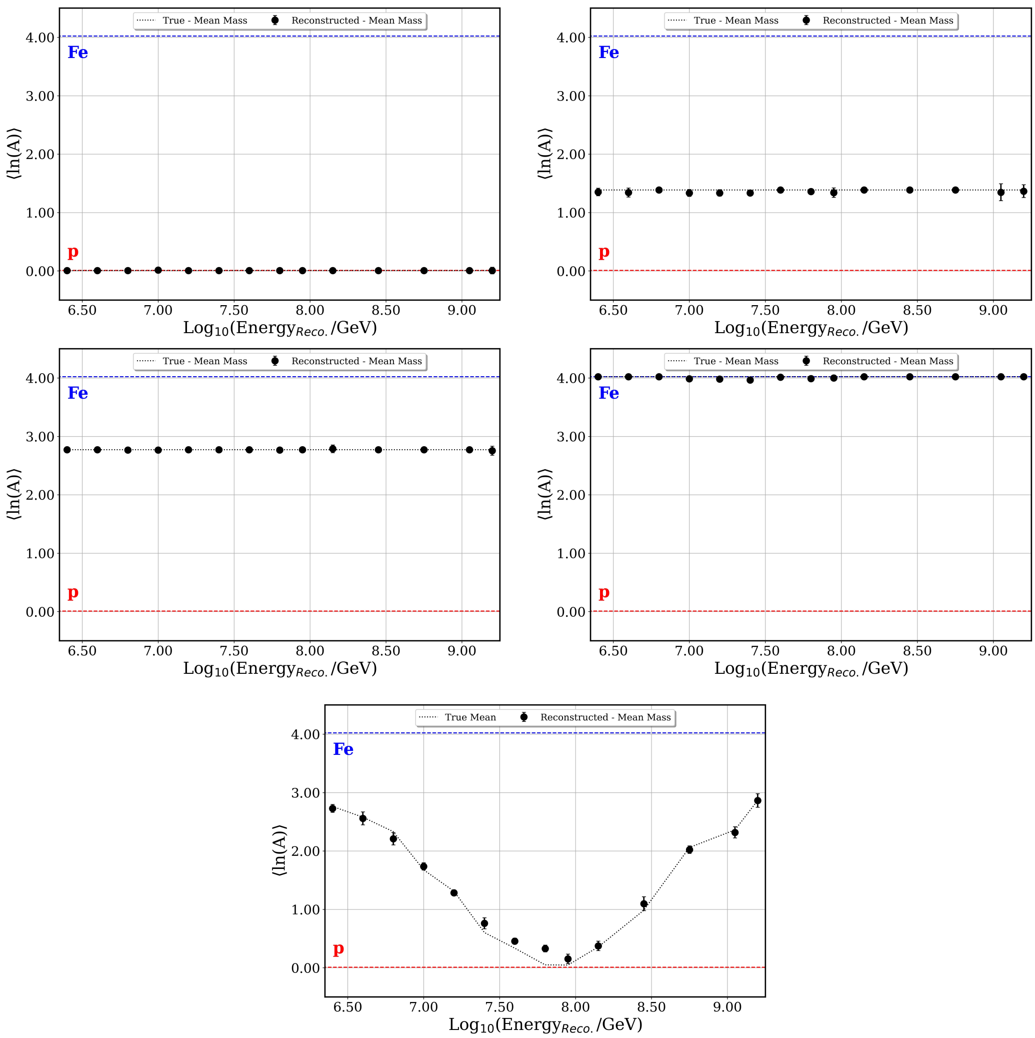


Figure 8.12: Reconstructed $\langle \ln(A) \rangle$ for the test-case of different underlying composition fractions, generated using a random sample of the test-dataset. The test dataset was divided into two equal parts randomly. One was used to generate KDE templates and the other was used to create custom fractions to be tested. The reconstruction for all of them is done using template KDEs for the four primaries. Top Left = Pure p, Top Right = Pure He, Middle Left = Pure O, Middle Right = Pure Fe and Bottom = Sinusoidal variation between p and Fe. The dotted line represents the true-mean mass in the energy bins and the dots represent the reconstructed mean mass.

9 RESULTS

karmaṇy-evādhikāras te mā phaleṣhu kadāchana
mā karma-phala-hetur bhūr mā te saṅgo 'stvakarmaṇi
(You have a right to perform your prescribed duties, but you are not entitled to the fruits of your actions.
Never consider yourself to be the cause of the results of your activities, nor be attached to inaction.)

— Bhagavad Gita: Chapter 2, Verse 47 [644]

Chapter 6 introduced the idea of using Graph Neural Networks (GNNs) at ICNO. Chapter 7 focused on developing multiple CR composition-sensitive parameters and describing inconsistencies observed in composition-expectation among those. Chapter 8 combined the GNNs introduced in Chapter 6, with composition sensitive EAS-observables discussed in Chapter 7. Hence, a method (a GNN-architecture, as shown in Figure 8.1) benefiting from both low-level and high-level information in EASs is developed. Section 8.4 introduced a method to get contributions or fractions (for different primary types) from the mass-predictions from GNN[1]. This chapter will present the results obtained from the Template-fitting procedure using ln(A) prediction discussed in Chapter 8. Figure 9.1 depicts the raw energy-spectrum[2] for the events (i. e. the burnsample) used for the upcoming discussions. The underlying events used for the upcoming plots use the quality cuts discussed in Section 5.3.3.

9.1 PREDICTING COMPOSITION

In order to obtain the fractional contribution of primary types in real data (burnsample) the KDE-templates for the Template-Fitting method (discussed in Section 8.4) are created, with SIBYLL 2.1 as the hadronic-model. The following text will discuss the fractional contributions obtained from such analysis, and the respective $\langle \ln(A) \rangle$. For an energy greater than $\mathrm{Log}_{10}(E_{\mathrm{Reco.}}/\mathrm{GeV}) \gtrsim 6.3$, the effective-area for this analysis becomes composition-independent [410]. Hence, only events with reconstructed-energy i. e. $\mathrm{Log}_{10}(E_{\mathrm{Reco.}}/\mathrm{GeV}) \gtrsim 6.3$ will be used.

The GNN-based architecture introduced in Chapter 8 was trained and tuned to get the maximal-possible separation between the primary types. The work introduced therein used the full shower footprint, along with high-level composition-sensitive features (see Figure 8.1). The IT-tanks (and IC DOMs) were mapped as the nodes of the

1 It is however important to notice that the Template-fitting method introduced in Equation 8.7 can be used for templates created using mass-prediction via GNNs as well as for any shower-observable. The "Data" in Equation 8.7 represents ln(A) prediction on burnsample with GNN, if KDE-templates are also generated using ln(A) predictions for simulations. In case the KDE-templates are generated using shower-observables, the "Data" represents the distribution of the particular shower-observable.

2 For details of energy-prediction, read Section 8.2.

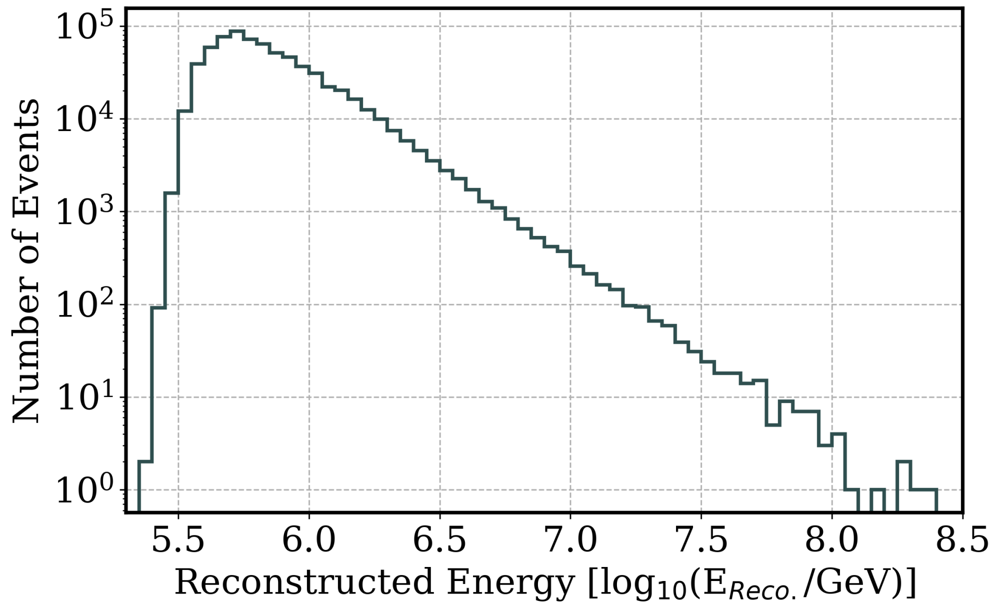


Figure 9.1: Raw energy-spectrum using 10% of 2012's real-data (burnsample), as a function of reconstructed-energy.

graph (read Section 8.1.1 for details.) and the edges were constructed based on a policy which focused on introducing inductive biases (read Section 8.1.2 for details). In addition to that reconstructed high-level EAS observables were also used (read Section 8.1.9 for details). The logarithmic-mass prediction from the network (for each event) was then used to perform the Template-fitting procedure introduced in Section 8.4. Figure A.11 shows the template-fits obtained with burnsample. The fractional contributions for various primary types obtained from this (using SIBYLL 2.1 as the hadronic model to create the KDE-templates for Template-fitting), are presented in Figure 9.2. The reconstructed fraction for proton shows a maxima around $\text{Log}_{10}(E_{Reco.}/\text{GeV}) \approx 6.5$. With increasing energies, the proton fraction drops to almost zero at the highest energies in the burnsample. For Helium, the maxima occurs at approximately $\text{Log}_{10}(E_{Reco.}/\text{GeV}) \approx 7$. Similar to a proton, with increasing energies the fractional contribution for Helium drops, with increasing energies. At bump at the highest energies is observed for Helium. It is currently unclear if it is because of limited statistics of the burnsample, or if there is any other underlying cause. For oxygen, the trend is unique. It shows multiple bumps in fractional contribution as a function of reconstructed energy. For Iron, the fractional contribution increases with an increase in energy. On comparing the contributions with the various flux models (discussed in Chapter 2), no single model fully explains the obtained fractions fully. However, the Iron fractions seem to favor GST-3gen over others [91]. On comparing this work with fractions obtained from previous work [20], shown in Figure A.12, a significant increase in Fe-expectation is observed.

Figure 9.3 shows the corresponding $\langle \ln(A) \rangle$ as a function of reconstructed energy. It also shows the comparison of the $\langle \ln(A) \rangle$ obtained from this work with the one from previous publication [20]. It is interesting to note that this work predicts a heavier composition than [20], for the same choice of hadronic model (i. e. SIBYLL 2.1). This can be seen from Figure 9.3, where this work (labeled as "This Work: SIBYLL 2.1") shows a greater $\langle \ln(A) \rangle$ than the previous work (labeled as "IC-2019: SIBYLL 2.1"),

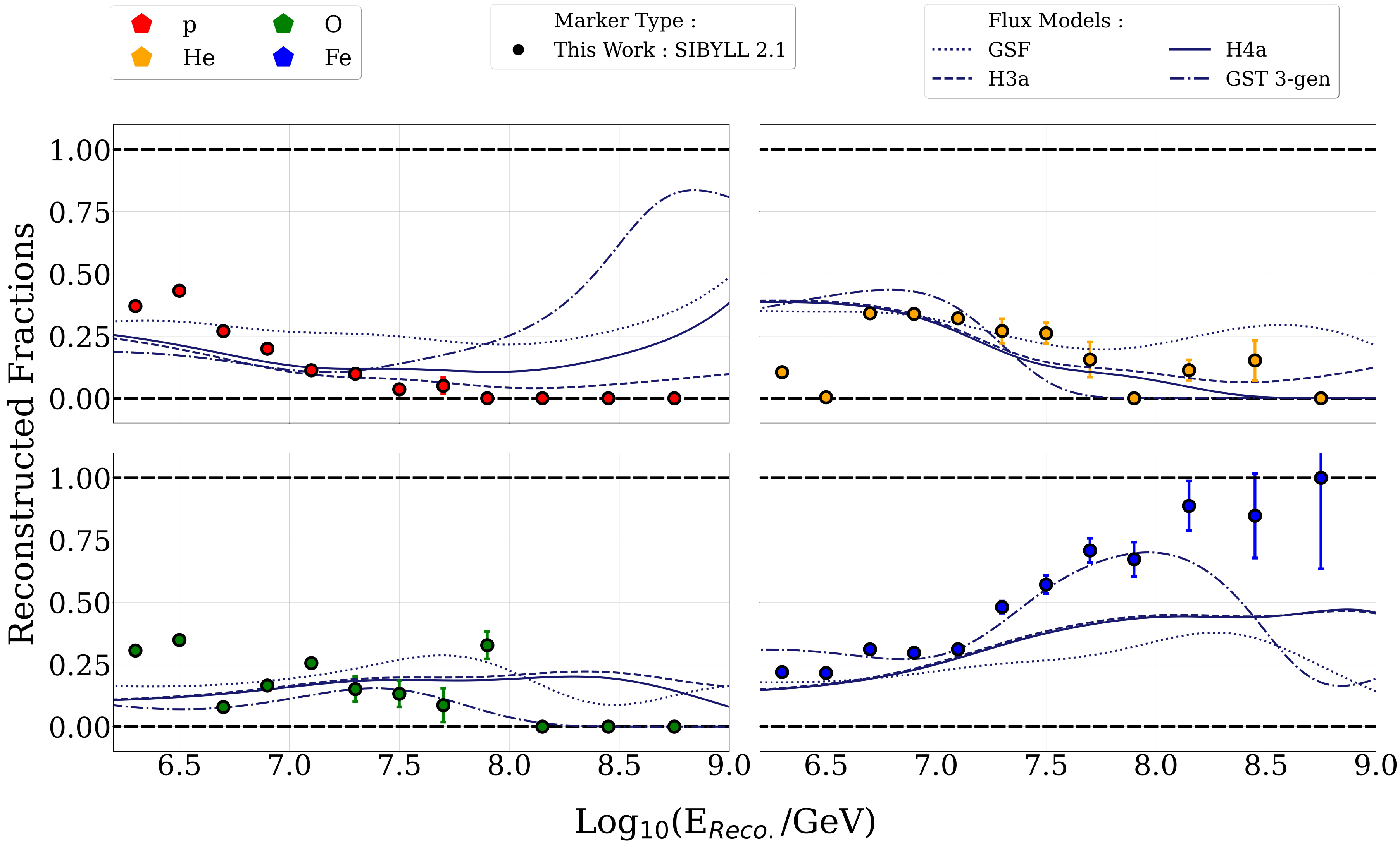


Figure 9.2: Reconstructed proton, helium, oxygen, and iron fractions in the burnsample (using ln(A) predictions from GNN-based architecture - Read Chapter 8 for details), as a function of reconstructed energy. The fractions are obtained by the Template-fitting method (discussed in Section 8.4). The template fits are shown in Figure A.11.

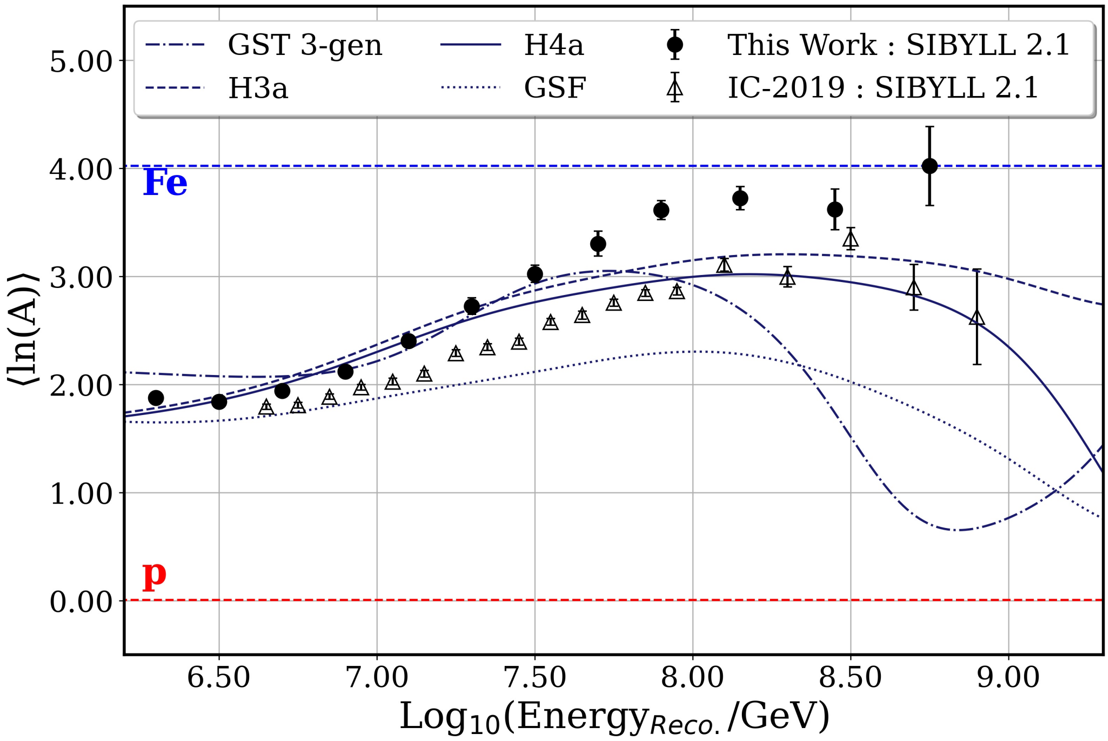


Figure 9.3: Mean logarithmic mass expectation from this work (using GNN predictions) and the comparison with previous ICoII publication [20], as a function of reconstructed energy. The dashed and dotted curves show the the mean logarithmic mass expectation for the choice of different flux-models and results from ICNO. The flux models have been discussed in Chapter 2.

for all energies. This can also be expected from Figure A.12, where this work predicts more contribution from Fe (in comparison to the previous work), with an increase in energy[3]. There can be multiple reasons/possible solutions which can explain the seen difference, namely:

- **Detector Uncertainties:** The MC-simulations used for this analysis (Chapter 7, Chapter 8, Chapter 9) used a fixed detector configuration[4]. However, it is already known that our detector knowledge is not perfect and a conservative estimate of the detector-systematics is also known [20]. The inclusion of the detector-systematics (and its effect on the uncertainty in mass prediction) for this work is planned for the future. Hence, the possibility of the two results being compatible on the inclusion of detector systematics remains to be seen[5].

- **Dataset Volume:** The previous work [20], used three year (from June 1, 2010, through May 2, 2013) of ICNO data to obtain the $\langle \ln(A) \rangle$ presented in Figure 9.3. In contrast, this work only uses 10% of data from the operation-year 2012 (spread throughout the year). Hence, this work only uses about 3% of the data used for the previous work. It remains to be seen if the shift is because of any systematic shifts in the limited dataset. It also remains to be seen if the sudden bumps seen in Oxygen contribution (in Figure 9.2) are also a consequence of this.

- **EAS Physics:** The previous work [20], was based on composition estimation using a MLP. The composition-sensitive parameters used for training the MLP were $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ (discussed in Section 7.1) and stochastic counting (discussed in Section 7.2). These parameters are generally only dependent on TeV muon multiplicity. Hence, the previous work was reliant on how well the MC-simulations are able to replicate the TeV muon-multiplicity. Any systematic bias or difference between MC-simulations and real data will bias/affect the training procedure (of the MLP), as well as the fractions obtained from the Template-fitting procedure. It is already known from observation at other observatories that a choice of different air-shower observables/detectors can lead to different composition expectations [645]. In contrast, this work uses shower-observables that capture different aspects of EAS-physics, like muon-multiplicity (Section 7.1, Section 7.2, Section 7.3), lateral-spread (Section 7.4.1), footprint-info (Section 8.1.1). Since the training as well as Template-fitting is now reliant on far more EAS-observables, it is much more likely that the final prediction will be less biased in a preferential direction, because of the choice of input made. A conclusive evidence of this is planned for a future work. Section 9.2 will present a compatibility check, by using the fractional contributions obtained from this work to check the overlap of EAS-observables with real-data.

It can be seen from the fractional-contribution plot (Figure 9.2) and the $\langle \ln(A) \rangle$ plot (Figure 9.3), that this work doesn't fully agree with one flux-model over the entire

3 As can be seen from Figure A.12, the additional contribution seems to come at the cost of reduced contribution(in comparison to the previous work) of the lighter primaries.
4 i. e. fixed values IT snow-correction (Equation 5.9), light-yield in IC etc.
5 The effect of detector-systematics on the previous work [20] is not shown in Figure 9.3. On a quick glance, on including the systematics from the previous work (Figure 17. in [20]), the two results are compatible for most energy-bins.

energy range considered. For the low-energies, the $\langle \ln(A) \rangle$ from this work favors H3a/H4a [90]. For $7 < \mathrm{Log}_{10}(\mathrm{E}_{\mathrm{Reco.}}/\mathrm{GeV}) \lessapprox 7.7$, the work seems to prefer GST 3-gen flux-model [91] more than the others. At higher energies the work seems to be in tension with any of the flux models considered. A bigger dataset than the burnsample, detailed MC-simulation test and a detailed detector-systematics study can help make a stronger statement in the future.

Figure A.13 presents the $\langle \ln(A) \rangle$ expectation for the choice of different hadronic-interaction models and compares it with results presented in Figure 9.3. The templates used for other hadronic models were generated using the same GNN model as before (trained on SIBYLL 2.1). This was done to make a comparison of different hadronic-models easier[6] with the data, as well as because of limited simulations availability for other hadronic models. The other hadronic models too predict a $\langle \ln(A) \rangle$ which is generally heavier (For SIBYLL 2.3 and EPOS-LHC - Always; For QGSJet-II-04 - For $\mathrm{Log}_{10}(\mathrm{E}_{\mathrm{Reco.}}/\mathrm{GeV}) \gtrapprox 7.3$) than the previous publication [20]. It is however important to notice that since the datasets for other hadronic models currently lack available simulations for Helium and Oxygen, the Template-fitting procedure tends to overestimate the contribution from proton and Iron. This is clearly visible from the corresponding template fits depicted in Figure A.14 (SIBYLL 2.3), Figure A.15 (EPOS-LHC) and Figure A.16 (QGSJet-II-04). Creating more MC-simulations at ICoII is an ongoing effort. Simulations from other primary types can hence be used to update mean mass expectations for other hadronic-models in future. Hence, from hereon only a comparison with SIBYLL 2.1 will be discussed.

Figure 9.4 presents the comparison of $\langle \ln(A) \rangle$ expectation from this work (using GNN predictions) and the previous ICoII publication [20] with results of $\langle \ln(A) \rangle$ expectation from KASCADE [33], for the choice of SIBYLL 2.1 as the hadronic-interaction model[7]. Comparing results with other experiments gives us the possibility to explore the possibility of approaching the true CR-composition as well as understand ICNO better. As can be seen from Figure 9.4, the results from KASCADE indicate an even heavier composition than this work or [20]. It is currently unclear if the difference is because of the difference in shower-observables[8], difference in observatory location(ICNO - Southern Hemisphere, KASCADE - Northern Hemisphere), or any other observatory-specific effects. It is however crucial to notice that this work indicates a much closer composition (in comparison to the previous work) to that from KASCADE. Section 9.2 will introduce an independent method as a sanity check to test the compatibility of results obtained from this work with real-data (burnsample).

6 Currently since burnsample predictions are only from a single GNN-model (trained on SIBYLL 2.1), any difference in KDE-templates or mean-mass expectation should primarily come from the difference in the inherent physics among the hadronic interaction models. In case different GNN-models are trained, the mean-mass expectation for burnsample will differ for different GNN-models. This makes disentangling the difference in mean-mass expectation because of different GNN-models and/or because of inherent differences in hadronic interaction models difficult.

7 Figure A.17 presents the $\langle \ln(A) \rangle$ comparison between this work, previous ICoII-publication [20], KASCADE [33], and Tunka [646] for different hadronic-models.

8 KASCADE was a surface-detector and was dependent on deposits from GeV-muons for composition analysis. In contrast, this work uses information from both GeV(at IT) as well as TeV muons(at IC). The previous work [20] was primarily dependent on TeV muons for composition sensitivity.

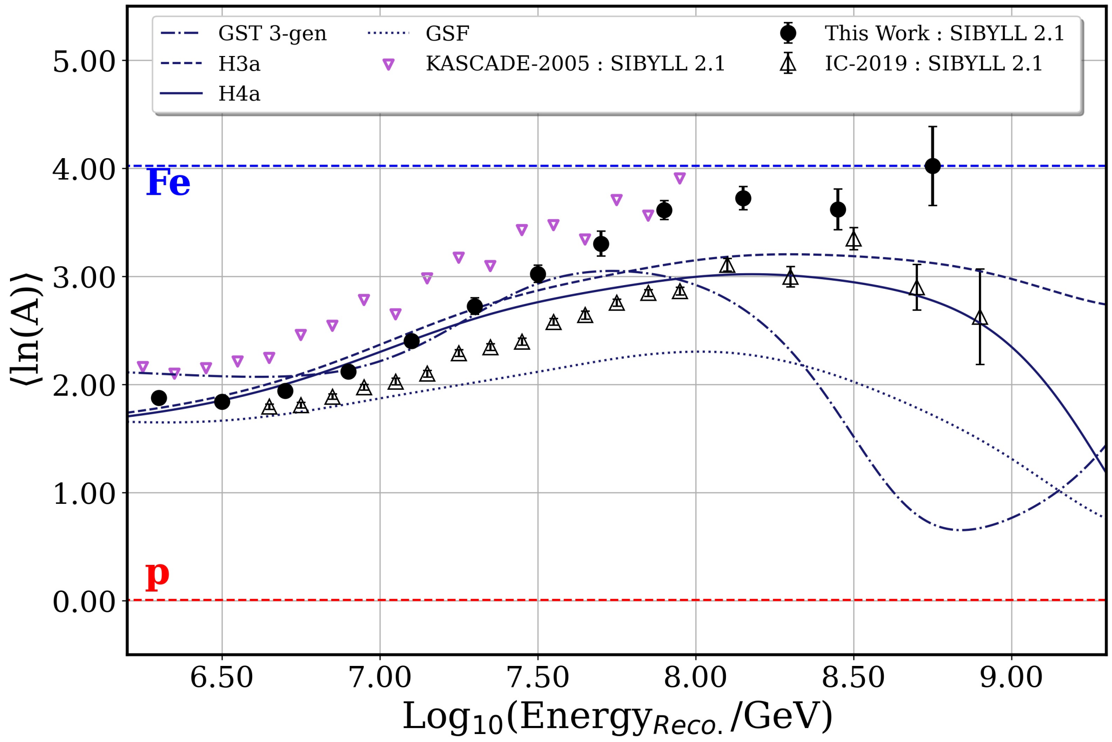


Figure 9.4: Comparison of $\langle \ln(A) \rangle$ expectation from this work (using GNN predictions) and the previous ICoIl publication [20] with results of $\langle \ln(A) \rangle$ from KASCADE [33] (error-bars not included) for SIBYLL 2.1, as a function of reconstructed energy. The data for KASCADE was accessed from KCDC (KASCADE Cosmic-ray Data Centre) [50].

9.2 COMPOSITION : CONSISTENCY CHECK

For a reliable composition estimation, the fractions obtained from a method should be able to reproduce distributions of EAS-observables from real data. In order to test this the steps indicated in Figure 9.5 are followed. The details of the steps are:

- **Step 1:** Choose a EAS-observable to check the compatibility of fractional contributions obtained from this work. $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}} / \frac{\mathrm{GeV}}{m}\right)$ (discussed in Section 7.1) and Mean-Radii (with a maximum DOM-radii cut of 200 m - discussed in Section 7.4.1) are among the choices for the EAS-observables[9].
- **Step 2:** Generate the KDE-templates for the shower-observable (in energy-bins), for real-data (burnsample) and the different primary-types[10].
- **Step 3:** Weight the simulation-KDEs with the primary-type fractions obtained from the method (Section 9.1 and Section A.8 here) to be tested for compatibility with real-data.

9 As discussed before, the choice was made because the two are among the representative examples (among the composition-sensitive observables discussed in Chapter 7) of composition-sensitive shower-observables depending on muon-multiplicity and their lateral-extent.

10 SIBYLL 2.1 - p, He, O and Fe; Others - p and Fe

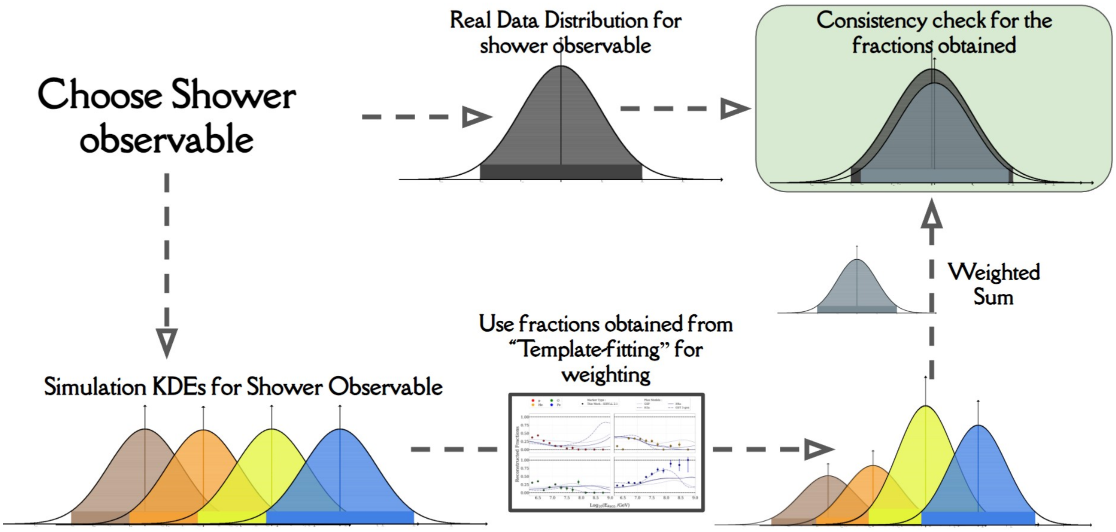


Figure 9.5: Outline of steps carried out for compatibility-check of primary-type fractions obtained from this work, as well as for comparison with other works.

- **Step 4:** Compare the weighted-sum[11] of simulated KDEs with the KDE generated from real-data (in Step 2). Calculate p-value using two-sample (burnsample-KDE and Weighted-sum KDE) Kolmogorov–Smirnov (KS)-test [647], with the null-hypothesis that the two distributions are sampled from the same underlying distribution. A p-value close to 1 indicates evidence in favor of null hypothesis.

- **Step 5:** Repeat Step 1 to 4 for other shower-observables.

Figure 9.6 and Figure 9.7 presents the results from the consistency check described earlier for $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}} / \frac{\mathrm{GeV}}{m}\right)$ (discussed in Section 7.1) and Mean-Radii (with a maximum DOM-radii cut of 200 m - discussed in Section 7.4.1) respectively. The x-axis in the plots represents the value of EAS-observable in energy bins and the y-axis represents the probability-density. The y-axis is linear for the plots. For Log-scale counterparts view Figure A.21 and Figure A.22. The solid black-line in the figures represents the real-data (burnsample) distribution for the EAS-observable. The solid magenta-line and the magenta-band represent the weighted-KDE (using fractions estimates from Section 9.1), and the corresponding propagated-error, respectively. The other three-lines represent the weighted-KDEs using fractions from H4a [90], GST 3-gen [91] and GSF [92] flux model. The number of events in the burnsample for each energy-bin is also indicated in the plot.

Figure 9.6 and Figure 9.7 indicate that the fractions obtained from this work using GNN-estimates are generally compatible with EAS-distributions of real-data(burnsample), for both of the shower-observables. The figures also suggest that the overlap with real data decreases with an increase in energy (the reduced overlap is more prominent for $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}} / \frac{\mathrm{GeV}}{m}\right)$ than Mean-Radii). This might be because of reduced statistics at higher energies or because of the different effect of detector-systematics for higher

11 with fractions obtained from Template-fractions

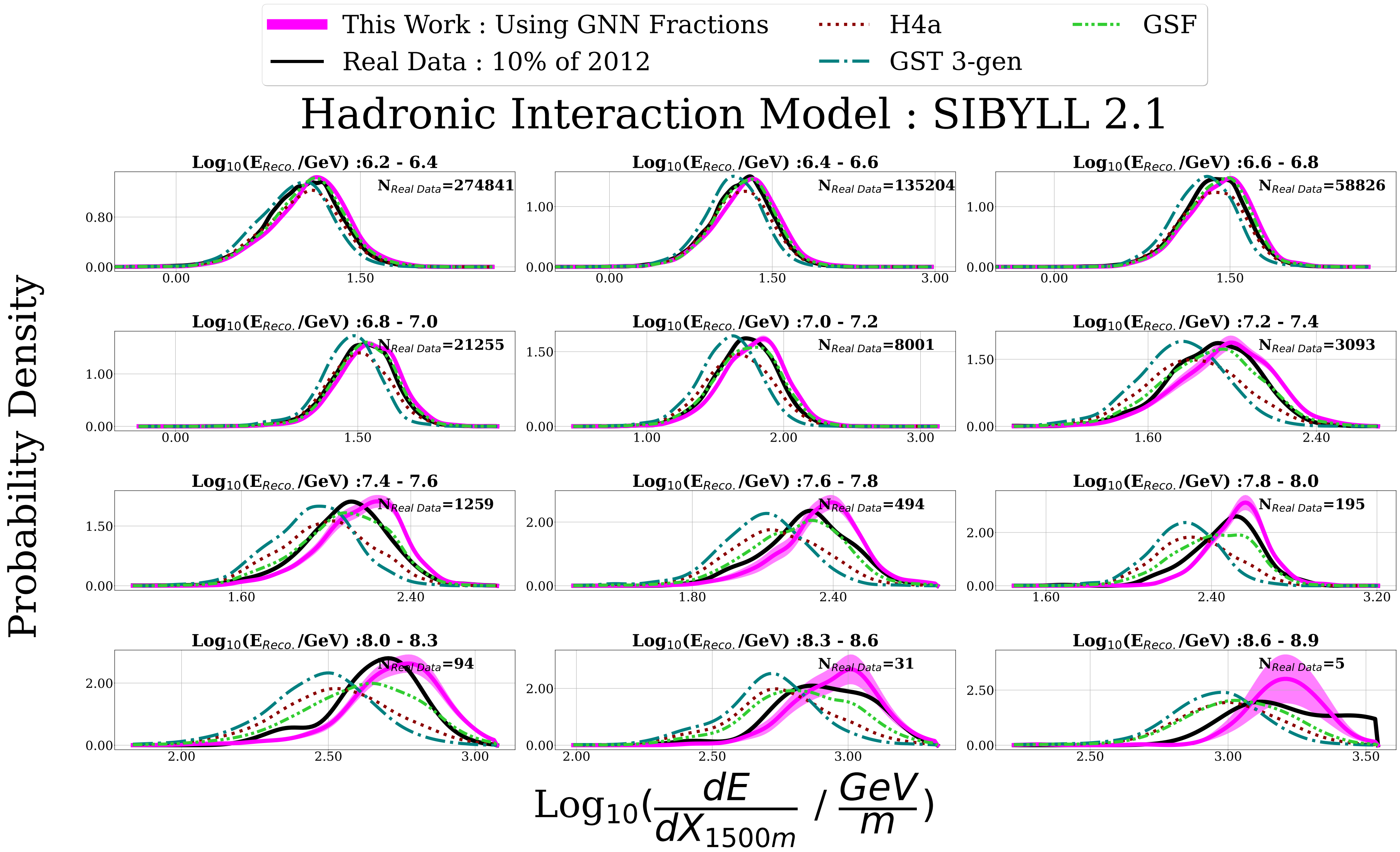


Figure 9.6: Compatibility check for composition-fractions obtained from GNN (details in Section 9.1) for $\text{Log}_{10}\left(\frac{\text{dE}}{\text{dX}_{1500\text{m}}}/\frac{\text{GeV}}{\text{m}}\right)$ and comparison with fractions from other flux-models, on a *Linear-scale*. To see the plot on *Log-scale*, see Figure A.21.

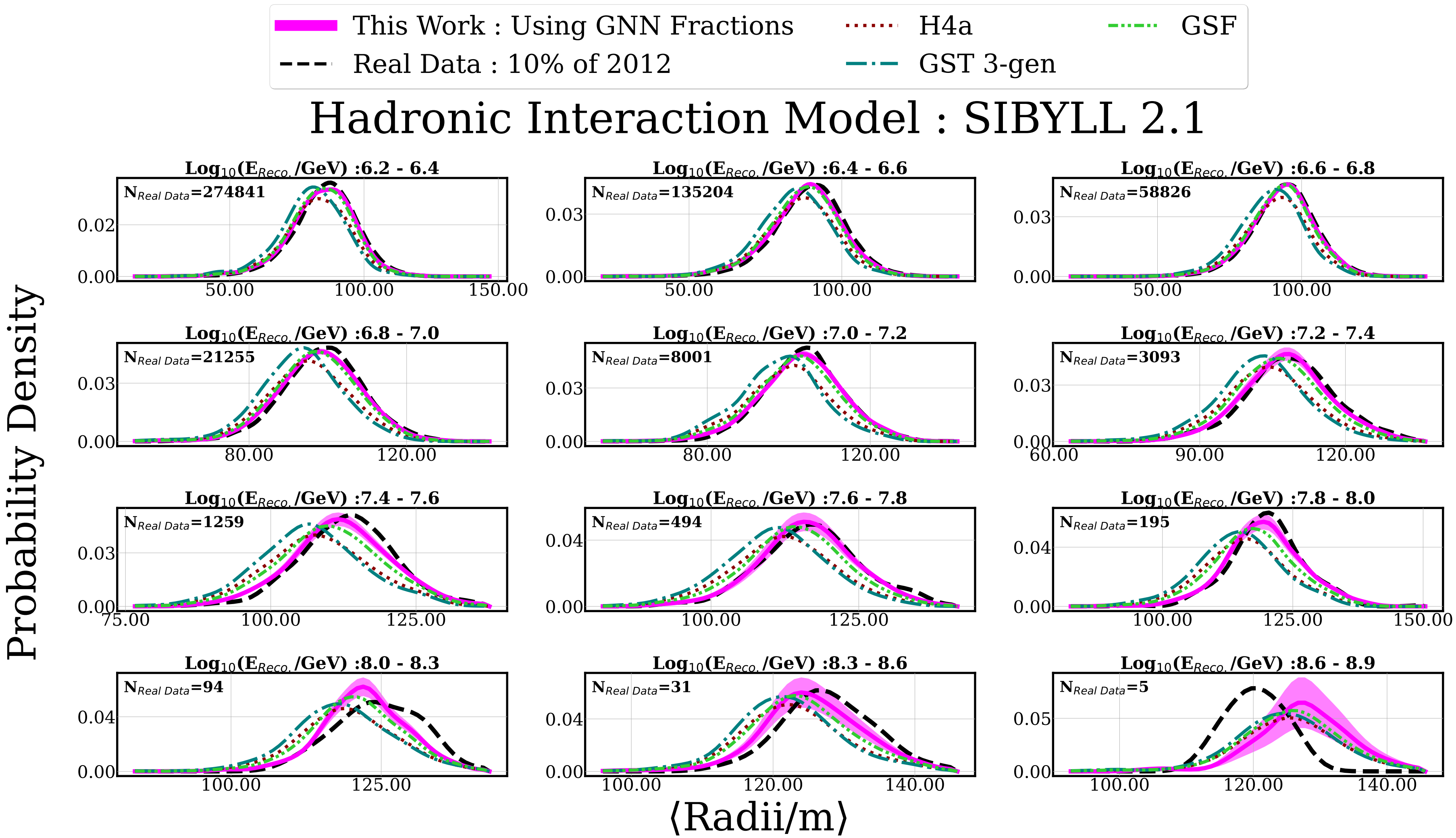


Figure 9.7: Compatibility check for composition-fractions obtained from GNN (details in Section 9.1) for *Mean-Radii* (with maximum DOM-radii cut of 200 m - discussed in Section 7.4.1) and comparison with fractions from other flux-models, on a *Linear-scale*. To see the plot on *Log-scale*, see Figure A.22.

Energy Bin	This Work (Bootstrap-Fraction)	H4a	GST	GSF
6.2 - 6.4	<u>0.0136</u> (0.576)	0.0108	0.0018	0.0108
6.4 - 6.6	0.0002 (0.568)	<u>0.0008</u>	0.0	0.001
6.6 - 6.8	<u>0.0</u> (0.584)	<u>0.0</u>	<u>0.0</u>	<u>0.0</u>
6.8 - 7.0	0.0151 (0.611)	<u>0.0195</u>	0.025	<u>0.0195</u>
7.0 - 7.2	<u>0.1061</u> (0.702)	0.0857	0.0066	<u>0.1061</u>
7.2 - 7.4	0.5388 (0.706)	0.0079	<u>0.718</u>	0.3782
7.4 - 7.6	0.0011 (0.527)	0.0715	0.0024	<u>0.1219</u>
7.6 - 7.8	<u>0.0018</u> (0.058)	0.0002	0.0012	0.0005
7.8 - 8.0	<u>0.1299</u> (0.444)	0.0061	<u>0.1299</u>	0.0019
8.0 - 8.3	<u>0.0288</u> (0.699)	0.0009	0.0088	0.0023
8.3 - 8.6	<u>0.138</u> (0.793)	0.0132	0.0089	0.0279
8.6 - 8.9	<u>0.018</u> (0.9)	0.0	0.0	0.0

Table 9.1: p-value (using KS-test) for compatibility test with $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$, using this work and other flux-models. "Bootstrap-Fraction" represents the fraction of bootstrapped-examples (with 1000 iterations), with a p-value greater than the listed p-value. The best p-values in an energy-bin are underlined.

energies (in comparison to low energies). Updating this work with greater statistics and detector-systematics is planned for the future.

Figure 9.6 and Figure 9.7 also show the comparison with three other flux-models namely H4a [90], GST 3-gen [91] and GSF [92]. For both $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ (Figure 9.6) and Mean-Radii (Figure 9.7), this work seems more compatible with real-data (burnsample) than other flux-models. This can also be seen at the tails of the distributions, in the log-scale plots ($\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ - Figure A.21; Mean Radii - Figure A.22). Table 9.1 and Table 9.2 present a quantitative measure of this. The tables represent the p-value comparison, using KS-test, for comparison between real data (burnsample) and the choice of weighting scheme (using reconstructed fractional contributions from this work or other flux-models). The null hypothesis is that the two distributions are sampled from the same underlying distribution. A p-value close to 1 indicates evidence in favor of the null hypothesis. To account for the error band evaluated for this work, Table 9.1 and Table 9.2 also list down "Bootstrap-Fraction". It is defined as the fraction of bootstrapped examples (1000 iterations), with a p-value greater than the one listed. As is expected from Figure 9.6 and Figure 9.7 and clear from the tables, for both $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ (Table 9.1) and Mean-Radii (Table 9.2), this work is generally more compatible with real-data at ICNO, than other flux-models. For energy bins where the p-value for this work is not the highest, the Bootstrap-Fraction is generally greater than 0.5, indicating that the majority of the bootstrapped examples give a p-value greater than the one listed in the tables.

The overall compatibility of the fractions obtained from this work (shown in Figure 9.2), among different shower-observables (Figure 9.6 and Figure 9.7) for SIBYLL 2.1 generate confidence in the reconstructed fractions. This study with a limited dataset

Energy Bin	This Work (Bootstrap-Fraction)	H4a	GST	GSF
6.2 - 6.4	<u>0.2279</u> (0.656)	0.1752	0.1752	0.1752
6.4 - 6.6	0.8814 (0.891)	0.6815	0.6815	<u>0.9856</u>
6.6 - 6.8	<u>0.0279</u> (0.628)	<u>0.0279</u>	0.0084	0.019
6.8 - 7.0	0.768 (0.819)	0.6447	<u>0.9524</u>	0.768
7.0 - 7.2	<u>0.9876</u> (0.959)	0.4999	0.4999	0.7499
7.2 - 7.4	<u>0.7838</u> (0.926)	0.2044	0.2044	0.3891
7.4 - 7.6	<u>0.9996</u> (0.996)	0.2849	0.8285	0.8285
7.6 - 7.8	<u>0.9878</u> (0.975)	0.6748	<u>0.9878</u>	0.9327
7.8 - 8.0	<u>0.8733</u> (0.826)	0.3465	0.6037	0.1732
8.0 - 8.3	<u>0.0761</u> (0.743)	0.0087	0.0046	0.0087
8.3 - 8.6	<u>0.1873</u> (0.794)	0.119	0.119	<u>0.1873</u>
8.6 - 8.9	<u>0.0047</u> (0.631)	0.0	0.0	0.0

Table 9.2: p-value (using KS-test) for compatibility test with Mean-Radii (with maximum DOM-radii cut of 200 m), using this work and other flux-models. "Bootstrap Fraction" represents the fraction of bootstrapped examples (with 1000 iterations), with a p-value greater than the listed p-value. The best p-values in an energy bin are underlined.

seems to indicate that, without detector-systematics and for SIBYLL 2.1, H4a is indeed not the most compatible flux model with real data at ICNO. On closely comparing the overlap of distributions obtained from this work among the shower-observables, Mean-radii seems to indicate a better overlap with the data. The slight difference for SIBYLL 2.1 seen here is closely connected to the discrepancy in composition-expectation discussed in Section 7.5. Since the difference exists for the choice of all flux models, it might indicate internal inconsistencies in hadronic models to replicate multiple shower-observables. It is unclear at this point if the slight difference between the two is because of different detector-systematic effects or an effect of choice of the hadronic-interaction model. A bigger dataset for real data and simulation for intermediate primary types for SIBYLL 2.3, EPOS-LHC, and QGSjet-II-04 can shed more light on this in the future.

Section A.8 discusses the $\langle \ln(A) \rangle$ expectation from just using the shower-observables. It is presented in Figure A.18. Using the shower observables separately yields lighter composition than the one obtained from GNN. It seems that the heavier composition is because of the footprint, and has influence from the GNN training. However, since detailed systematic studies (hadronic interaction models, detector-systematics) with greater statistics have yet to be performed, it is difficult to conclude which is the true composition.

10 CONCLUSIONS

...bol, ki thoda waqt bahut hai
jism-o-zuban ki maut se pehle
bol, ki sach zinda hai ab tak
bol, jo kuch kehna hai kehle
(... Speak, this little time is plenty
Before the death of body and tongue.
Speak, for truth is still alive
Speak, say whatever is to be said.)

— Bol (Speak Out) : Faiz Ahmad Faiz [648]

This work focused on studying the elemental composition of cosmic-ray (CR) at IceCube Neutrino Observatory (ICNO), in an integrated operation of its surface detector i. e. IceTop (IT) and the in-ice component i. e. IceCube (IC). The observatory is sensitive to CRs between PeV to EeV energy range. This is considered as the energy range wherein the transition from Galactic to extragalactic origins is anticipated as a potential explanation for the shape and features of the observed cosmic ray spectrum.

In order to study the elemental composition, the work used detailed MC-simulations of EASs for the use case of CR-analysis at ICNO. Tested quality-cuts were used to ensure that the events selected for analysis are cleaned as well as feasibly reconstructable. The cleaned simulations were then used to tackle the problem of studying elemental composition in two different ways. The first method explored different shower-observables from the footprint of CR-initiated Extensive Air-Showers (EASs). The study provided an opportunity to benefit from the energy loss of TeV muons in IC. Using EAS-based physics motivations the work focused on studying observables which study different aspects of TeV muons in EASs. The most prominent among them are shower observables which are based on the muon multiplicity and the lateral extent of TeV muon bundle in the shower core. This work revealed an internal inconsistency among hadronic models to explain the multiplicity and lateral spread of TeV muons together. This is a useful physics insight which might be of interest to developing the next generation of hadronic interaction models to describe the production of TeV muons in air showers.

A major part of this work focused on developing an approach for CR-mass estimation at ICNO which has the possibility to not only benefit from the composition-sensitive observables developed in this work (and earlier) but also benefits from capturing useful correlation in the low-level footprint signal information. Moreover, another requirement of this analysis, which is successfully fulfilled, was to develop an approach that is flexible to detector upgrades of ICNO (namely IceCube-Gen2). To do this a Graph Neural Network (GNN)-based approach is developed as a part of this work. The GNN-based approach is EAS-physics informed and is developed in a way to include maximal inductive biases to enhance CR-composition sensitivity and its discrimination power in an efficient manner. A Gradient-Boosted Decision Trees (GBDT) based approach is utilized for the energy estimate of the primary in the CRs.

The composition estimate from the GNN-based approach predicts a heavier composition than one associated before at ICNO, with increasing energies. The heavier composition is a consequence of a greater expectation of Fe-like nuclei from this work. If this is indeed the case, then this indicates the presence of the highest energy galactic cosmic rays in the energy range between PeV to EeV CRs. Even though the composition expectation is in tension with previous results, the expectation is much closer to the results obtained by another independent observatory *i. e.* KASCADE for the choice of the same hadronic interaction model (*i. e.* SIBYLL 2.1).

Another important observation from this work was that the composition expectation from the GNN is also in tension with the composition expectation from individual air-shower observables. It is very important to notice that this analysis only uses about 3% of the data used for the previous publication from ICNO. Hence with the limited dataset, it is very difficult to make a concrete statement about the differences seen in this work, as well as make a statement on an astrophysical consequence. A future update of this work is planned to add more measured data, with an in-depth study into the composition-sensitive observables developed in this work to look for any hidden biases. A detailed study of systematics for this work is also envisioned.

Part I

APPENDIX

A PHYSICS

A.1 CORSIKA ATMOSPHERE

The standard choice of atmosphere in CORSIKA is made up of 78.1% N_2, 21.0% O_2, and about 0.9% Ar. To represent the density profile of the atmosphere with altitude, a five-layered model is utilized. This model allows for a depiction of the density structure of the atmosphere that is satisfactorily consistent with real-world measurements. In the bottom four layers, the density profile is described by an exponential decrease of mass overburden, X(h), with altitude h. Whereas for the fifth layer, it follows a linearly decreasing trend. This can be expressed as:

$$X(h) = \begin{cases} a_i + b_i e^{-h/c_i} & \text{for layer 1 to 4} \\ a_5 - b_5 \frac{h}{c_5} & \text{for layer 5} \end{cases} \tag{A.1}$$

The values of the parameters (a_i, b_i, c_i) are carefully chosen to ensure continuity and differentiability of Equation A.1. The density can be obtained from the overburden by the relation

$$\rho(h) = -\frac{dX(h)}{dh} \tag{A.2}$$

A.2 MUON RANGE @ ICECUBE

The mean energy-loss equation for muons (Equation 2.3) can be used to get the range of the muon. This is given by:

$$x_f = \frac{\log(1 + E_i \cdot b/a)}{b} \tag{A.3}$$

For IC, a = 0.212 $\frac{\text{GeV}}{\text{mwe}}$ and b = 0.251 $\cdot 10^{-3}$ $\frac{1}{\text{mwe}}$ [649], where Meter Water Equivalent (mwe) is used to quantify the attenuation because of underground detection. This equation can now be used to get the cutoff energy at which only 0.1% of muons reach IC. The cut-off energy is given by:

$$E_{cut}(x) = (e^{b \cdot x} - 1)a/b \tag{A.4}$$

where x is the distance measured in mwe. For IC, the effective length(i. e. x_f), considering ice bubbles and other effects, is about 1119.74 mwe [650, 651]. Using this in Equation A.4, the cut-off energy turns out to be approximately 274 GeV.

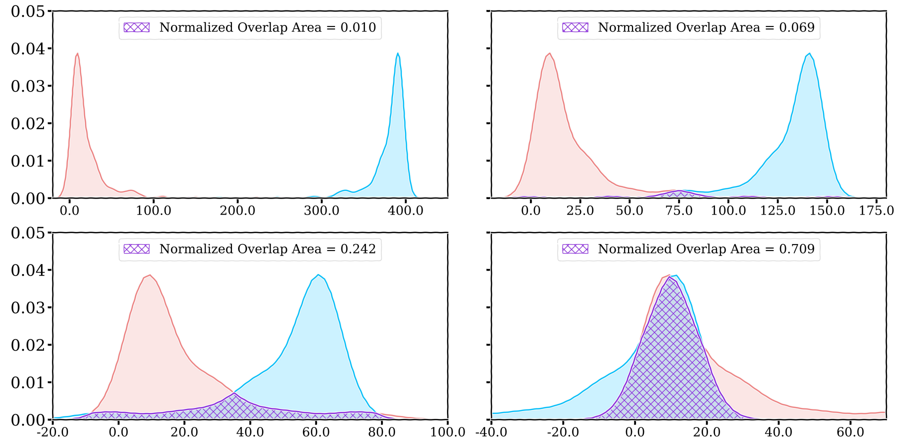


Figure A.1: Normalized Overlap Area (NOA) for two arbitrary normalized distributions. The location of the red distribution (left) is kept fixed in each panel and the bluish distribution (right) is moved. The legends indicate the NOA for each panel. Read Section A.4 for details.

A.3 FIGURE OF MERIT : FOM

Figure of Merit (FOM) gives a measure of separation between two distributions[1] (**i** and **j**) and is given by:

$$\mathrm{FOM} = \frac{|\mu_i - \mu_j|}{\sqrt{\sigma_i^2 + \sigma_j^2}} \tag{A.5}$$

where μ_k and σ_k respectively denotes the mean and standard deviation of the k^{th} distribution. It is crucial to note that the definition is only ideal as a measure if the two distributions follow a Gaussian-like distribution around their means. Slightly different definitions can be used in case the underlying distributions are skewed (e.g. Section A.4, [110])

A.4 NORMALIZED OVERLAP AREA: NOA

FOM (Section A.3) is only an ideal measure of separation if two distributions are Gaussian-like around their means. In order to come up with an alternative measure of overlap, **Normalized Overlap Area (NOA)** is chosen. It is the overlapping area of two normalized (i. e. area under curve = 1) distributions. The greater the overlap, the greater the NOA value. A smaller value is better for separation/classification. Hence, NOA has the opposite behavior to FOM. The change in NOA with change in overlap for two aribtrary distributions is depicted using Figure A.1. The two distributions are normalized individually. The red distribution (left) is kept fixed in each panel and the mean of the blue distribution (right) is moved, keeping the shape of both distributions fixed. Trapezoidal-rule [654] is used to estimate the overlap area. With an increase in overlap, the NOA value increases and vice versa.

1 The measure is also common among other cosmic-ray observatories [652, 653].

A.5 MEAN RADII AND MEAN CHARGE

A.6 TESTING TEMPLATE FITTING

Section 8.4 introduced a Template-Fitting method used to reconstruct the contribution of each-primary type in a dataset with unknown primary-type contributions/fractions. Before applying the method on burnsample it is essential to test the validity of the method. In order to do that blinded datasets were created and the test was performed. Pure-p, Pure-He, Pure-O, Pure-Fe, and a sinusoidal variation between p and Fe were tested. Figure 8.12 presents the result of such a test by showing the comparisons between the true mean-mass and reconstructed mean-mass, as a function of reconstructed energy. Figure A.6 (Pure-p), Figure A.7 (Pure-He), Figure A.8 (Pure-O), Figure A.9 (Pure-Fe) and Figure A.10 (sinusoidal variation between p and Fe) represent the template-fits obtained in the different energy-bins.

A.7 TEMPLATE FITS ON REAL DATA : USING GNN PREDICTIONS

The Template-Fitting introduced in Section 8.4 can be used to reconstruct the primary-type contributions from real data (burnsample here). This can be done by using GNN predictions of logarithmic mass i. e. ln(A) for each CR-initiated EAS/event. Figure 9.2 presents the reconstructed primary-type fractions from this work(using ln(A) predictions from GNN), as a function of reconstructed energy. Figure A.11 presents the underlying KDE template-fits in each energy-bin. Chapter 9 presents a detailed discussion of the reconstructed fractions from the burnsample and also the comparison with other works. Figure A.12 presents the comparison of the primary-type fractions obtained from this work, and the last relevant ICoII publication [20] (systematics not included). Figure A.13 presents the comparison of $\langle \ln(A) \rangle$ expectation for the choice of different hadronic-models, and the results from the previous ICoII publication [20] for SIBYLL 2.1. The corresponding template fits are depicted in Figure A.14 (SIBYLL 2.3), Figure A.15 (EPOS-LHC) and Figure A.16 (QGSJet-II-04). Figure A.17 presents the comparison of $\langle \ln(A) \rangle$ expectation from this work (using GNN predictions) and the previous ICoII publication [20] with results from other observatories. For details of the plots presented here, read Section 9.1.

A.8 TEMPLATE FITS ON REAL DATA : USING AIR-SHOWER OBSERVABLES

Similar to Section 9.1, the Template-Fitting method can also be used to get primary-type contributions from different EAS observables. Different shower-observables allow us to capture different aspects of EAS-physics. A difference in mass expectation between different shower observables can indicate internal inconsistencies between different hadronic models (to replicate different shower observables at the same time) and/or differences in the effect of detector systematics on different shower observables. Figure A.18 presents the $\langle \ln(A) \rangle$ expectation for the choice of SIBYLL 2.1 hadronic interaction model (and the comparison with $\langle \ln(A) \rangle$ expectation from GNN predictions), for the choice of two EAS-observables at ICNO. $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}} / \frac{GeV}{m}\right)$ (discussed in Section 7.1) and Mean-Radii (with a maximum DOM-radii cut of 200 m - discussed in

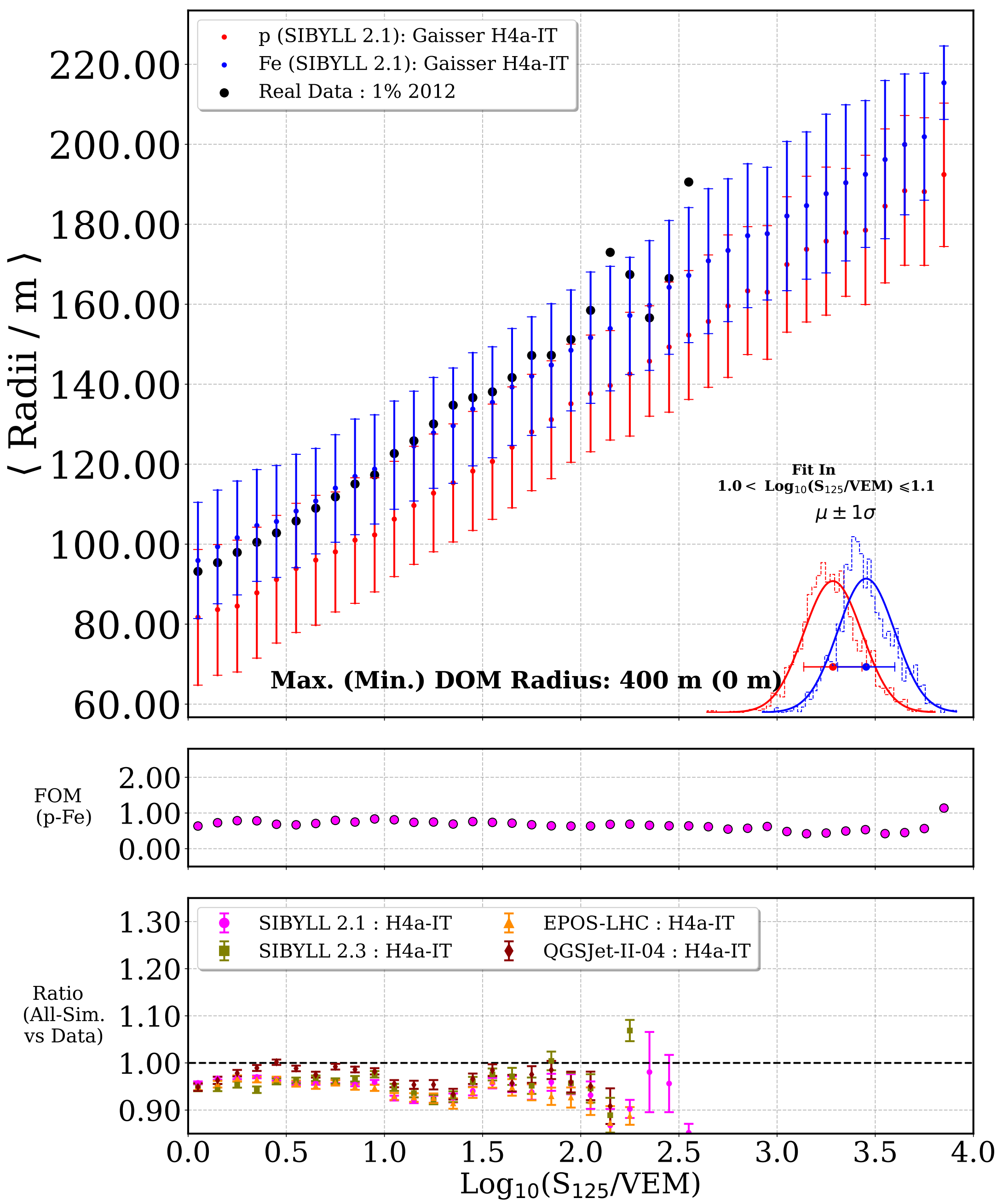


Figure A.2: ***Top Panel:*** Composition sensitivity of Mean Radii (Section 7.4.1) as a function of $\text{Log}_{10}(S_{125}/\text{VEM})$ for SIBYLL 2.1 (weighted to H4a-IT), with maximum DOM distance cut of 400 m. Inset Plot shows the fit in a bin. ***Middle Panel:*** FOM for p-Fe separation. ***Bottom Panel:*** Data-MC overlap for different hadronic models and H4a-IT flux-model.
Read Section 7.1, for more details of underlying physics inputs that go into different panels and Section 7.3 for the physics interpretations. The underlying events are IT-IC coincident events and pass the IT-IC quality cuts discussed in Section 5.3.3.

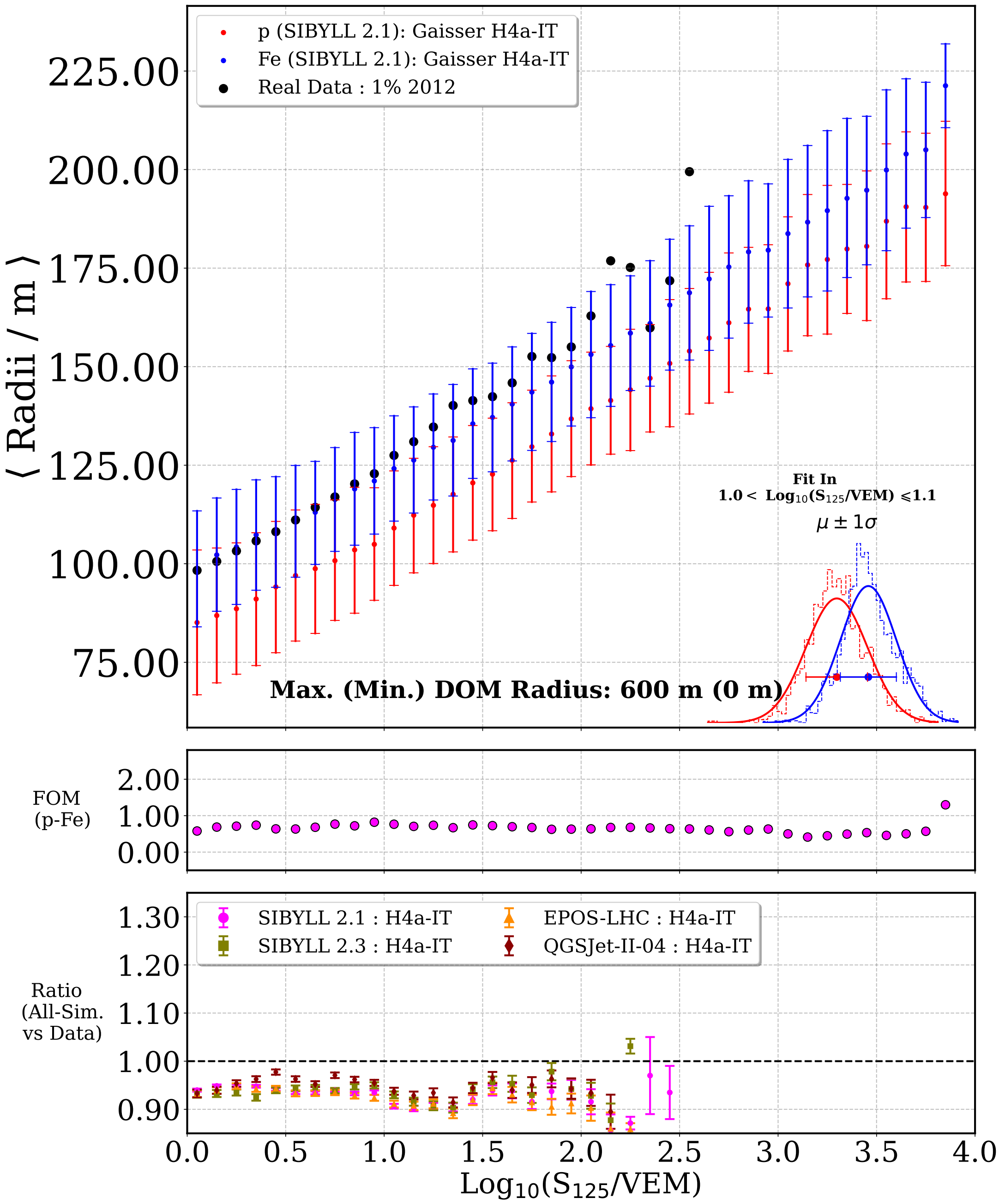


Figure A.3: ***Top Panel:*** Composition sensitivity of Mean Radii (Section 7.4.1) as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ for SIBYLL 2.1 (weighted to H4a-IT), with maximum DOM distance cut of 600 m. Inset Plot shows the fit in a bin. ***Middle Panel:*** FOM for p-Fe separation. ***Bottom Panel:*** Data-MC overlap for different hadronic models and H4a-IT flux-model.
Read Section 7.1, for more details of underlying physics inputs that go into different panels and Section 7.4.1 for the physics interpretations. The underlying events are IT-IC coincident events and pass the IT-IC quality cuts discussed in Section 5.3.3.

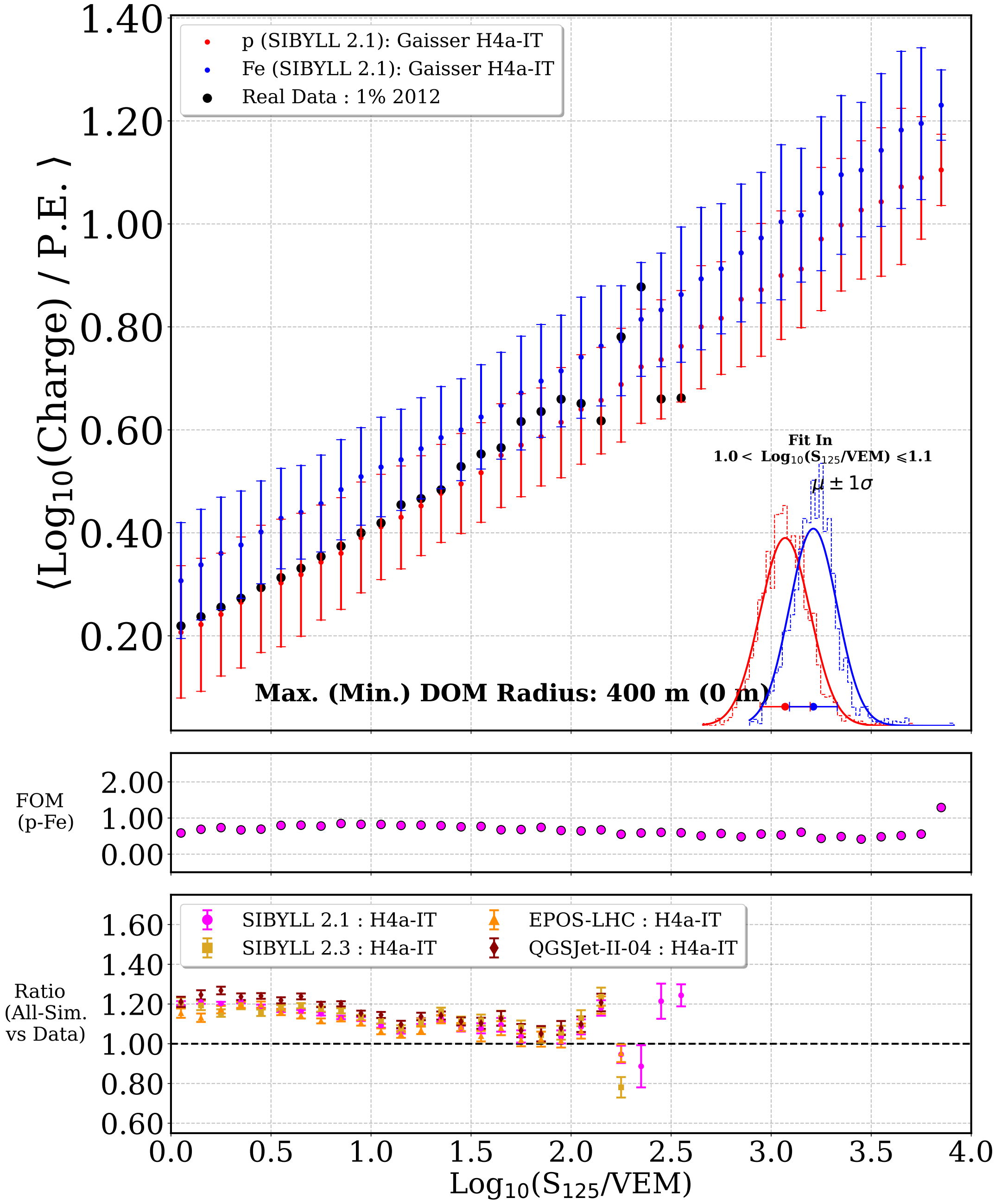


Figure A.4: ***Top Panel:*** Composition sensitivity of **Mean Charge** (Section 7.4) as a function of $\text{Log}_{10}(S_{125}/\text{VEM})$ for SIBYLL 2.1 (weighted to H4a-IT), with maximum DOM distance cut of 600 m. Inset Plot shows the fit in a bin. ***Middle Panel:*** FOM for p-Fe separation. ***Bottom Panel:*** Data-MC overlap for different hadronic models and H4a-IT flux-model.
Read Section 7.1, for more details of underlying physics inputs that go into different panels and Section 7.4.2 for the physics interpretations. The underlying events are IT-IC coincident events and pass the IT-IC quality cuts discussed in Section 5.3.3.

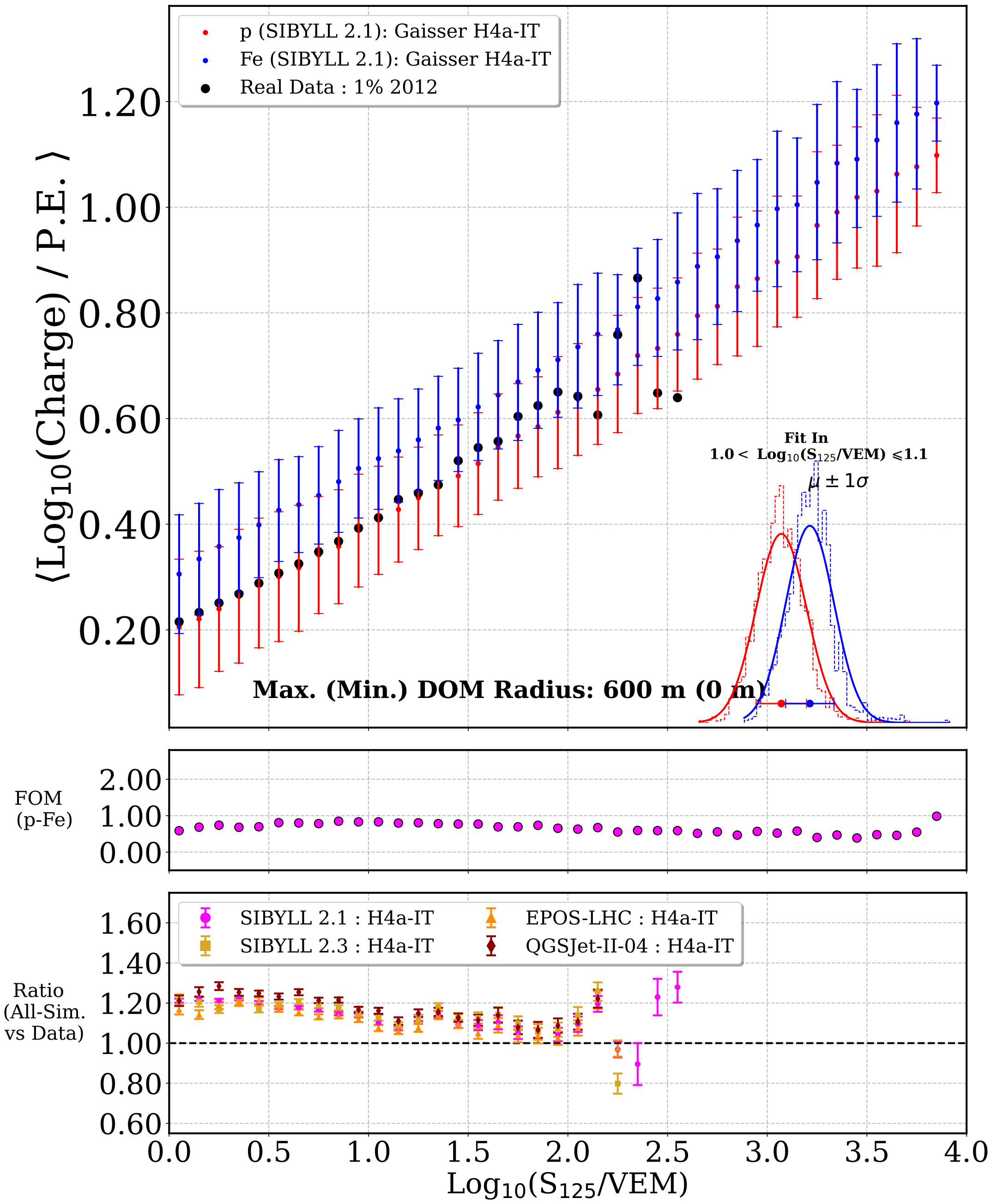


Figure A.5: ***Top Panel:*** Composition sensitivity of **Mean Charge** (Section 7.4) as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$ for SIBYLL 2.1 (weighted to H4a-IT), with maximum DOM distance cut of 600 m. Inset Plot shows the fit in a bin. ***Middle Panel:*** FOM for p-Fe separation. ***Bottom Panel:*** Data-MC overlap for different hadronic models and H4a-IT flux-model.
Read Section 7.1, for more details of underlying physics inputs that go into different panels and Section 7.4.2 for the physics interpretations. The underlying events are IT-IC coincident events and pass the IT-IC quality cuts discussed in Section 5.3.3.

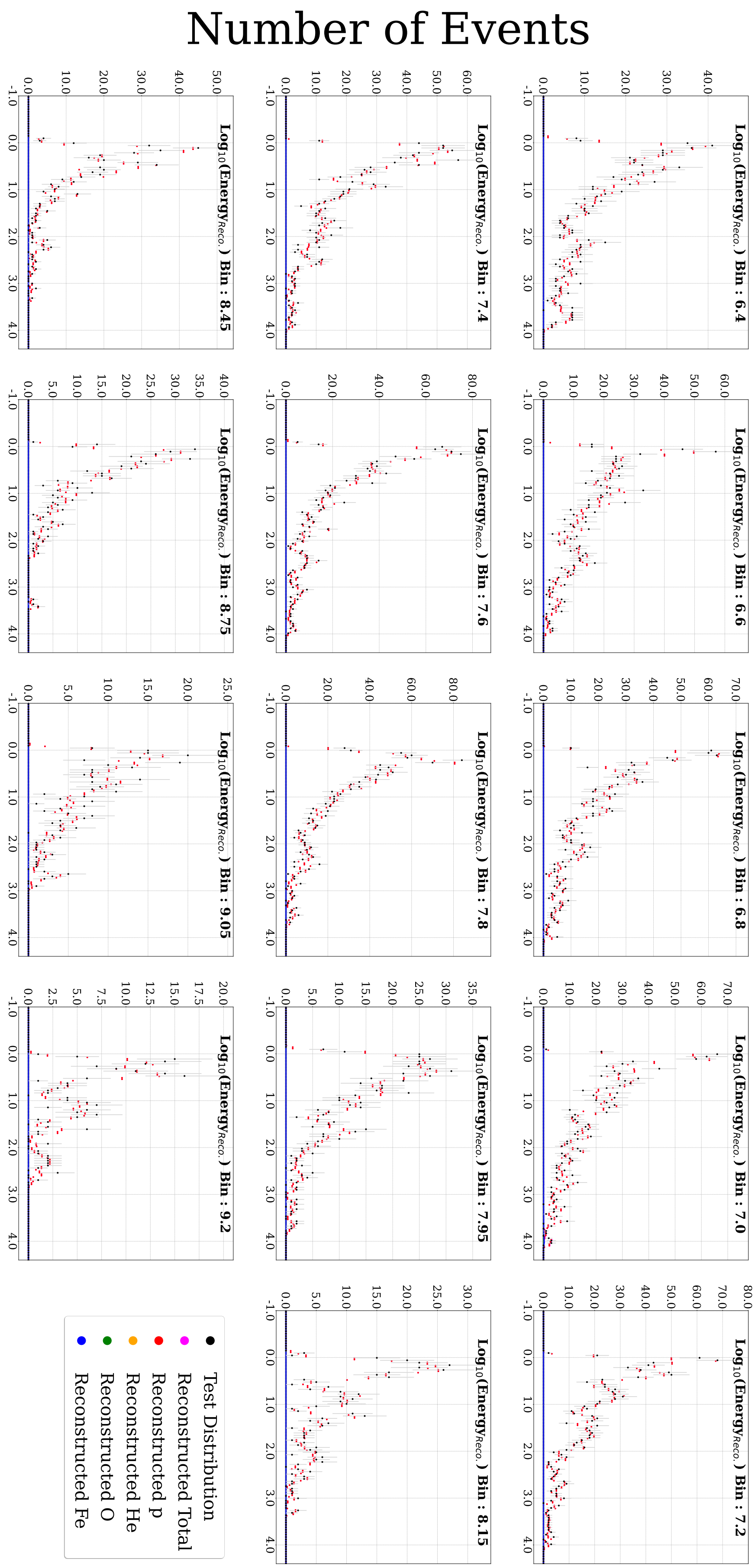


Figure A.6: Underlying KDE-template fits for the case of pure Proton. For details read Section A.6.

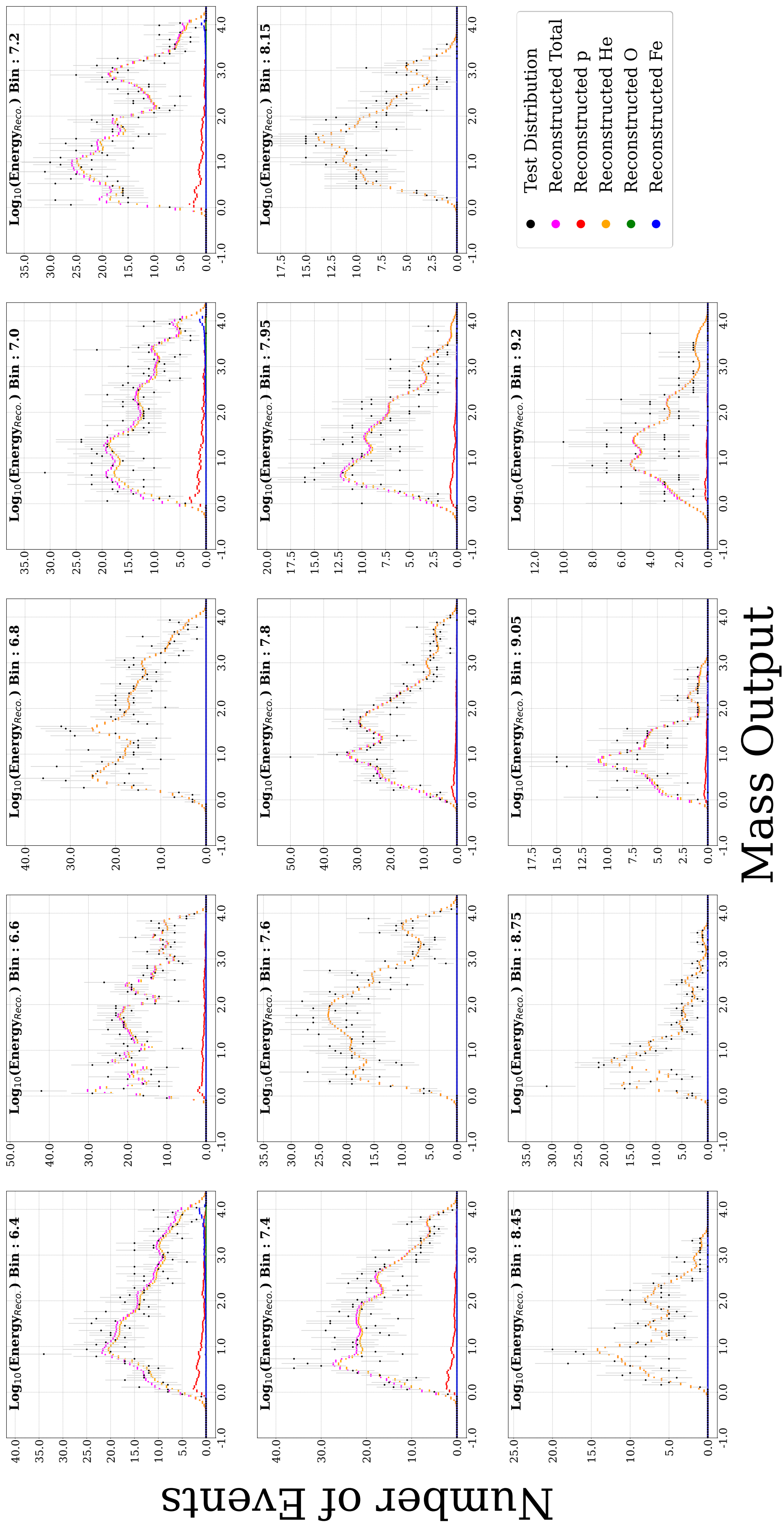


Figure A.7: Underlying KDE-template fits for the case of pure Helium. For details read Section A.6.

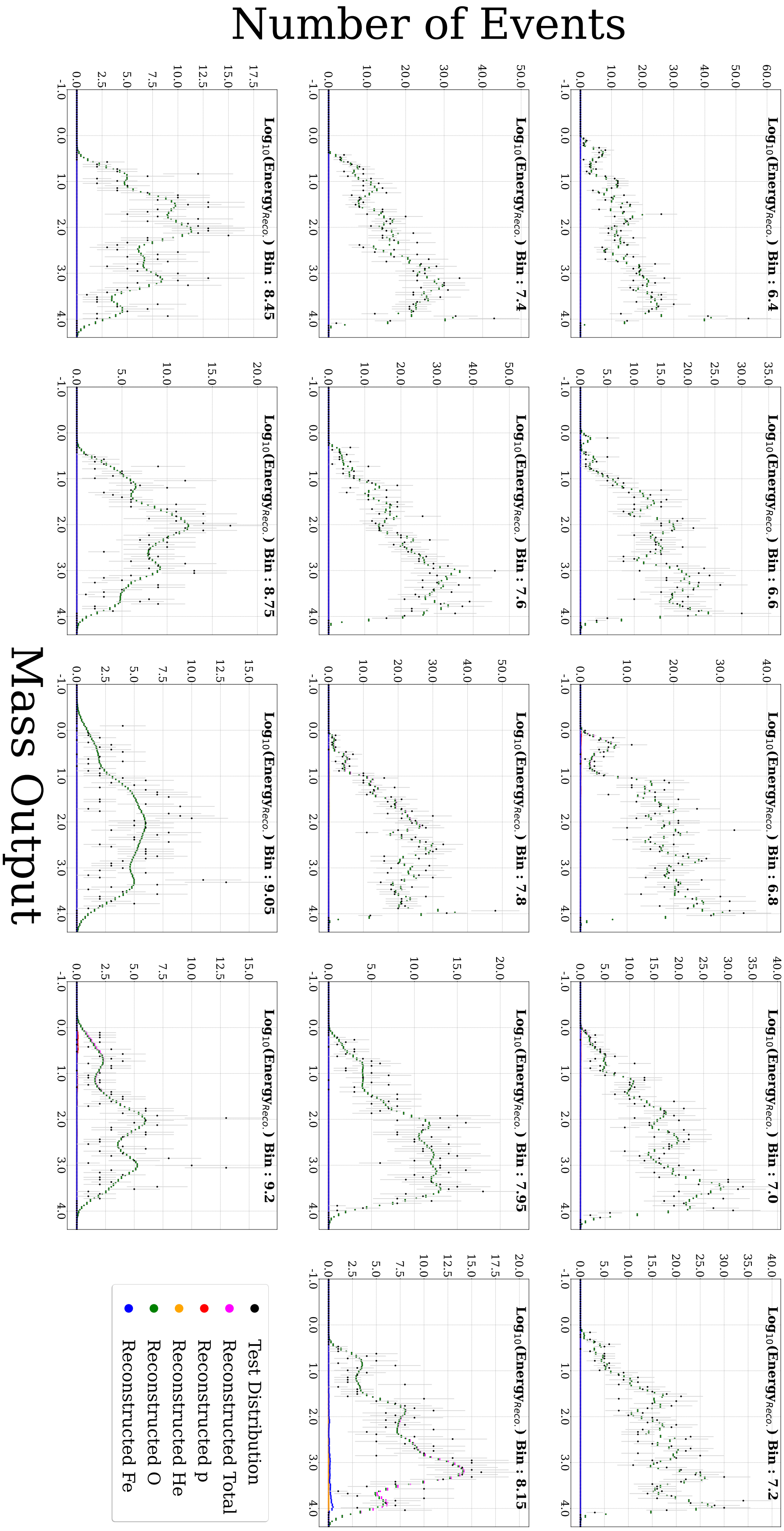


Figure A.8: Underlying KDE-template fits for the case of pure Oxygen. For details read Section A.6.

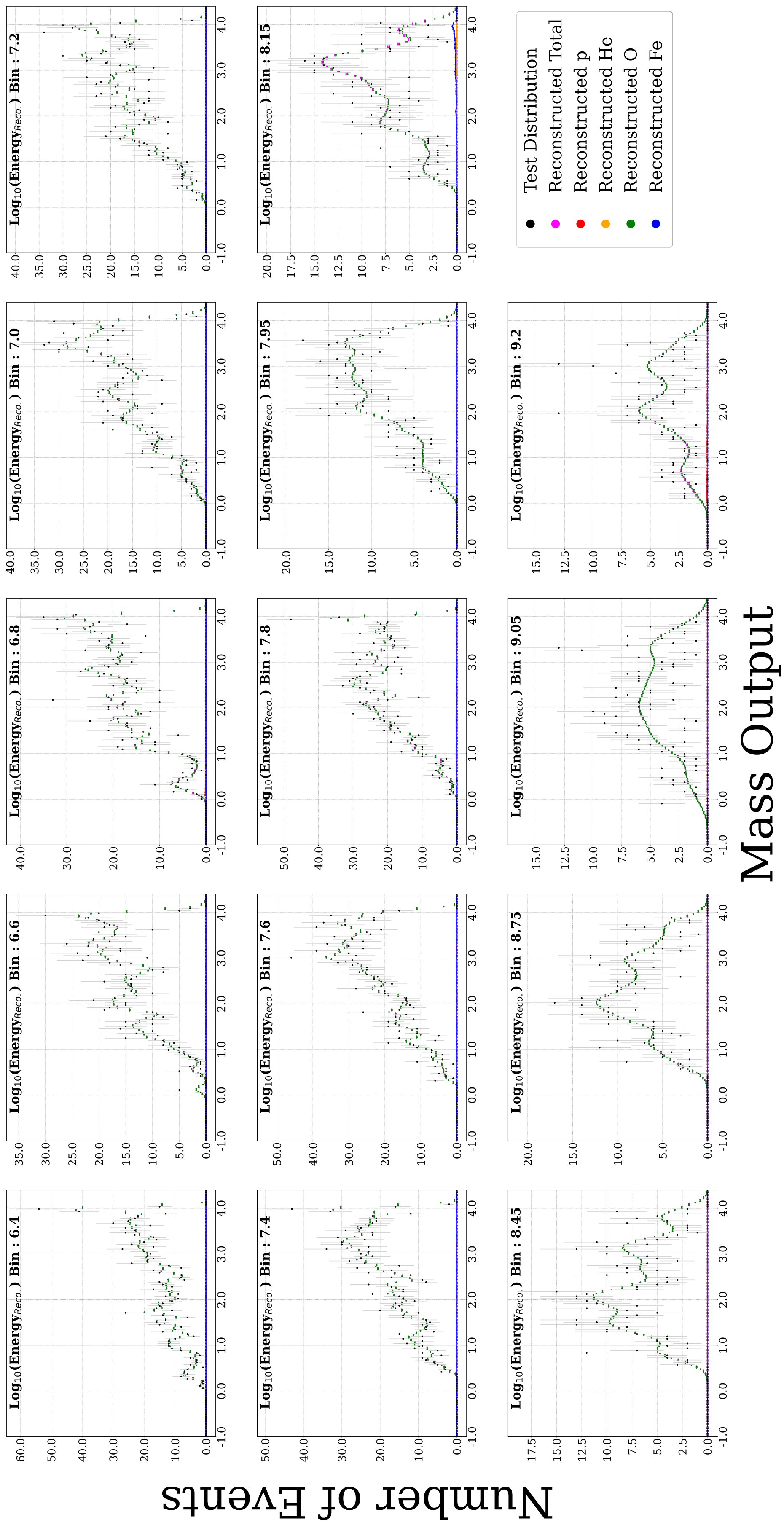


Figure A.9: Underlying KDE-template fits for the case of pure Iron. For details read Section A.6.

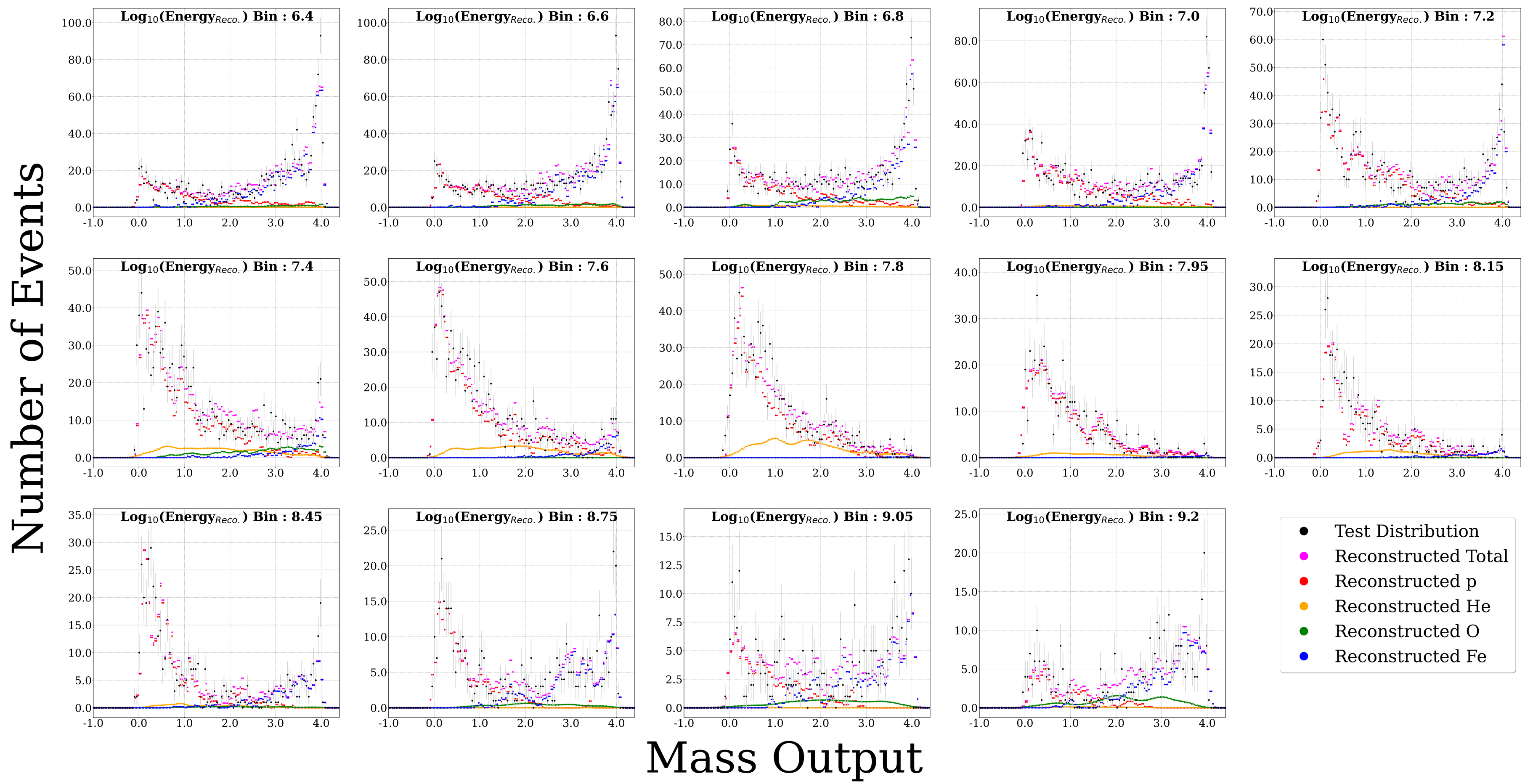


Figure A.10: Underlying KDE-template fits for the case of sinusoidal variation between Proton and Iron. For details read Section A.6.

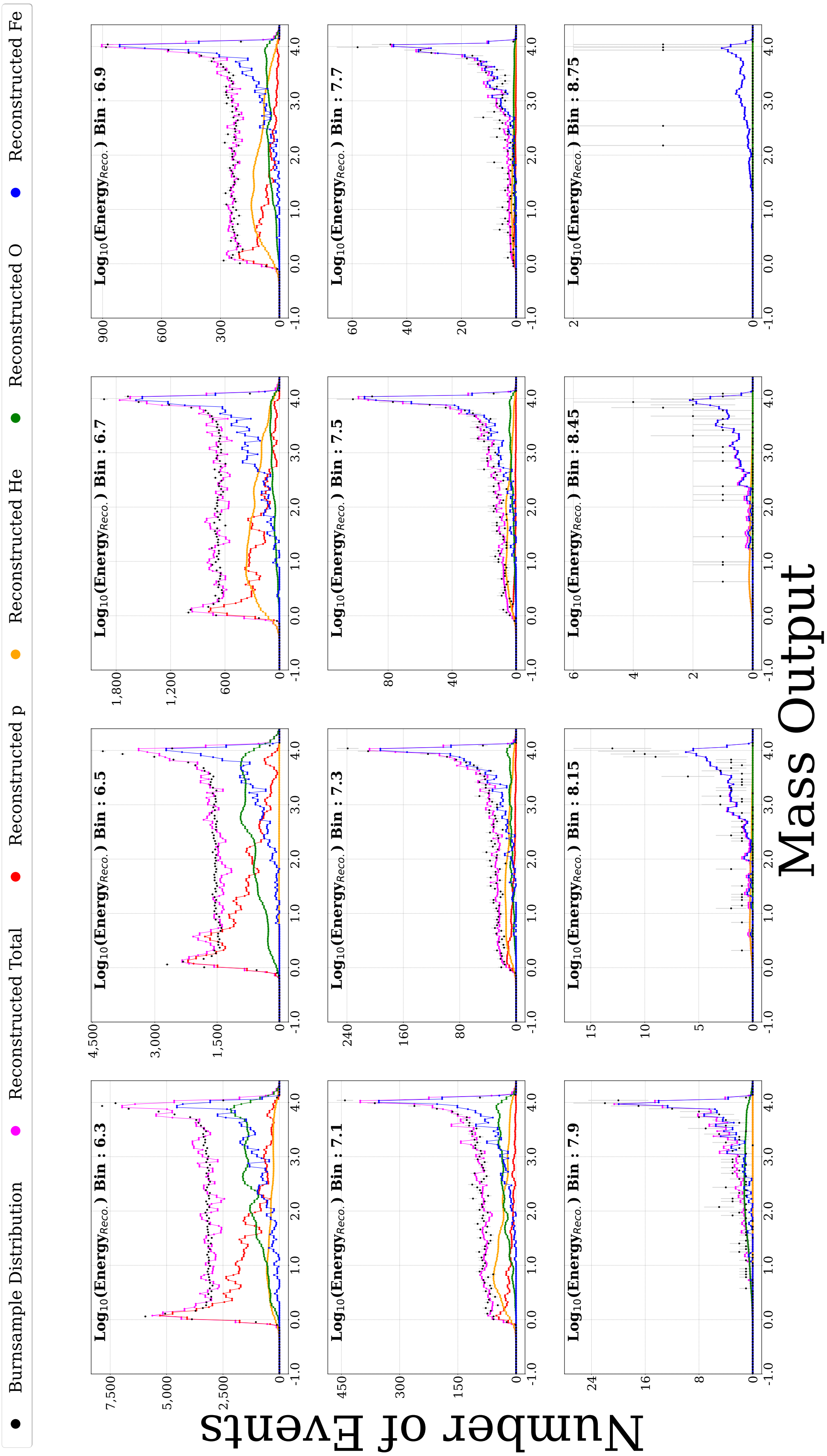


Figure A.11: KDE-template fits for burnsample using ln(A) predictions from GNN, and using SIBYLL 2.1 as the hadronic interaction model for template creation. For details read Section A.7.

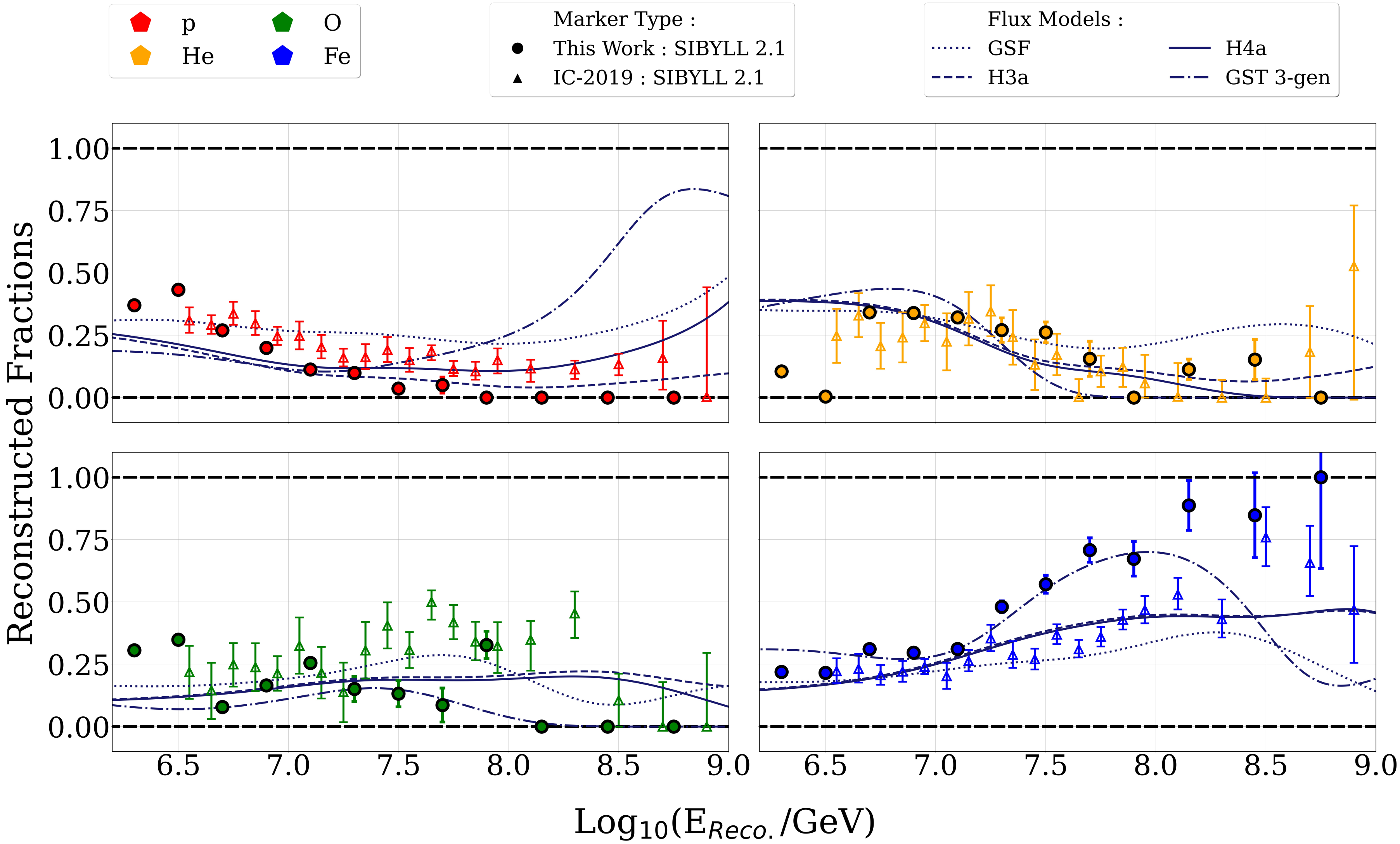


Figure A.12: Reconstructed proton, helium, oxygen and iron fractions in the burnsample (using ln(A) predictions from GNN-based architecture - Read Chapter 8 for details) and comparison with last relevant ICoII publication [20] (systematics not included), as a function of reconstructed energy. The fractions are obtained using the Template-Fitting method (discussed in Section 8.4), by templates created using SIBYLL 2.1 as the hadronic model. The fractions from this work are labeled as "This Work: SIBYLL 2.1", and the ones from [20] are called as labeled as "IC-2019: SIBYLL 2.1". For details read Section A.7.

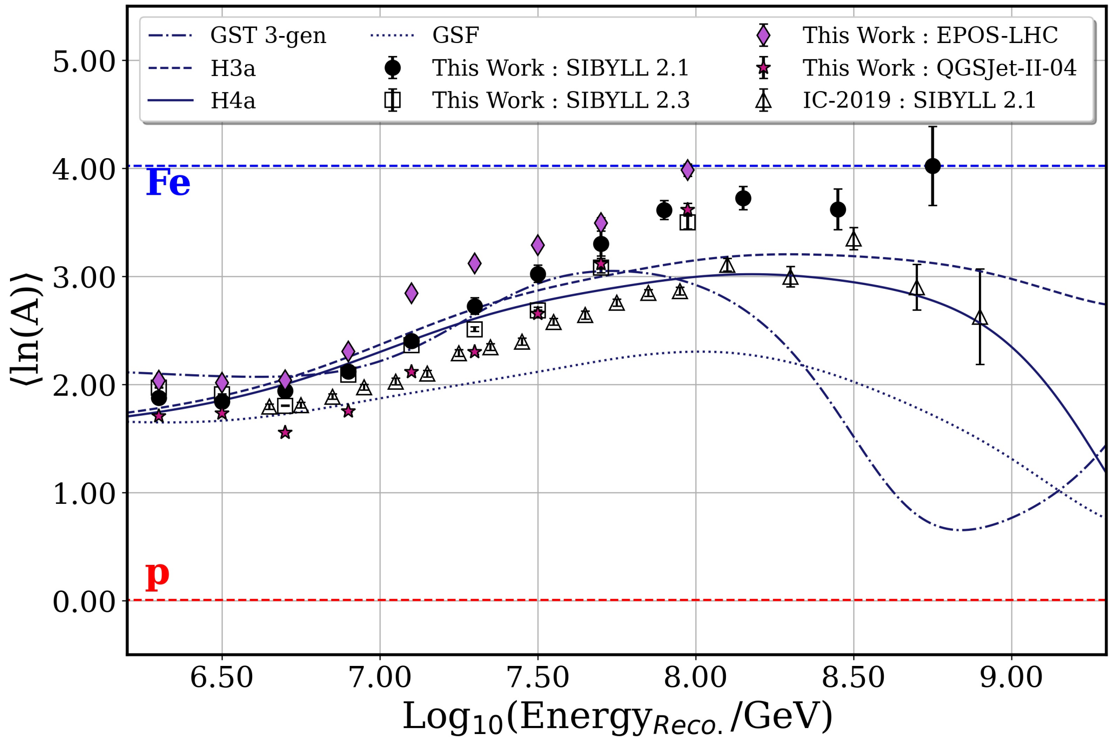


Figure A.13: Comparison of $\langle \ln(A) \rangle$ expectation from this work(using GNN predictions) and the previous ICoII publication [20] for the choice of different hadronic interaction models, as a function of reconstructed energy.

Section 7.4.1) are the two chosen EAS-observables, and are same as for the discussion in Section 7.5. The choice was made because the two are among the representative examples of composition-sensitive shower observables depending on muon multiplicity and their lateral extent. The underlying template fits are shown in Figure A.19 and Figure A.20.

On comparing the composition expectation obtained from Template-fitting separately from different EAS-observables, we see a trend towards heavier composition with an increase in energy. However, we see a significantly lighter composition expectation than that from the GNN-prediction as well as from previous ICoII-publication [20]. Moreover, the composition expectation from different EAS-observables also don't match. Similar to the differences seen in Figure 9.3, the differences(with the results from GNN and among the observables) might come from the unknown effect of detector-systematics, limited dataset volume in burnsample and/or from differences/internal-inconsistencies in hadronic model/s to describe different EAS-observables together. Since the exploration of the first two is planned for future work, we will discuss the possibility of looking for differences/internal inconsistencies in the hadronic model/s.

The fraction obtained from different EAS-observables separately will tend to be biased by how well/close the MC-simulations are able to replicate the distributions in real data for that particular observable. However, since the CR-composition of a dataset should be fixed, the obtained fractional contributions from any shower-observable need to be consistent among the shower-observables. Performing a multi-observable constrained template fit provides the possibility to obtain a composition that is consistent among shower observables. GNN-based architecture introduced in Chapter 8

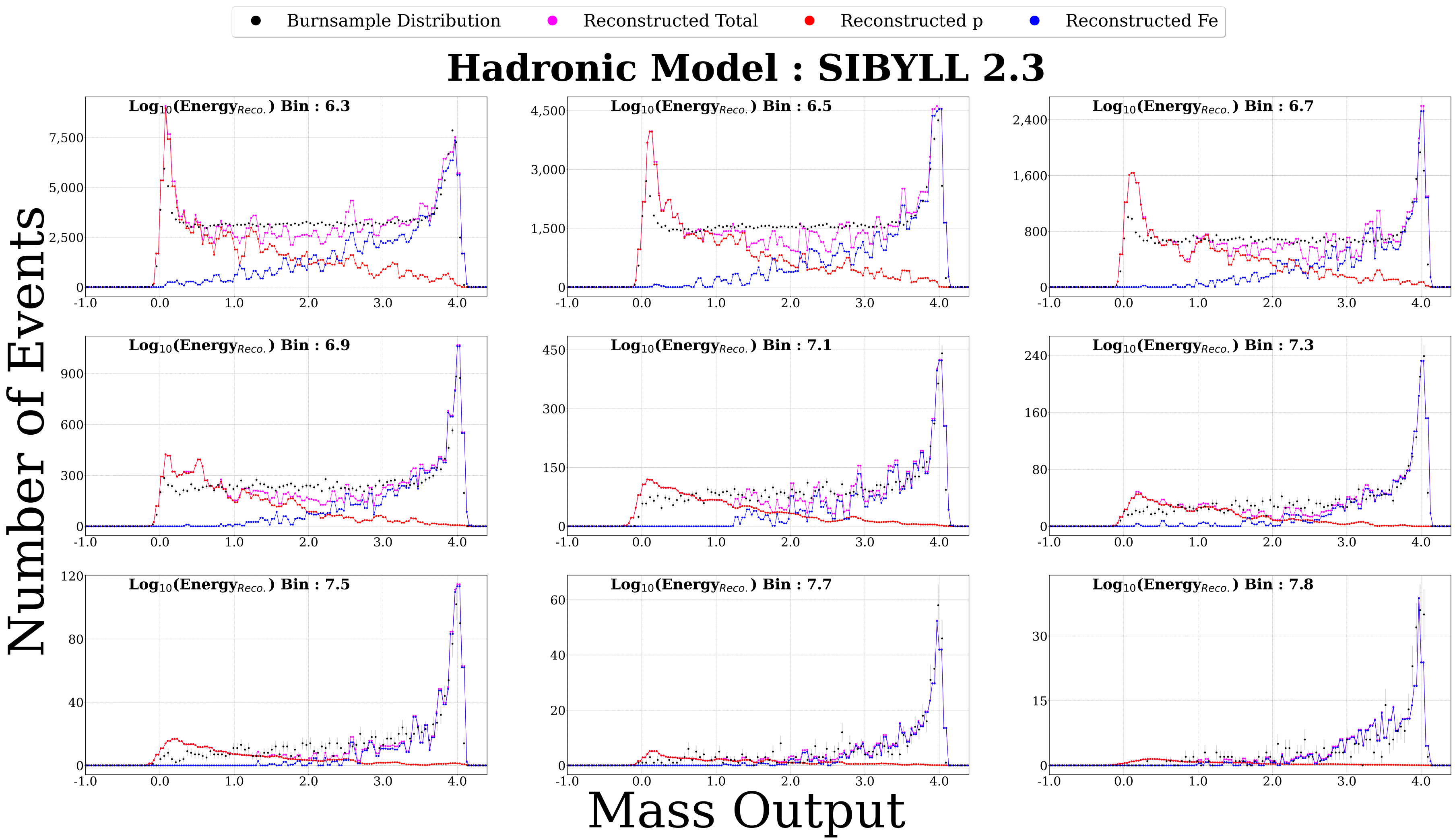


Figure A.14: KDE-template fits for burnsample using ln(A) predictions from GNN, and using SIBYLL 2.3 as the hadronic interaction model for template creation. For details read Section A.7.

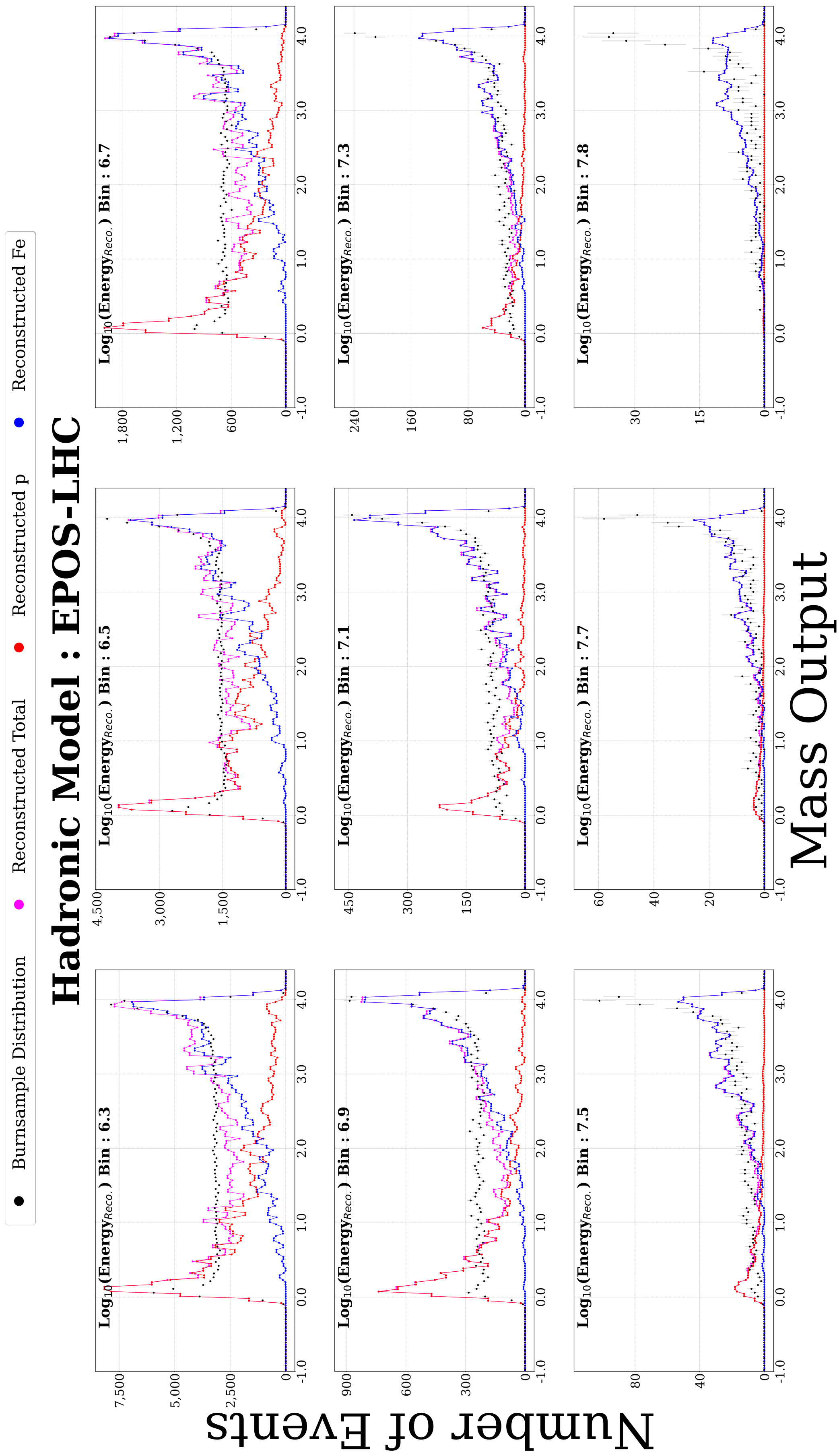


Figure A.15: KDE-template fits for burnsample using ln(A) predictions from GNN, and using EPOS-LHC as the hadronic interaction model for template creation. For details read Section A.7.

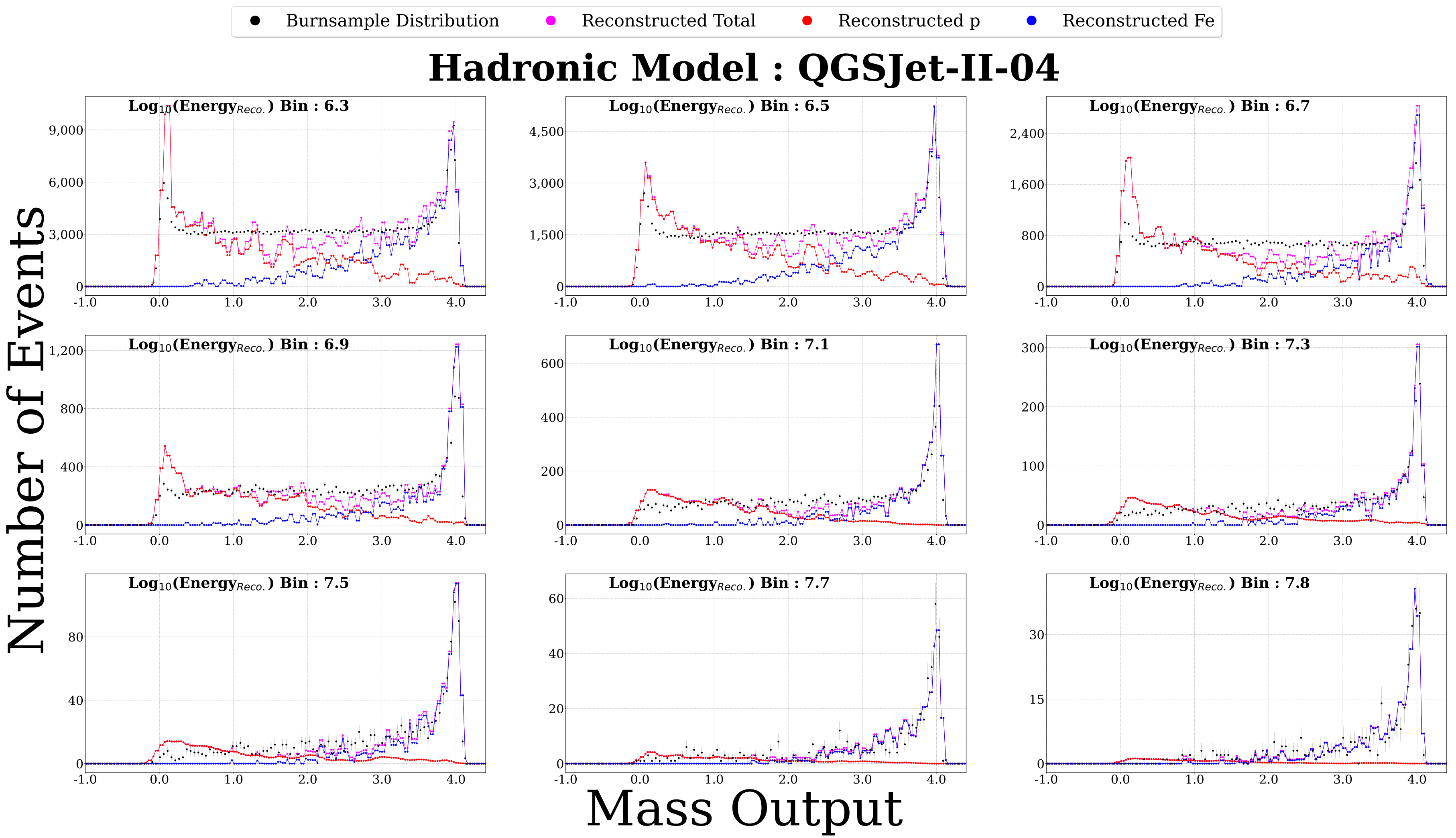


Figure A.16: KDE-template fits for burnsample using ln(A) predictions from GNN, and using QGSJet-II-04 as the hadronic interaction model for template creation. For details read Section A.7.

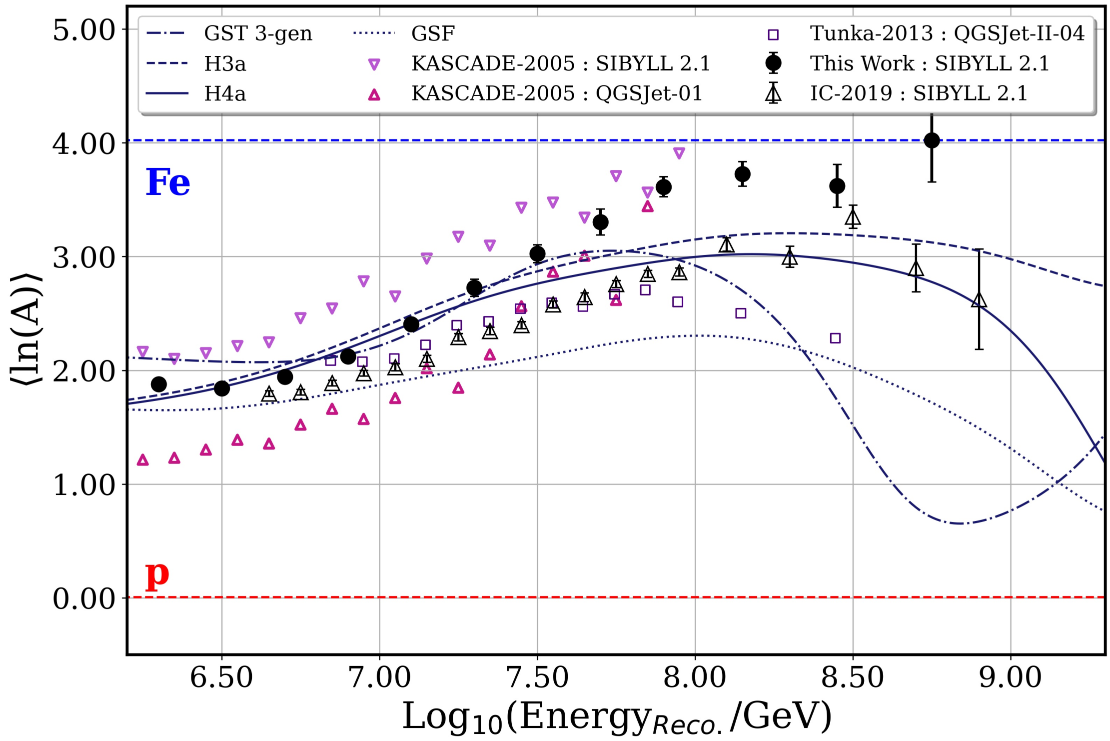


Figure A.17: Comparison of $\langle \ln(A) \rangle$ expectation from this work (using GNN predictions) and the previous ICoII publication [20] with results of $\langle \ln(A) \rangle$ from KASCADE [33], Tunka [646] (error-bars not included) for different hadronic-models, as a function of reconstructed energy. The data for KASCADE was accessed from KCDC (KASCADE Cosmic-ray Data Centre) [50].

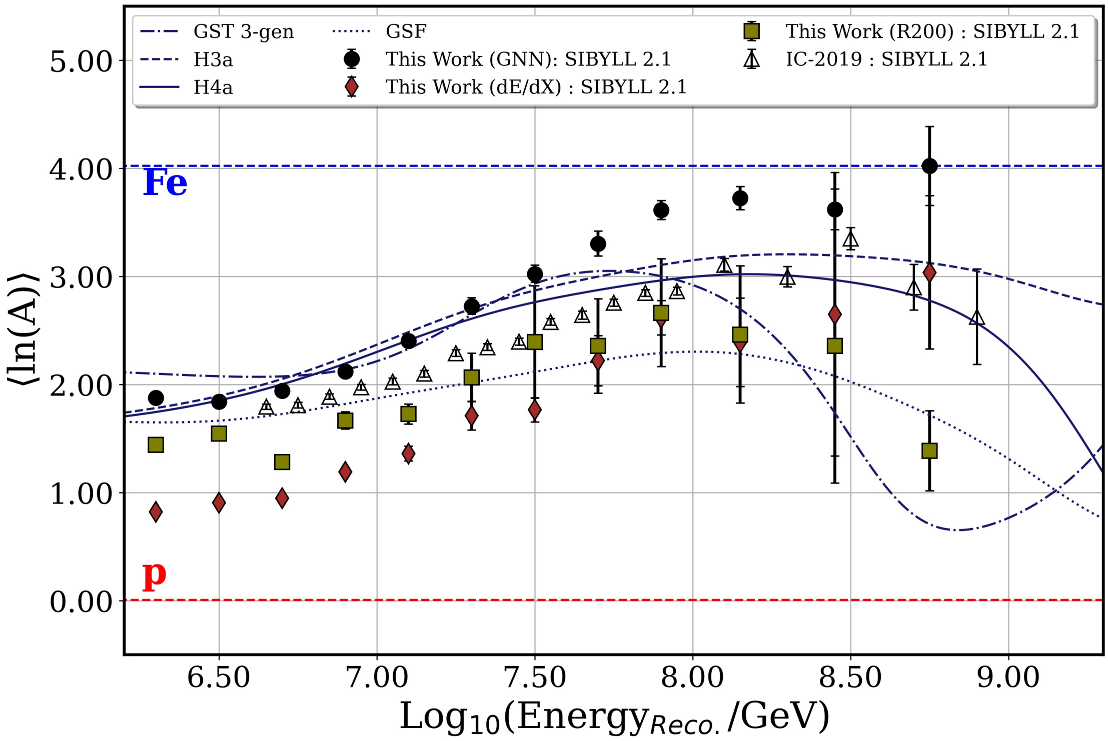


Figure A.18: Comparison of $\langle \ln(A) \rangle$ expectation from this work using EAS-observables i. e. $\text{Log}_{10}\left(\frac{dE}{dX_{1500m}}/\frac{GeV}{m}\right)$ (Labelled as "This Work (dE/dx): SIBYLL 2.1". The details of the observable are discussed in Section 7.1) and Mean-Radii (with a maximum DOM-radii cut of 200 m. Labelled as "This Work (R200): SIBYLL 2.1". The details of the observable are discussed in Section 7.4.1), to those from GNN predictions from this work and the previous ICoU publication [20] for SIBYLL 2.1, as a function of reconstructed energy. The underlying template fits for the EAS-observables are shown in Figure A.19 and Figure A.20.

uses multiple shower-observables which capture different aspects of EAS-physics, like muon-multiplicity (Section 7.1, Section 7.2, Section 7.3), lateral-spread(Section 7.4.1), footprint-info(Section 8.1.1). The training is done to predict logarithmic mass i. e. ln(A). The GNN-training procedure acts as a multi-dimensional constrained fit since the ln(A) for an event should be a single-fixed value, irrespective of the differences among shower observables (because of possible internal inconsistencies in hadronic models). This also makes it more likely that the final prediction through GNN-based architecture will be less biased in a preferential direction, because of the choice of input made. Section 9.2 presents a quantitative compatibility check for GNN-based fraction predictions, putting more trust into the results obtained therein.

A.9 COMPATIBILITY CHECK: FRACTION ESTIMATES

Section 9.2 introduced a method to check the compatibility of primary-type fraction estimates from any custom method with real-data. It also presented(Figure 9.6 and Figure 9.7) the results from such checks using two shower observables, $\mathrm{Log}_{10}\left(\frac{dE}{dX_{1500m}} / \frac{GeV}{m}\right)$ (discussed in Section 7.1) and Mean-Radii (with a maximum DOM-radii cut of 200 m - discussed in Section 7.4.1), on a *Linear-scale*. Figure A.21 and Figure A.22 presents the corresponding plots on a *Log-scale*. This allows us to look at the overlap at the edges. Read Section 9.2 for more details.

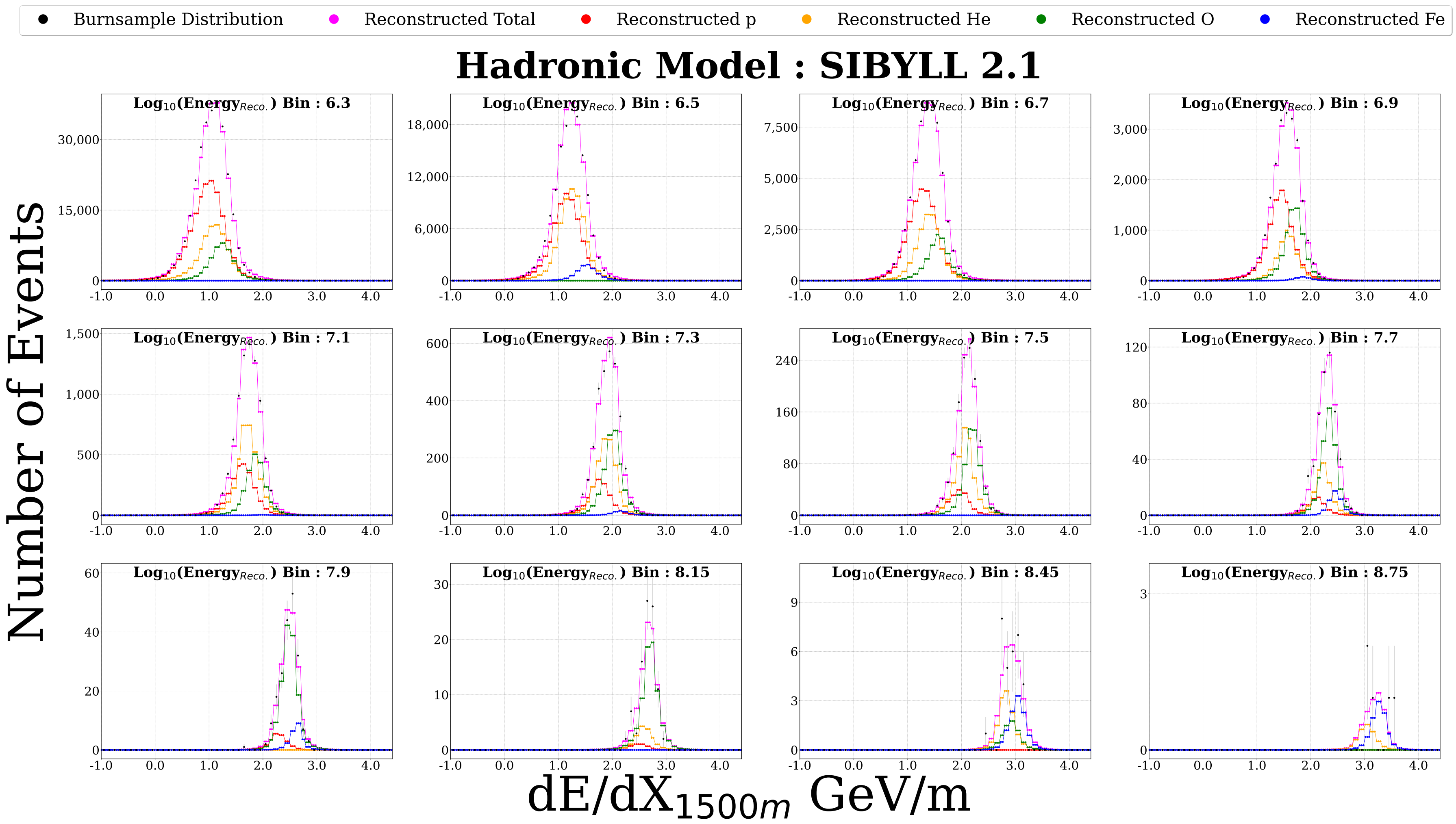


Figure A.19: KDE-template fits for burnsample, using MC-distributions of $\text{Log}_{10}\left(\frac{\text{dE}}{\text{dX}_{1500\text{m}}}/\frac{\text{GeV}}{\text{m}}\right)$ (discussed in Section 7.1), and using SIBYLL 2.1 as the hadronic interaction model for template creation. For details read Section A.8.

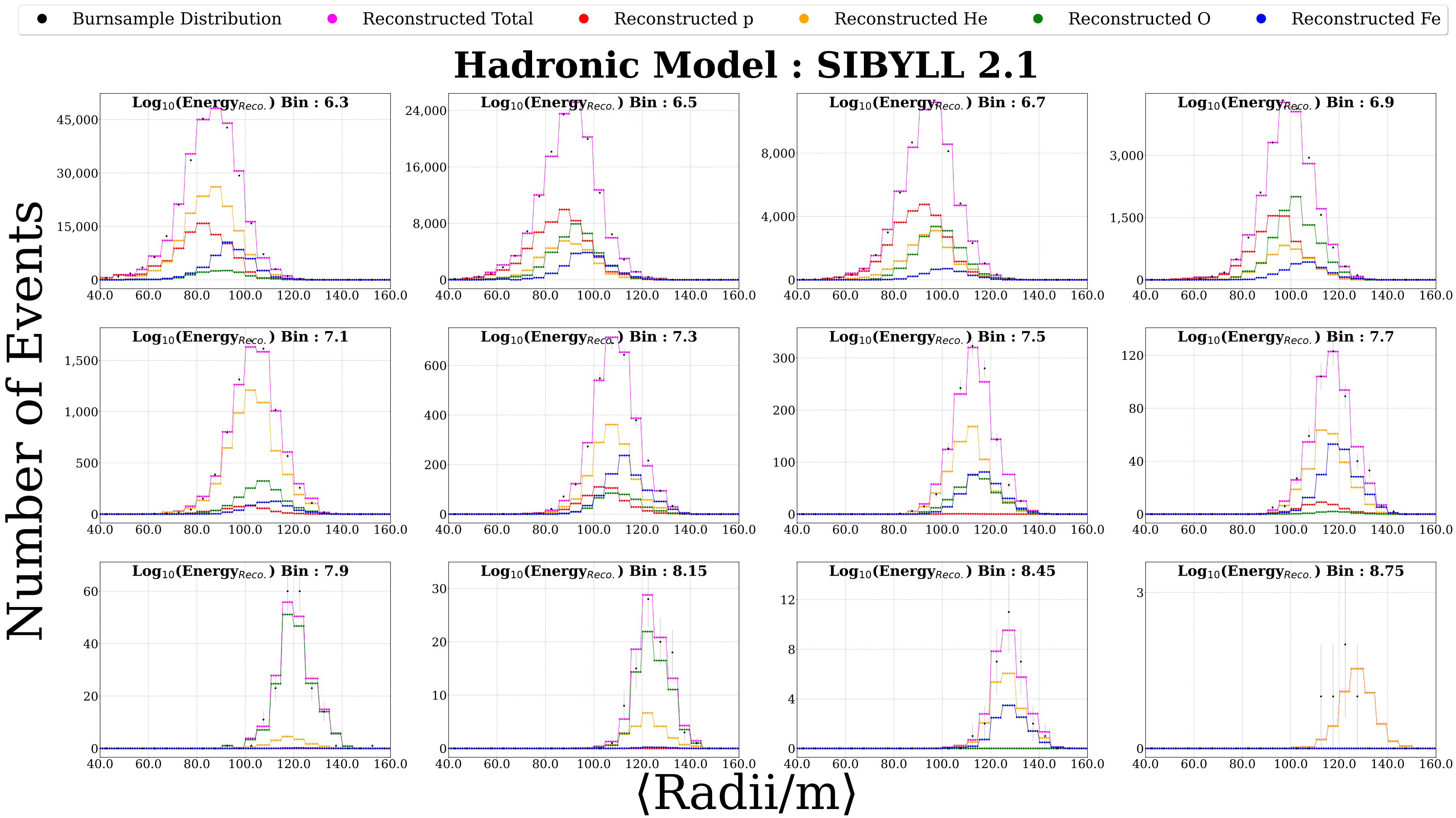


Figure A.20: KDE-template fits for burnsample, using MC-distributions of Mean-Radii(with a maximum DOM-radii cut of 200 m - discussed in Section 7.4.1), and using SIBYLL 2.1 as the hadronic interaction model for template creation. For details read Section A.8.

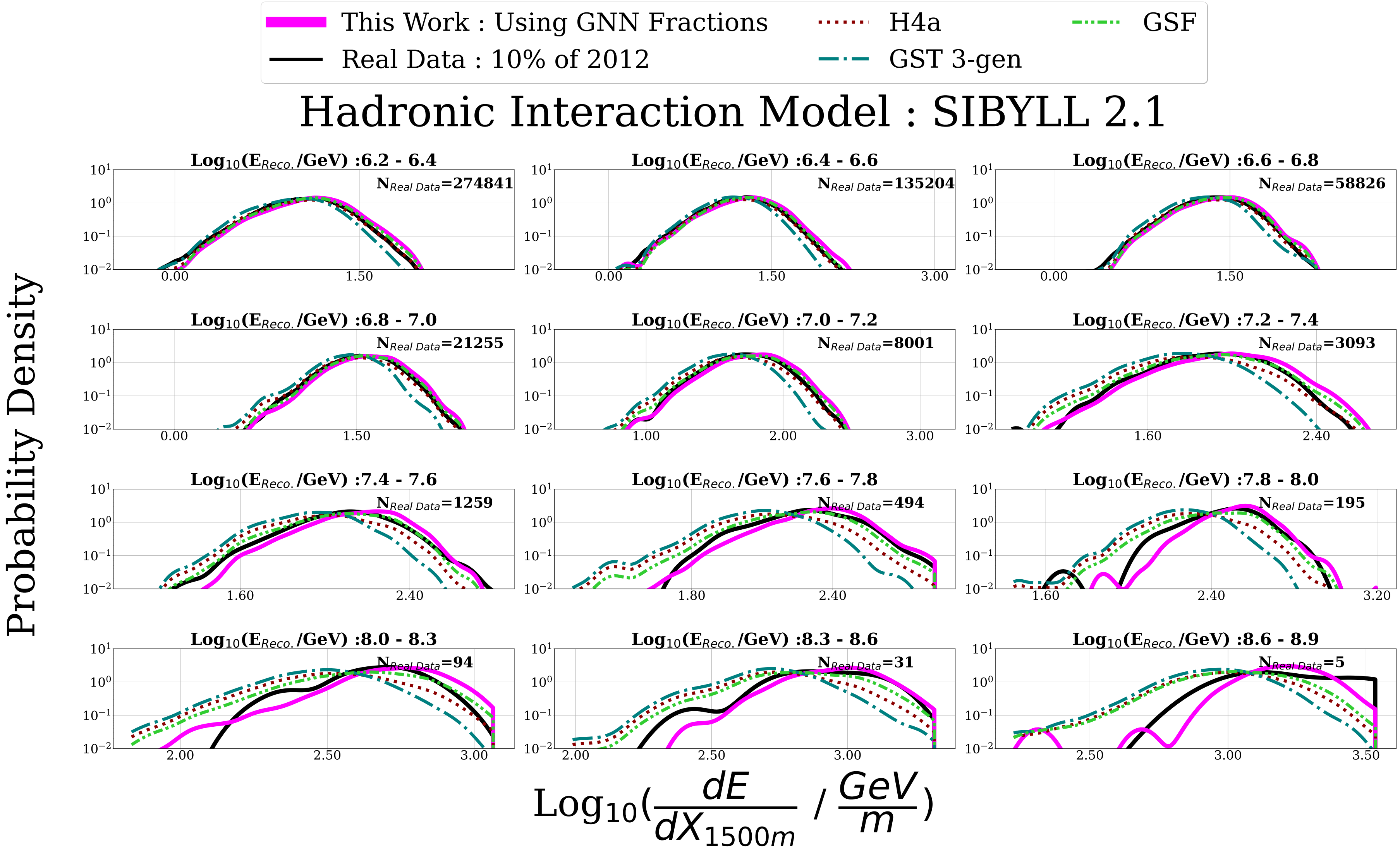


Figure A.21: Compatibility check using for composition-fractions obtained from GNN (details in Section 9.1) for $\mathrm{Log}_{10}\left(\frac{\mathrm{dE}}{\mathrm{dX}_{1500\mathrm{m}}}/\frac{\mathrm{GeV}}{\mathrm{m}}\right)$ (discussed in Section 7.1) and comparison with fractions from other flux-models, on a *Log-scale*. To see the plot on *Linear-scale*, see Figure 9.6.

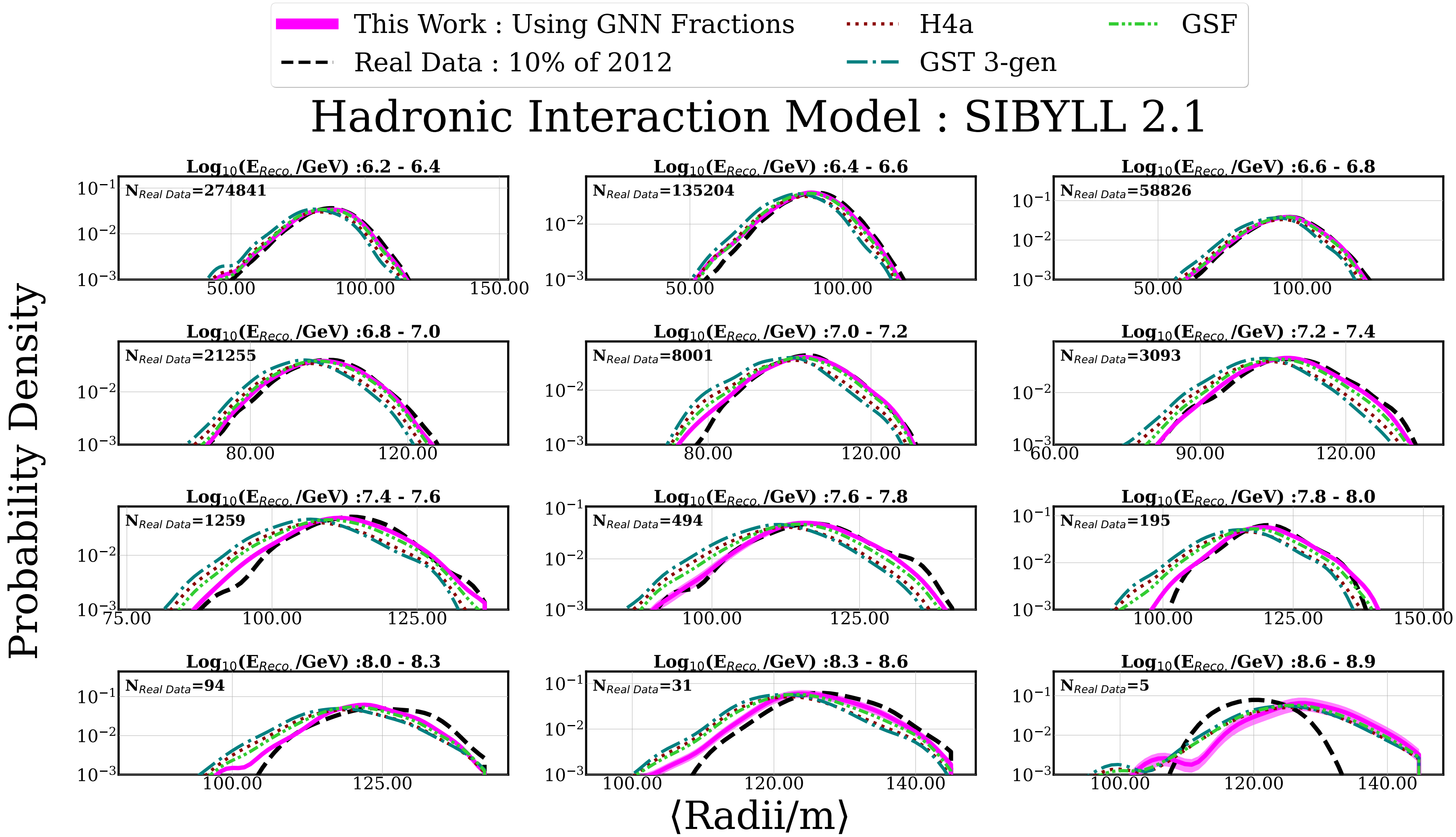


Figure A.22: Compatibility check using for composition-fractions obtained from GNN (details in Section 9.1) for *Mean-Radii* (with maximum DOM-radii cut of 200 m - discussed in Section 7.4.1) and comparison with fractions from other flux-models, on a *Log-scale*. To see the plot on *Linear-scale*, see Figure 9.7.

B DETECTOR

B.1 MILLIPEDE

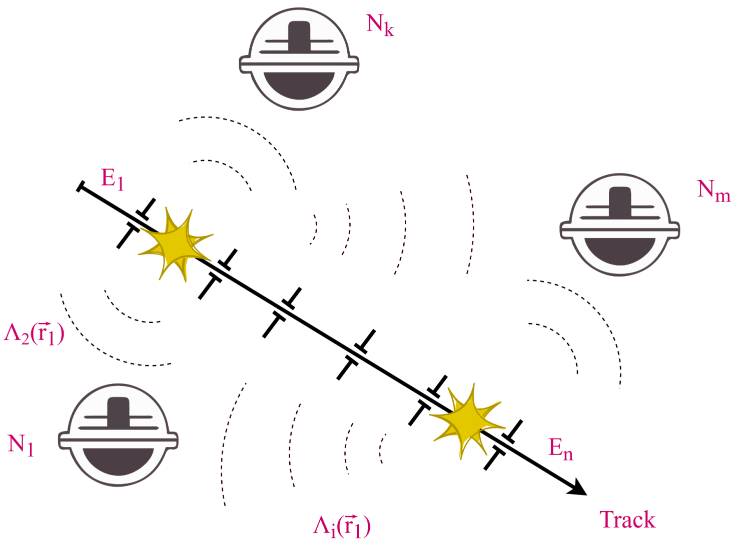


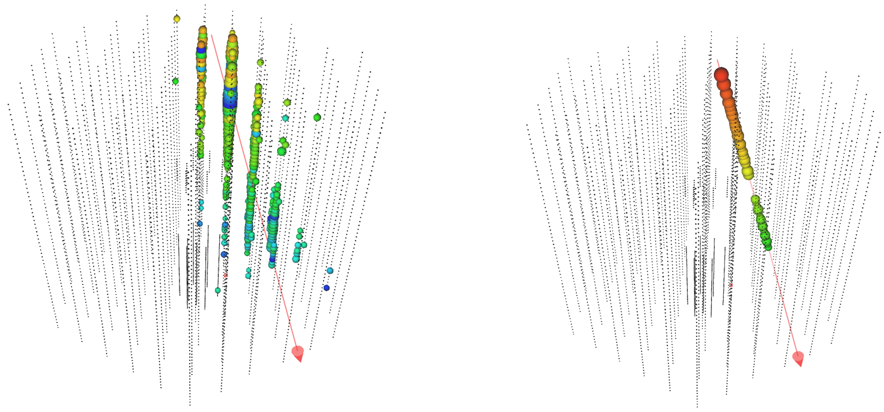
Figure B.1: Energy loss along the segmented track. N_k denotes the number of photons measured at the k^{th} DOM. The yellow stars denote energy deposits. $\Lambda(\vec{r_k}, \vec{r_i})$ is the light-yield by the energy deposit.

The energy loss in a single track segment can contribute to charge deposit at multiple DOMs. Correspondingly the charge deposit at a single DOM is an accumulation of contributions from different segments. The combination of light deposits at a DOM from energy deposit along different track segments is linear [655]. Hence an unfolding can be performed, to reconstruct the energy loss in the segments. At ICNO, it is performed using a framework termed as **MILLIPEDE**[1]. This is illustrated in Figure B.1 and is given by:

$$N_k = \rho + \sum_{i=1}^{n} \Lambda(\vec{r_k}, \vec{r_i}) \cdot E_i \tag{B.1}$$

where N_k denotes the total number of measured photons at the k^{th} DOM, located at $\vec{r_k}$. ρ denotes the average expectation for the noise-induced charge measured at a DOM. The $\Lambda(\vec{r_k}, \vec{r_i}) \cdot E_i$ term denotes the contribution to the total number of photons by the E_i energy-deposit at $\vec{r_i}$. Our knowledge of the detector is captured by $\Lambda(\vec{r_k}, \vec{r_i})$. The light-yield by the energy deposit is encapsulated within $\Lambda(\vec{r_k}, \vec{r_i})$. The term also

1 Named as such because of the resemblance of segmented energy-reconstruction along the track with the segmented body of a millipede.



(a) Charge footprint

(b) Unfolded Energy Loss

Figure B.2: Charge footprint and unfolded energy loss (using MILLIPEDE) of the in-ice detector response for a p-initiated shower, with MC energy of 1.1 PeV and $(\theta, \phi) = (15.6°, 219.8°)$. The size of spheres is indicative of the amount of charge deposit (at DOM) and unfolded energy loss (along the track) respectively. Color represent time (red = early, blue = Late). The red arrow represents the reconstructed track.

captures the angular profile of Cherenkov light radiation, since the direction of the energy deposit is also crucial in determining the photon deposit at a DOM. In addition to this the absorption and scattering of photons during their propagation from $\vec{r_i}$ to $\vec{r_k}$, and DOM efficiency is taken into account. MC-simulations and data from the calibration runs using flasher-LEDs are used to construct a multi-dimensional binned histogram parametrization. A multi-dimensional penalized spline surface technique is then used to efficiently compute interpolated values[2] [404]. Parameterization is one of the most important sources of systematic uncertainties since it is heavily dependent on our knowledge of the detector (scattering coefficient, absorption coefficient, DOM efficiency, etc.). The energy deposit E_i can be obtained using Equation B.1. A non-monotonic Poisson likelihood minimization technique [656] is used to limit the solutions in a phase space of physically possible solutions (e.g. positive energies). [657] discusses another approach to estimating the energy deposit by accounting for our incomplete knowledge of inherently stochastic energy losses. This is done by constructing probability distributions for measuring different numbers of photons. However, it gives the same solution as the one obtained by Equation B.1. Application of MILLIPEDE for an example air-shower is illustrated in Figure B.2.

B.2 DDDDR - DATA DERIVED DIFFERENTIAL DEPOSITION RECONSTRUCTION

Data Derived Differential Deposition Reconstruction (DDDDR) is generally used to characterize the energy loss spectrum of a muon-bundle[3] and was used to understand muon-multiplicity at IceCube[658]. It calculates the differential energy loss along the

2 Currently the parametrization doesn't account for azimuthal anisotropy and for the known tilt in the ice layers.

3 or individual muon

track by considering an exponential light attenuation in bulk ice. The energy deposit along the segmented track is given by:

$$\frac{dE_{\mu-\mathrm{bundle}}}{dX} = \frac{q_{DOM}}{\epsilon_{DOM}} \cdot f_{scale} \cdot \begin{cases} r_0 & \text{for } r < r_0 \\ r \cdot e^{\frac{r-r_0}{\lambda_{att.}(Z)}} & \text{for } r > r_0 \end{cases} \tag{B.2}$$

where q_{DOM} is the charge deposit at a DOM located at a perpendicular distance of r from the track. ϵ is the DOM efficiency. The approximation of exponential decay is not valid closer to the track (and near the dust layer) and is corrected using the depth-dependent r_0-factor = $19m + 0.01 \cdot z$ factor. Here, z is the vertical coordinate (height) measured in IceCube's coordinate (centered at the middle of the IC detector at 1949 meters below the surface). $f_{scale} \approx 0.020\ \mathrm{GeV/PE} \cdot m^2$ is a scaling factor obtained from MC-simulations. $\lambda_{att.}$ is the effective attenuation length. Since, IceCube is not a single transparent homogeneous volume of Ice, $\lambda_{att.}$ shows the dependence of the location in the detector and is obtained for each horizontal layer in the ice[4]. [658] presents the validity of using the reconstruction. This study doesn't use the reconstructed energy deposits. Only the measured charges and the perpendicular distance of DOMs from the track are used. For details of how they are used read Section 7.4.

4 In reality the individual ice-layers are not totally horizontal everywhere.

C MACHINE LEARNING

C.1 WEISFEILER-LEHMAN (WL) TEST

WL Test produces a representation (represented by color later) for a graph by an iterative approach and then compares it with the representation of the other graph. If the representation of the graphs are not same, they are non-isomorphic and hence distinguishable. However, two distinguishable graphs can have the same canonical representation. The steps involved in getting the canonical form are namely[1]:

- Color (or label) each node in a graph with the same color.
- Collect the colors from the node and the adjacent nodes, into a multiset of colors. Hash the colors into a new color.
- Repeat the earlier steps until a stable coloring is obtained.
- Repeat the earlier steps for the other graph and compare the coloring with the other graph.

If the coloring of the two graphs is different is different, the graphs are non-isomorphic. However, the same can't be claimed vice-versa.

C.2 CNN @ ICECUBE

Most DL-analysis at IC are currently reliant on using Convolutional Neural Networks (CNNs). For this they split the in-ice detector into 3 components. These are the "Main-array", "Lower DeepCore" and "Upper DeepCore". These are then transformed into an orthogonal geometry, allowing them to be used as input for the NN. This is depicted in Figure C.1.

C.3 HOMOPHILY

As described in Section 8.1.3, Homophily is a quantitative measure of the probability of nodes (in a graph), sharing the same labels, and being located in close proximity to one another within the graph. Section 8.1.3 also presents the measure of Homophily for multiple node features used in the work. Figure C.2 and Figure C.3 present the Homophily for "Charge weighted y coordinate" and "Total Charge" (with increased number of connections in comparison to those shown in Section 8.1.3), as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$. For details read Section 8.1.3.

1 The following points are adapted from Medium.com article written by Michael Bronstein titled "Expressive power of graph neural networks and the Weisfeiler-Lehman test".

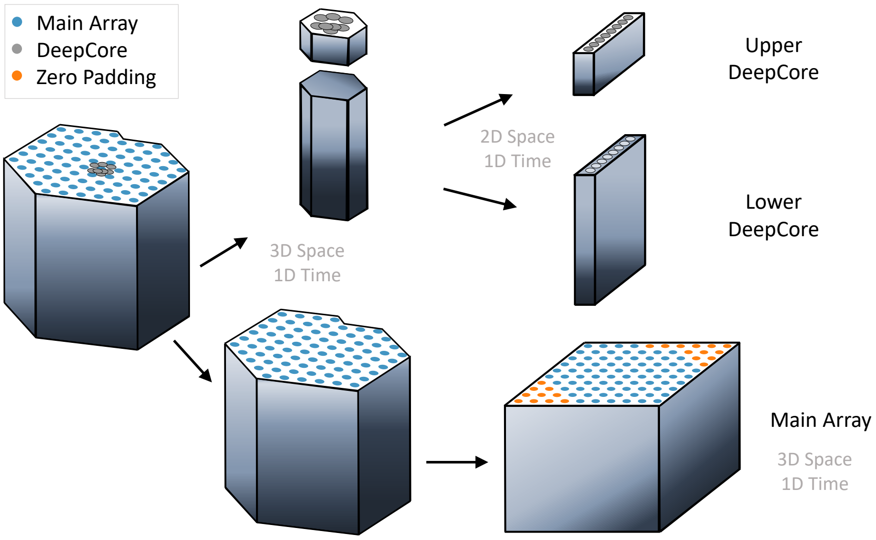


Figure C.1: Division of IceCube array into the main array and two components of DeepCore for CNN-based implementation at IceCube. This is done to treat different detector geometries separately. The transformation from hexagonal to orthogonal geometry (by padding with zeros - orange dots) is also shown. Illustration from [198].

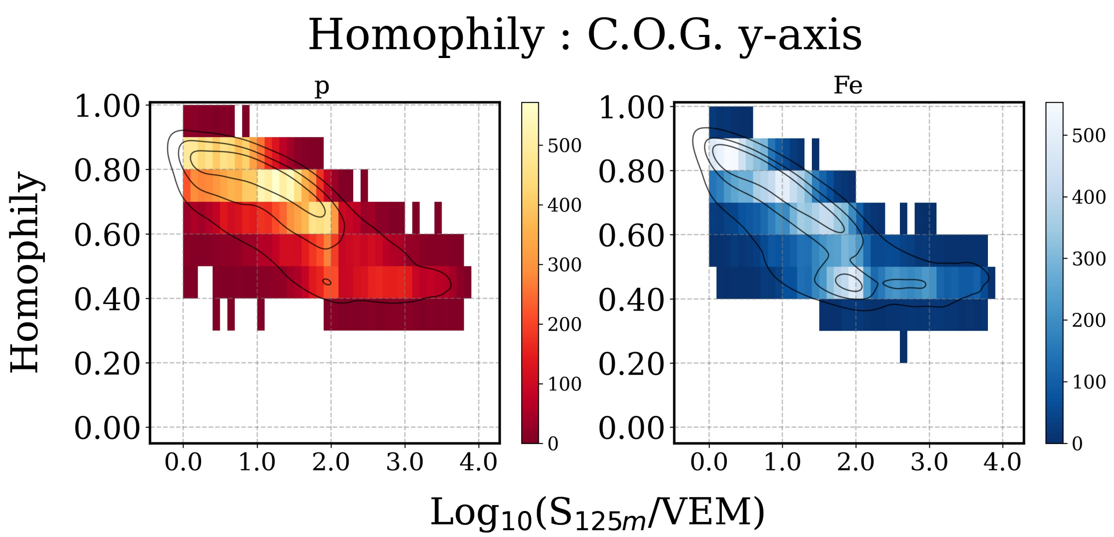


Figure C.2: Homophily for Charge weighted y coordinate (with the coordinate-origin at the charge-weighted COG) as a function of $\mathrm{Log}_{10}(S_{125}/\mathrm{VEM})$, for p and Fe. The maximum number of connections any node (DOM) in the dataset can make (with neighboring nodes) is 10. The underlying events are from the training dataset (unweighted). The color represents the counts in a bin. The overlayed black curves represent the 2D KDEs. For interpretation read Section 8.1.3.

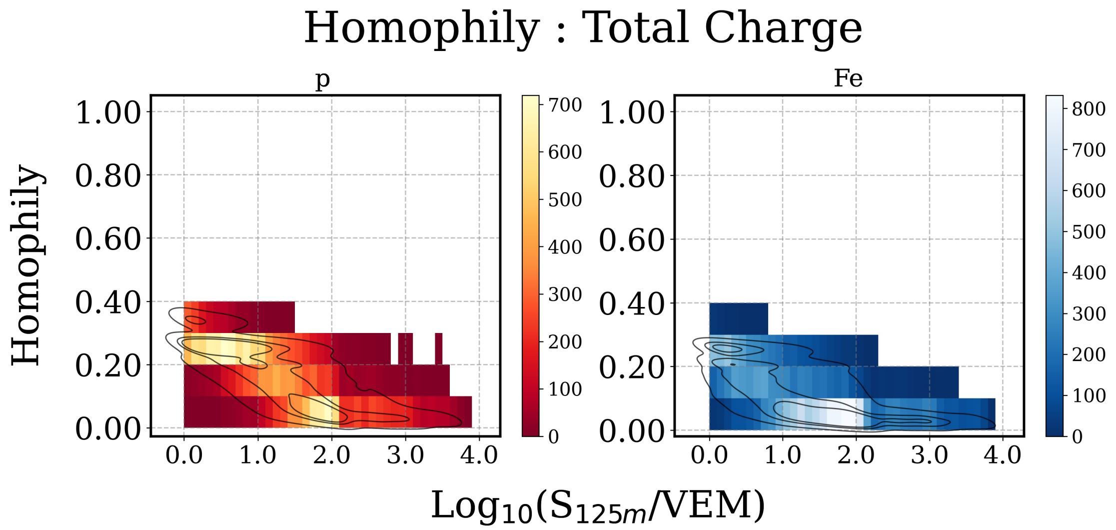


Figure C.3: Homophily for Total Charge as a function of $\text{Log}_{10}(S_{125}/\text{VEM})$, for p and Fe. The maximum number of connections any node (DOM) in the dataset can make (with neighboring nodes) is 40. The underlying events are from the training dataset (unweighted). The color represents the counts in a bin. The overlayed black curves represent the 2D KDEs. For interpretation read Section 8.1.3.

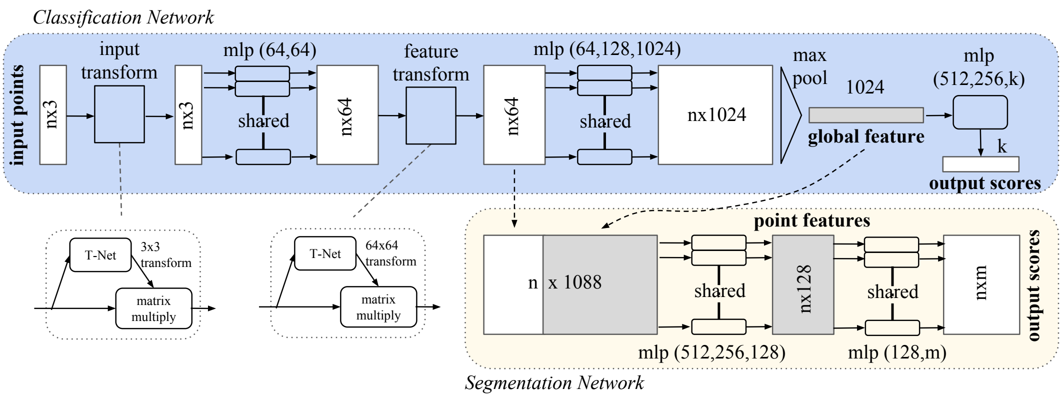


Figure C.4: PointNet Architecture for classification and segmentation. Illustration from [599].

C.4 POINTNET

PointNet is a DL-architecture designed to work with point clouds for the purpose of object classification, and/or image segmentation[599]. The PointNet architecture can be seen in Figure C.4. Similar to any standard GNN, the architecture also has permutational-invariance of node-labeling encoded into it. In addition, transformation-invariance (rotation and translation) is also encoded into the architecture. Read [599] for the details of the architecture.

C.5 TRANSFORMER AND ATTENTION

As mentioned in Section 8.1.5 the effectiveness of the Transformer based model can be attributed to their underlying Attention Mechanism. Transformers were introduced to replace Recurrent Neural Networks (RNNs). The following is a list of issues which RNNs face and mentions how Attention Mechanism (in Transformers) allays those[2]:

- For a useful language-based model it is necessary that the model can properly work with long text input/output. RNNs suffer with understanding or producing long-term dependencies [659][3]. This makes the RNN-based language models fairly useless for most practical purposes. Transformers prevent this by the use of attention. The attention mechanism allows for learning dependencies or correlation between far-off segments in input with the same likelihood as if they were located closer i.e. Attention Mechanism allows the network to only choose the dependencies which are beneficial to network's performance. This allows for the network to learn dependencies of long as well short-range.

- Almost all kind of neural-networks use back-propagation [440] for training. It upgrades the gradients associated with the learnable parameters during the backward pass of a neural network. RNNs suffer from vanishing and exploding gradients [473] making the training process very difficult. Since Transformers, rely on working will all segments on the input at the same time (in contrast to the recursive manner of RNNs), gradients are generally well-behaved.

- Because of the recursive nature of the processing, the parallel computation cannot be done easily on the input in RNNs. This slows down the training process significantly. Transformers don't suffer from any such problem.

2 The following points are adapted from Medium.com article written by Arjun Sarkar titled "All you need to know about 'Attention' and 'Transformers'"

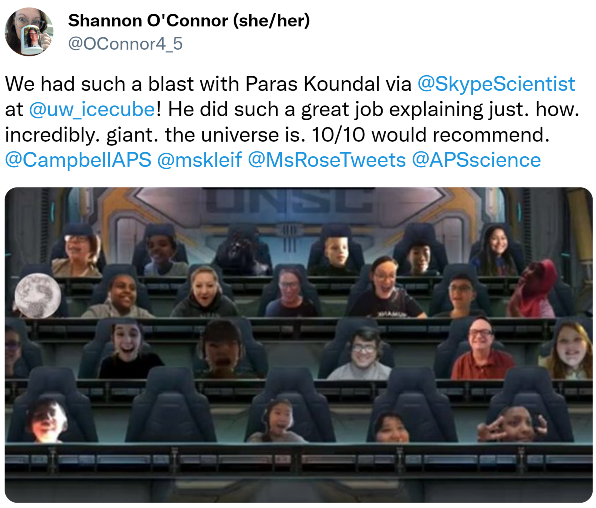
3 Vanishing gradient during back-propagation are primarily the reason for this. Long Short Term Memory (LSTM) networks were proposed to solve this [659]. However, LSTM also compress the information, from dependencies at all different steps/times, into a fixed vector, making the task of understanding long-term dependencies difficult

D SCIENCE COMMUNICATION

Over the course of my doctoral degree, I have also enjoyed communicating science with a variety of audiences. It has not only given me immense pleasure but has also helped me train to turn abstract theories into tangible narratives. Some of these experiences are namely:

- **International Cosmic Day, 2020:** I was involved in discussing the amazing science of cosmic-rays (CRs), and Python Coding with a few school students from various countries in Europe. It was a very enriching experience.

- **Skype A Scientist**: I was involved with the "Skype a Scientist" platform and I gave a small introductory talk about astrophysics and the IceCube Observatory to fourth-grade students in a school in U.S.A., in January 2021. It was a wonderful experience and was also appreciated by the kids and the teacher.



- **Nishtha Sahni:** She was a student of IIT Bombay, India, and was selected by KIT as a summer-intern student under IAESTE scholarship from DAAD. She worked remotely in the Summer of 2021 using open data from the KASCADE-experiment available at KCDC platform [50] and used neural networks for CR-composition estimate. Later in December 2021, she was at Karlsruhe, where she worked in close collaboration with my colleague to test the viability of tree-based methods to improve cosmic-ray analysis at IceCube. She presented her work in internal meetings at KIT. She will soon be joining the University of Delaware as a Doctoral Researcher.

- **Mathias Hilfiker:** He was a student from the University of Torino and was at KIT through the ERASMUS program. We checked the feasibility of neural networks-based methods for composition analysis at IceCube Observatory. He presented his work in internal meetings at KIT. He is currently a Doctoral Researcher at Université du Luxembourg, working closely with AstraZeneca for innovative drug discovery using Quantum-Machine-Learning force fields for studying molecule-protein interactions.

- **Miro Joensuu:** He was a school student at the time, working on a project at CERN organized by the German organization Netzwerk Teilchenwelt. I supervised him remotely on a cosmic-ray analysis. He presented his work in a meeting at CERN. He is currently a student at Universität Heidelberg, Heidelberg.

- **Harsh Choudhary:** He was a student of IIT Bombay, India, and was selected by KIT as a summer-intern student under IAESTE scholarship from DAAD. He worked closely with me testing and developing some key methods in Machine Learning, which although not used in this thesis hold strong potential for developments in the field of Machine Learning. The ideas will be pursued in the future. He will soon start as a Doctoral researcher at Czech Technical University in Prague. He will work on Physics-informed neural networks.

- **Others:** During my doctoral degree I have had many opportunities to present my work publicly. I have dearly enjoyed each of these experiences.

BIBLIOGRAPHY

[1] William Shakespeare. *Twelfth Night: or, What You Will*. URL: https://www.gutenberg.org/files/1526/1526-h/1526-h.htm.

[2] Viktor F Hess. "Uber Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten." *Phys. Zeits.* 13 (1912), pp. 1084–1091. URL: https://inspirehep.net/literature/1623161.

[3] Werner Kolhörster. "Messungen der durchdringenden Strahlung im Freiballon in grosseren Hohen." *Phys. Zeits.* 14 (1913), pp. 1153–1156. URL: https://babel.hathitrust.org/cgi/pt?id=mdp.39015021268936.

[4] Jörg R. Hörandel. "Early cosmic-ray work published in German." *AIP Conference Proceedings* 1516.1 (Feb. 2013), pp. 52–60. DOI: 10.1063/1.4792540.

[5] Walther Bothe and Werner Kolhörster. "Das Wesen der Höhenstrahlung." *Zeitschrift für Physik* 56.11 (1929), pp. 751–777. DOI: 10.1007/bf01340137.

[6] Jacob Clay. "Ultra radiation (penetrating radiation) III: Annual variation and variation with the geographical latitude." *Proc. R. Acad. Amsterdam*. Vol. 33. 1930, pp. 711–718.

[7] Marcel Schein, William P. Jesse, and E. O. Wollan. "The Nature of the Primary Cosmic Radiation and the Origin of the Mesotron." *Physical Review* 59.7 (Apr. 1941), pp. 615–615. DOI: 10.1103/physrev.59.615.

[8] Péter Mészáros, Derek B. Fox, Chad Hanna, and Kohta Murase. "Multi-Messenger Astrophysics." *Nature Rev. Phys.* 1 (2019), pp. 585–599. DOI: 10.1038/s42254-019-0101-z.

[9] Stef Verpoest. "Multiplicity of TeV muons in air showers measured with IceTop and IceCube." 423 (2023), p. 074. DOI: 10.22323/1.423.0074.

[10] A. Gide and P. Louÿs. *Le traité du Narcisse (théorie du symbole)*. Librairie de l'art indépndant, 1891. URL: https://books.google.de/books?id=vkqbjgEACAAJ.

[11] AC Cummings et al. "Galactic cosmic rays in the local interstellar medium: Voyager 1 observations and model results." *The Astrophysical Journal* 831.1 (2016), p. 18. DOI: 10.3847/0004-637x/831/1/18.

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

[13] Carl D. Anderson. "The Positive Electron." *Phys. Rev.* 43 (6 1933), pp. 491–494. DOI: 10.1103/PhysRev.43.491.

[14] Michael Price. *Marking Time*. 2023. URL: https://www.science.org/content/article/marking-time-cosmic-ray-storms-can-pin-precise-dates-history-ancient-egypt-vikings (visited on 06/11/2023).

[15] Taylor Kubota. *Cosmic rays may have left indelible imprint on early life*. 2020. URL: https://news.stanford.edu/2020/05/20/cosmic-rays-may-shaped-life/ (visited on 06/11/2023).

[16] Dimitra Atri. "Investigating the biological potential of galactic cosmic ray-induced radiation-driven chemical disequilibrium in the Martian subsurface environment." *Scientific Reports* 10.1 (July 2020). DOI: 10.1038/s41598-020-68715-7.

[17] Christoph Welling, Philipp Frank, Torsten A. Enßlin, and Anna Nelles. "Reconstructing non-repeating radio pulses with Information Field Theory." *JCAP* 04 (2021), p. 71. DOI: 10.1088/1475-7516/2021/04/071.

[18] Valerio Verzi. "Measurement of the energy spectrum of ultra-high energy cosmic rays using the Pierre Auger Observatory." *Proceedings of 36th International Cosmic Ray Conference — PoS(ICRC2019)*. Vol. 358. 2019, p. 450. DOI: 10.22323/1.358.0450.

[19] R. Alfaro et al. "All-particle cosmic ray energy spectrum measured by the HAWC experiment from 10 to 500 TeV." *Phys. Rev. D* 96.12 (2017), p. 122001. DOI: 10.1103/PhysRevD.96.122001.

[20] M. G. Aartsen et al. "Cosmic ray spectrum and composition from PeV to EeV using 3 years of data from IceTop and IceCube." *Phys. Rev. D* 100.8 (2019), p. 082002. DOI: 10.1103/PhysRevD.100.082002.

[21] Marcel René Finger. "Reconstruction of energy spectra for different mass groups of high-energy cosmic rays." PhD thesis. Karlsruher Institut für Technologie (KIT), 2011. URL: https://d-nb.info/1014279917/34.

[22] Mario E. Bertaina. "KASCADE-Grande energy spectrum of cosmic rays interpreted with post-LHC hadronic interaction models." *Proceedings of The 34th International Cosmic Ray Conference — PoS(ICRC2015)*. Vol. 236. 2016, p. 359. DOI: 10.22323/1.236.0359.

[23] Nikolay Gorbunov et al. "Energy spectra of abundant cosmic-ray nuclei in the NUCLEON experiment." *Adv. Space Res.* 64.12 (2019), pp. 2546–2558. DOI: 10.1016/j.asr.2019.10.004.

[24] John Matthews. "Highlights from the Telescope Array Experiment." *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Vol. 301. 2017, p. 1096. DOI: 10.22323/1.301.1096.

[25] M. Amenomori et al. "The All-particle spectrum of primary cosmic rays in the wide energy range from 10^{14} eV to 10^{17} eV observed with the Tibet-III air-shower array." *Astrophys. J.* 678 (2008), pp. 1165–1179. DOI: 10.1086/529514.

[26] N. M. Budnev et al. "The primary cosmic-ray energy spectrum measured with the Tunka-133 array." *Astropart. Phys.* 117 (2020), p. 102406. DOI: 10.1016/j.astropartphys.2019.102406.

[27] M. Aguilar et al. "Precision Measurement of the Proton Flux in Primary Cosmic Rays from Rigidity 1 GV to 1.8 TV with the Alpha Magnetic Spectrometer on the International Space Station." *Phys. Rev. Lett.* 114 (2015), p. 171103. DOI: 10.1103/PhysRevLett.114.171103.

[28] Y. Shikaze et al. "Measurements of 0.2 to 20-GeV/n cosmic-ray proton and helium spectra from 1997 through 2002 with the BESS spectrometer." *Astropart. Phys.* 28 (2007), pp. 154–167. DOI: 10.1016/j.astropartphys.2007.05.001.

[29] O. Adriani et al. "Direct Measurement of the Cosmic-Ray Proton Spectrum from 50 GeV to 10 TeV with the Calorimetric Electron Telescope on the International Space Station." *Phys. Rev. Lett.* 122.18 (2019), p. 181102. DOI: 10.1103/PhysRevLett.122.181102.

[30] Y. S. Yoon et al. "Proton and Helium Spectra from the CREAM-III Flight." *Astrophys. J.* 839.1 (2017), p. 5. DOI: 10.3847/1538-4357/aa68e4.

[31] Q. An et al. "Measurement of the cosmic-ray proton spectrum from 40 GeV to 100 TeV with the DAMPE satellite." *Sci. Adv.* 5.9 (2019). DOI: 10.1126/sciadv.aax3793.

[32] Carlos Juan Arteaga-Velázquez et al. "Measurements of the muon content of EAS in KASCADE-Grande compared with SIBYLL 2.3 predictions." *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Vol. 301. 2017, p. 316. DOI: 10.22323/1.301.0316.

[33] T. Antoni et al. "KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems." *Astropart. Phys.* 24 (2005), pp. 1–25. DOI: 10.1016/j.astropartphys.2005.04.001.

[34] O. Adriani et al. "PAMELA Measurements of Cosmic-ray Proton and Helium Spectra." *Science* 332 (2011), pp. 69–72. DOI: 10.1126/science.1199172.

[35] M. Aguilar et al. "Towards Understanding the Origin of Cosmic-Ray Electrons." *Phys. Rev. Lett.* 122.10 (2019), p. 101101. DOI: 10.1103/PhysRevLett.122.101101.

[36] O. Adriani et al. "Extended Measurement of the Cosmic-Ray Electron and Positron Spectrum from 11 GeV to 4.8 TeV with the Calorimetric Electron Telescope on the International Space Station." *Phys. Rev. Lett.* 120.26 (2018), p. 261102. DOI: 10.1103/PhysRevLett.120.261102.

[37] G. Ambrosi et al. "Direct detection of a break in the teraelectronvolt cosmic-ray spectrum of electrons and positrons." *Nature* 552 (2017), pp. 63–66. DOI: 10.1038/nature24475.

[38] S. Abdollahi et al. "Cosmic-ray electron-positron spectrum from 7 GeV to 2 TeV with the Fermi Large Area Telescope." *Phys. Rev. D* 95.8 (2017), p. 082007. DOI: 10.1103/PhysRevD.95.082007.

[39] A. Archer et al. "Measurement of Cosmic-ray Electrons at TeV Energies by VERITAS." *Phys. Rev. D* 98.6 (2018), p. 062004. DOI: 10.1103/PhysRevD.98.062004.

[40] M. Aguilar et al. "Antiproton Flux, Antiproton-to-Proton Flux Ratio, and Properties of Elementary Particle Fluxes in Primary Cosmic Rays Measured with the Alpha Magnetic Spectrometer on the International Space Station." *Phys. Rev. Lett.* 117.9 (2016), p. 091103. DOI: 10.1103/PhysRevLett.117.091103.

[41] K. Abe et al. "Measurement of the cosmic-ray antiproton spectrum at solar minimum with a long-duration balloon flight over Antarctica." *Phys. Rev. Lett.* 108 (2012), p. 051102. DOI: 10.1103/PhysRevLett.108.051102.

[42] O. Adriani et al. "Measurement of the flux of primary cosmic ray antiprotons with energies of 60-MeV to 350-GeV in the PAMELA experiment." *Pisma Zh. Eksp. Teor. Fiz.* 96 (2012), pp. 693–699. DOI: 10.1134/S002136401222002X.

[43] O. Adriani et al. "Cosmic-Ray Positron Energy Spectrum Measured by PAMELA." *Phys. Rev. Lett.* 111 (2013), p. 081102. DOI: 10.1103/PhysRevLett.111.081102.

[44] M. Ackermann et al. “Measurement of separate cosmic-ray electron and positron spectra with the Fermi Large Area Telescope.” *Phys. Rev. Lett.* 108 (2012), p. 011103. DOI: 10.1103/PhysRevLett.108.011103.

[45] R. Abbasi et al. “The IceCube high-energy starting event sample: Description and flux characterization with 7.5 years of data.” *Phys. Rev. D* 104 (2021), p. 022002. DOI: 10.1103/PhysRevD.104.022002.

[46] Mattia Di Mauro. “The origin of the Fermi-LAT γ-ray background.” *14th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories*. Vol. 3. 2017, pp. 3098–3104. DOI: 10.1142/9789813226609_0393.

[47] D. Maurin, F. Melot, and R. Taillet. “A database of charged cosmic rays.” *Astron. Astrophys.* 569 (2014), A32. DOI: 10.1051/0004-6361/201321344.

[48] D. Maurin, H. Dembinski, J. Gonzalez, Ioana C. Maris, and F. Melot. “Cosmic-Ray Database Update: Ultra-High Energy, Ultra-Heavy, and Antinuclei Cosmic-Ray Data (CRDB v4.0).” *Universe* 6.8 (2020), p. 102. DOI: 10.3390/universe6080102.

[49] Cecilia Pizzolotto, Valeria Di Felice, Domenico D’Urso, Simonetta Dari, Daniele Navarra, Roberto Primavera, and Bruna Bertucci. “Looking for cosmic ray data? The ASI Cosmic Ray Database.” *PoS* ICRC2017 (2018), p. 227. DOI: 10.22323/1.301.0227.

[50] A. Haungs et al. “The KASCADE Cosmic-ray Data Centre KCDC: Granting Open Access to Astroparticle Physics Research Data.” *Eur. Phys. J. C* 78.9 (2018), p. 741. DOI: 10.1140/epjc/s10052-018-6221-2.

[51] Carmelo Evoli. *"The Cosmic-Ray Energy Spectrum"*. Dec. 2020. DOI: 10.5281/zenodo.4396125.

[52] W. D. Apel et al. “Kneelike structure in the spectrum of the heavy component of cosmic rays observed with KASCADE-Grande.” *Phys. Rev. Lett.* 107 (2011), p. 171104. DOI: 10.1103/PhysRevLett.107.171104.

[53] T Antoni et al. “The Cosmic ray experiment KASCADE.” *Nucl. Instrum. Meth. A* 513 (2003), pp. 490–510. DOI: 10.1016/S0168-9002(03)02076-X.

[54] D. J. Bird et al. “Evidence for correlated changes in the spectrum and composition of cosmic rays at extremely high-energies.” *Phys. Rev. Lett.* 71 (1993), pp. 3401–3404. DOI: 10.1103/PhysRevLett.71.3401.

[55] R. U. Abbasi et al. “First observation of the Greisen-Zatsepin-Kuzmin suppression.” *Phys. Rev. Lett.* 100 (2008), p. 101101. DOI: 10.1103/PhysRevLett.100.101101.

[56] J. Abraham et al. “Measurement of the Energy Spectrum of Cosmic Rays above 10^{18} eV using the Pierre Auger Observatory.” *Phys. Lett. B* 685 (2010), pp. 239–246. DOI: 10.1016/j.physletb.2010.02.013.

[57] W. D. Apel et al. “The spectrum of high-energy cosmic rays measured with KASCADE-Grande.” *Astropart. Phys.* 36 (2012), pp. 183–194. DOI: 10.1016/j.astropartphys.2012.05.023.

[58] M. G. Aartsen et al. “Measurement of the cosmic ray energy spectrum with IceTop-73.” *Phys. Rev. D* 88.4 (2013), p. 042004. DOI: 10.1103/PhysRevD.88.042004.

[59] S. P. Knurenko, Z. E. Petrov, R. Sidorov, I. Ye. Sleptsov, S. K. Starostin, and G. G. Struchkov. "Cosmic ray spectrum in the energy range 10^{15}-10^{18} eV and the second knee according to the small Cherenkov setup at the Yakutsk EAS array" (Oct. 2013). arXiv: 1310.1978 [astro-ph.HE].

[60] V. V. Prosin et al. "Tunka-133: Results of 3 year operation." *Nucl. Instrum. Meth. A* 756 (2014), pp. 94–101. DOI: 10.1016/j.nima.2013.09.018.

[61] T. Abu-Zayyad et al. "Measurement of the cosmic ray energy spectrum and composition from 10^{17} to $10^{18.3}$ eV using a hybrid technique." *Astrophys. J.* 557 (2001), pp. 686–699. DOI: 10.1086/322240.

[62] Douglas R. Bergman and John W. Belz. "Cosmic Rays: The Second Knee and Beyond." *J. Phys. G* 34 (2007), R359. DOI: 10.1088/0954-3899/34/10/R01.

[63] Dmitri Ivanov. "Energy Spectrum Measured by the Telescope Array." *Proceedings of 36th International Cosmic Ray Conference — PoS(ICRC2019)*. Vol. 358. 2019, p. 298. DOI: 10.22323/1.358.0298.

[64] Kenneth Greisen. "End to the cosmic ray spectrum?" *Phys. Rev. Lett.* 16 (1966), pp. 748–750. DOI: 10.1103/PhysRevLett.16.748.

[65] G. T. Zatsepin and V. A. Kuzmin. "Upper limit of the spectrum of cosmic rays." *JETP Lett.* 4 (1966), pp. 78–80. URL: http://jetpletters.ru/ps/1624/index.shtml.

[66] B. Peters. "Primary cosmic radiation and extensive air showers." *Il Nuovo Cimento* 22.4 (Nov. 1961), pp. 800–819. DOI: 10.1007/bf02783106.

[67] Philip Morrison. "On the Origins of Cosmic Rays." *Reviews of Modern Physics* 29.2 (Apr. 1957), pp. 235–243. DOI: 10.1103/revmodphys.29.235.

[68] G. Cocconi. "Intergalactic space and cosmic rays." *Il Nuovo Cimento* 3.6 (June 1956), pp. 1433–1442. DOI: 10.1007/bf02785019.

[69] Alexander Aab et al. "Observation of a Large-scale Anisotropy in the Arrival Directions of Cosmic Rays above 8×10^{18} eV." *Science* 357.6537 (2017), pp. 1266–1270. DOI: 10.1126/science.aan4338.

[70] G. Giacinti, M. Kachelrieß, and D. V. Semikoz. "Explaining the Spectra of Cosmic Ray Groups above the Knee by Escape from the Galaxy." *Phys. Rev. D* 90.4 (2014), p. 041302. DOI: 10.1103/PhysRevD.90.041302.

[71] A. D. Erlykin and A. W. Wolfendale. "A single source of cosmic rays in the range 10^{15}-eV to 10^{16}-eV." *J. Phys. G* 23 (1997), pp. 979–989. DOI: 10.1088/0954-3899/23/8/012.

[72] A. De Rujula. "A solution of a hoary conundrum: The Origin and properties of cosmic rays." *Nucl. Phys. B Proc. Suppl.* 165 (2007), pp. 93–102. DOI: 10.1016/j.nuclphysbps.2006.11.015.

[73] VS Ptuskin et al. "Diffusion and drift of very high energy cosmic rays in galactic magnetic fields." *Astronomy and Astrophysics* 268 (1993), pp. 726–735. URL: https://adsabs.harvard.edu/full/1993A%26A...268..726P.

[74] Julian Candia, Esteban Roulet, and Luis N. Epele. "Turbulent diffusion and drift in galactic magnetic fields and the explanation of the knee in the cosmic ray spectrum." *JHEP* 12 (2002), p. 033. DOI: 10.1088/1126-6708/2002/12/033.

[75] David d'Enterria, Ralph Engel, Tanguy Pierog, Sergey Ostapchenko, and Klaus Werner. "Constraints from the first LHC data on hadronic event generators for ultra-high energy cosmic-ray physics." *Astropart. Phys.* 35 (2011), pp. 98–113. DOI: 10.1016/j.astropartphys.2011.05.002.

[76] Rafael Alves Batista et al. "Open Questions in Cosmic-Ray Research at Ultrahigh Energies." *Front. Astron. Space Sci.* 6 (2019), p. 23. DOI: 10.3389/fspas.2019.00023.

[77] Michael Unger. "Highlights from the Pierre Auger Observatory." *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Vol. 301. 2017, p. 1102. DOI: 10.22323/1.301.1102.

[78] V. Berezinsky, A. Z. Gazizov, and S. I. Grigorieva. "On astrophysical solution to ultrahigh-energy cosmic rays." *Phys. Rev. D* 74 (2006), p. 043005. DOI: 10.1103/PhysRevD.74.043005.

[79] Veniamin Berezinsky, A. Z. Gazizov, and S. I. Grigorieva. "Dip in UHECR spectrum as signature of proton interaction with CMB." *Phys. Lett. B* 612 (2005), pp. 147–153. DOI: 10.1016/j.physletb.2005.02.058.

[80] Alexander Aab et al. "Depth of Maximum of Air-Shower Profiles at the Pierre Auger Observatory: Measurements at Energies above $10^{17.8}$ eV." *Phys. Rev. D* 90.12 (2014), p. 122005. DOI: 10.1103/PhysRevD.90.122005.

[81] A. Aab et al. "Depth of maximum of air-shower profiles at the Pierre Auger Observatory. II. Composition implications." *Phys. Rev. D* 90.12 (2014), p. 122006. DOI: 10.1103/PhysRevD.90.122006.

[82] R. Aloisio, D. Boncioli, A di Matteo, A. F. Grillo, S. Petrera, and F. Salamida. "Cosmogenic neutrinos and ultra-high energy cosmic ray models." *JCAP* 10 (2015), p. 006. DOI: 10.1088/1475-7516/2015/10/006.

[83] Denis Allard, E. Parizot, and A. V. Olinto. "On the transition from galactic to extragalactic cosmic-rays: spectral and composition features from two opposite scenarios." *Astropart. Phys.* 27 (2007), pp. 61–75. DOI: 10.1016/j.astropartphys.2006.09.006.

[84] Denis Allard, A. V. Olinto, and E. Parizot. "Signatures of the extragalactic cosmic-ray source composition from spectrum and shower depth measurements." *Astron. Astrophys.* 473 (2007), pp. 59–66. DOI: 10.1051/0004-6361:20077478.

[85] C. De Donato and G. A. Medina-Tanco. "Experimental constraints on the astrophysical interpretation of the cosmic ray Galactic-extragalactic transition region." *Astropart. Phys.* 32 (2009), pp. 253–268. DOI: 10.1016/j.astropartphys.2009.09.004.

[86] M. Kachelrieß, O. Kalashev, S. Ostapchenko, and D. V. Semikoz. "Minimal model for extragalactic cosmic rays and neutrinos." *Phys. Rev. D* 96.8 (2017), p. 083006. DOI: 10.1103/PhysRevD.96.083006.

[87] Veniamin Berezinsky, A. Gazizov, and M. Kachelriess. "Second dip as a signature of ultrahigh energy proton interactions with cosmic microwave background radiation." *Phys. Rev. Lett.* 97 (2006), p. 231101. DOI: 10.1103/PhysRevLett.92.231101.

[88] M. Kachelriess and D. V. Semikoz. "Cosmic Ray Models." *Prog. Part. Nucl. Phys.* 109 (2019), p. 103710. DOI: 10.1016/j.ppnp.2019.07.002.

[89] Michael Friedlander. "A century of cosmic rays." *Nature* 483.7390 (Mar. 2012), pp. 400–401. DOI: 10.1038/483400a.

[90] Thomas K. Gaisser. "Spectrum of cosmic-ray nucleons, kaon production, and the atmospheric muon charge ratio." *Astropart. Phys.* 35 (2012), pp. 801–806. DOI: 10.1016/j.astropartphys.2012.02.010.

[91] Thomas K. Gaisser, Todor Stanev, and Serap Tilav. "Cosmic Ray Energy Spectrum from measurements of Air Showers." *Front. Phys.* 8 (2013), pp. 748–758. DOI: 10.1007/s11467-013-0319-7.

[92] Hans Dembinski, Ralph Engel, Anatoli Fedynitch, Thomas Gaisser, Felix Riehn, and Todor Stanev. "Data-driven model of the cosmic-ray flux and mass composition from 10 GeV to 10^{11} GeV." *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Vol. 301. 2017, p. 533. DOI: 10.22323/1.301.0533.

[93] Anthony M. Hillas. "Cosmic Rays: Recent Progress and some Current Questions" (July 2006). arXiv: astro-ph/0607109.

[94] R. U. Abbasi et al. "Indications of Proton-Dominated Cosmic Ray Composition above 1.6 EeV." *Phys. Rev. Lett.* 104 (2010), p. 161101. DOI: 10.1103/PhysRevLett.104.161101.

[95] H. S. Ahn et al. "Energy spectra of cosmic-ray nuclei at high energies." *Astrophys. J.* 707 (2009), pp. 593–603. DOI: 10.1088/0004-637X/707/1/593.

[96] Vladimir Lenok. "Measurement of the cosmic-ray energy spectrum using a novel approach to model the aperture of radio arrays." PhD thesis. Karlsruher Institut für Technologie (KIT), 2022. DOI: 10.5445/IR/1000143479.

[97] Johannes Albrecht et al. "The Muon Puzzle in cosmic-ray induced air showers and its connection to the Large Hadron Collider." *Astrophys. Space Sci.* 367.3 (2022), p. 27. DOI: 10.1007/s10509-022-04054-5.

[98] Peter K. F. Grieder. *Exentsive Air Showers and High Energy Phenomena*. Springer Berlin Heidelberg, 2010. DOI: 10.1007/978-3-540-76941-5.

[99] Ralph Engel, Dieter Heck, and Tanguy Pierog. "Extensive air showers and hadronic interactions at high energy." *Ann. Rev. Nucl. Part. Sci.* 61 (2011), pp. 467–489. DOI: 10.1146/annurev.nucl.012809.104544.

[100] T. Antoni et al. "A non-parametric approach to infer the energy spectrum and the mass composition of cosmic rays." *Astroparticle Physics* 16.3 (Jan. 2002), pp. 245–263. DOI: 10.1016/s0927-6505(01)00111-6.

[101] M. Amenomori et al. "Are protons still dominant at the knee of the cosmic-ray energy spectrum?" *Phys. Lett. B* 632 (2006), pp. 58–64. DOI: 10.1016/j.physletb.2005.10.048.

[102] S. Müller, R. Engel, T. Pierog, and M. Roth. "Impact of muon detection thresholds on the separability of primary cosmic rays." *Astropart. Phys.* 97 (2018), pp. 174–185. DOI: 10.1016/j.astropartphys.2017.11.005.

[103] Karl-Heinz Kampert and Michael Unger. "Measurements of the Cosmic Ray Composition with Air Shower Experiments." *Astropart. Phys.* 35 (2012), pp. 660–678. DOI: 10.1016/j.astropartphys.2012.02.004.

[104] JW Elbert. "Multiple muons produced by cosmic ray interactions." *DUMAND Summer Workshop, Volume 2-UHE Interactions, Neutrino Astronomy*. Vol. 2. 1979, pp. 101–121. URL: https://ui.adsabs.harvard.edu/abs/1979duma....2..101E.

[105] J. W. Elbert. "Cosmic ray multiple muon events in deep detectors." *Proceedings of 16th International Cosmic Ray Conference — PoS(ICRC1979)*. 1979, pp. 405–409. URL: https://ui.adsabs.harvard.edu/abs/1979ICRC...10..405E.

[106] T. K. Gaisser, Todor Stanev, Phyllis Freier, and C. Jake Waddington. "Nucleus-nucleus Collisions and Interpretation of Cosmic Ray Cascades above 100-TeV." *Phys. Rev. D* 25 (1982), pp. 2341–2350. DOI: 10.1103/PhysRevD.25.2341.

[107] T. K. Gaisser and T. Stanev. "Muon bundles in underground detectors." *Nucl. Instrum. Meth. A* 235 (1985), pp. 183–192. DOI: 10.1016/0168-9002(85)90260-8.

[108] C. Forti, H. Bilokon, B. d'Ettore Piazzoli, T. K. Gaisser, L. Satta, and T. Stanev. "Simulation of atmospheric cascades and deep underground muons." *Phys. Rev. D* 42 (1990), pp. 3668–3689. DOI: 10.1103/PhysRevD.42.3668.

[109] P. Lipari. "TeV muons in hadronic showers." *Astroparticle Physics* 1.4 (Dec. 1993), pp. 399–415. DOI: 10.1016/0927-6505(93)90006-y.

[110] Xinhua Bai, Dmitry Chirkin, Thomas Gaisser, Todor Stanev, and David Seckel. "Muon bundle energy loss in deep underground detector" (July 2009). URL: https://galprop.stanford.edu/elibrary/icrc/2009/preliminary/pdf/icrc0642.pdf.

[111] Thomas K. Gaisser and Stef Verpoest. "Profiles of energetic muons in the atmosphere." *Astropart. Phys.* 133 (2021), p. 102630. DOI: 10.1016/j.astropartphys.2021.102630.

[112] Dennis Soldin. "Laterally Separated Muons from Cosmic Ray Air Showers Measured with the IceCube Neutrino Observatory." PhD thesis. Wuppertal University, 2017. URL: https://elpub.bib.uni-wuppertal.de/edocs/dokumente/fbc/physik/diss2017/soldin/dc1740.pdf.

[113] Bruno Rossi and William B. Fretter. "High-Energy Particles." *American Journal of Physics* 21.3 (Mar. 1953), pp. 236–236. DOI: 10.1119/1.1933408.

[114] R. M. Sternheimer. "Density Effect for the Ionization Loss in Various Materials." *Physical Review* 103.3 (Aug. 1956), pp. 511–515. DOI: 10.1103/physrev.103.511.

[115] R. M. Sternheimer and R. F. Peierls. "General expression for the density effect for the ionization loss of charged particles." *Phys. Rev. B* 3 (1971), pp. 3681–3692. DOI: 10.1103/PhysRevB.3.3681.

[116] H. Bethe and W. Heitler. "On the Stopping of fast particles and on the creation of positive electrons." *Proc. Roy. Soc. Lond. A* 146 (1934), pp. 83–112. DOI: 10.1098/rspa.1934.0140.

[117] A. A. Petrukhin and V. V. Shestakov. "The influence of the nuclear and atomic form factors on the muon bremsstrahlung cross section." *Canadian Journal of Physics* 46.10 (May 1968), S377–S380. DOI: 10.1139/p68-251.

[118] R P Kokoulin and A A Petrukhin. "Influence of the nuclear form factor on the cross section of electron pair production by high energy muons." Jan. 1971. URL: https://www.osti.gov/biblio/4563918.

[119] K. Kobayakawa. "Fluctuations and nuclear interactions in the energy loss of cosmic-ray muons." *Il Nuovo Cimento B Series 10* 47.2 (Feb. 1967), pp. 156–184. DOI: 10.1007/bf02710720.

[120] L B Bezrukov and E V Bugaev. "Nucleon shadowing effects in photonuclear interactions." *Sov. J. Nucl. Phys. (Engl. Transl.)* (May 1981). URL: https://www.osti.gov/biblio/5909968.

[121] H. Bethe. "Theory of the Passage of Fast Corpuscular Rays Through Matter." *Annalen Phys.* 5 (1930), pp. 325–400. DOI: 10.1002/andp.19303970303.

[122] D. E. Groom and S. R. Klein. "Passage of particles through matter." *The European Physical Journal C* 15.1–4 (Mar. 2000), 163–173. DOI: 10.1007/bf02683419.

[123] A. Crispin and G. N. Fowler. "Density effect in the ionization energy loss of fast charged particles in matter." *Rev. Mod. Phys.* 42 (1970), pp. 290–316. DOI: 10.1103/RevModPhys.42.290.

[124] R. M. Sternheimer, M. J. Berger, and S. M. Seltzer. "Density Effect for the Ionization Loss of Charged Particles in Various Substances." *Atom. Data Nucl. Data Tabl.* 30 (1984), pp. 261–271. DOI: 10.1016/0092-640X(84)90002-0.

[125] A. Van Ginneken. "Energy Loss and Angular Characteristics of High-Energy Electromagnetic Processes." *Nucl. Instrum. Meth. A* 251 (1986), p. 21. DOI: 10.1016/0168-9002(86)91146-0.

[126] R P Kokoulin and A A Petrukhin. "Analysis of the cross-section of direct pair production by fast muons." *Acta Phys. Acad. Sci. Hung. 29: Suppl. 4, 277-84(1970).* (Jan. 1970). URL: https://www.osti.gov/biblio/4651250.

[127] P. A. Cerenkov. "Visible radiation produced by electrons moving in a medium with velocities exceeding that of light." *Phys. Rev.* 52 (1937), pp. 378–379. DOI: 10.1103/PhysRev.52.378.

[128] P.J. Nahin. *"Oliver Heaviside: The Life, Work, and Times of an Electrical Genius of the Victorian Age"*. Johns Hopkins University Press. URL: https://www.press.jhu.edu/books/title/2069/oliver-heaviside.

[129] Michael F L'annunziata. *"Radioactivity: introduction and history, from the Quantum to Quarks"*. URL: https://www.sciencedirect.com/book/9780444634894/radioactivity.

[130] J V Jelley. *"Cerenkov radiation and its applications"*. Vol. 6. 7. IOP Publishing, July 1955, pp. 227–232. DOI: 10.1088/0508-3443/6/7/301.

[131] Christiaan Huygens. *Traité de la lumière*. Chez Pierre vander Aa, marchand libraire, 1690. DOI: 10.5479/sil.294285.39088000545160.

[132] K. Greisen. "Cosmic ray showers." *Ann. Rev. Nucl. Part. Sci.* 10 (1960), pp. 63–108. DOI: 10.1146/annurev.ns.10.120160.000431.

[133] Frederick Reines. "Neutrino interactions." *Ann. Rev. Nucl. Part. Sci.* 10 (1960), pp. 1–26. DOI: 10.1146/annurev.ns.10.120160.000245.

[134] Moisey Alexandrovich Markov et al. "On high energy neutrino physics." 1960, pp. 578–581. URL: http://tinyurl.com/456aruts.

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

[136] A. Roberts. "The Birth of high-energy neutrino astronomy: A Personal history of the DUMAND project." *Rev. Mod. Phys.* 64 (1992), pp. 259–312. DOI: 10.1103/RevModPhys.64.259.

[137] Thomas K. Gaisser, Francis Halzen, and Todor Stanev. "Particle astrophysics with high energy neutrinos." *Physics Reports* 258.3 (July 1995), pp. 173–236. DOI: 10.1016/0370-1573(95)00003-y.

[138] John G. Learned and Karl Mannheim. "High-Energy Neutrino Astrophysics." *Annual Review of Nuclear and Particle Science* 50.1 (2000), pp. 679–749. DOI: 10.1146/annurev.nucl.50.1.679.

[139] Francis Halzen and Dan Hooper. "High-energy neutrino astronomy: The Cosmic ray connection." *Rept. Prog. Phys.* 65 (2002), pp. 1025–1078. DOI: 10.1088/0034-4885/65/7/201.

[140] J. Babson et al. "Cosmic Ray Muons in the Deep Ocean." *Phys. Rev. D* 42 (1990), pp. 3613–3620. DOI: 10.1103/PhysRevD.42.3613.

[141] V Balkanov et al. "The BAIKAL neutrino project: status report." *Nuclear Physics B - Proceedings Supplements* 91.1-3 (Jan. 2001), pp. 438–444. DOI: 10.1016/s0920-5632(00)00973-7.

[142] G. Aggouras et al. "A measurement of the cosmic-ray muon flux with a module of the NESTOR neutrino telescope." *Astropart. Phys.* 23 (2005), pp. 377–392. DOI: 10.1016/j.astropartphys.2005.02.001.

[143] J. A. Aguilar et al. "First results of the Instrumentation Line for the deep-sea ANTARES neutrino telescope." *Astropart. Phys.* 26 (2006), pp. 314–324. DOI: 10.1016/j.astropartphys.2006.07.004.

[144] R. M. Bionta et al. "Observation of a Neutrino Burst in Coincidence with Supernova SN 1987a in the Large Magellanic Cloud." *Phys. Rev. Lett.* 58 (1987), p. 1494. DOI: 10.1103/PhysRevLett.58.1494.

[145] K. S. Hirata et al. "Observation in the Kamiokande-II Detector of the Neutrino Burst from Supernova SN 1987a." *Phys. Rev. D* 38 (1988), pp. 448–458. DOI: 10.1103/PhysRevD.38.448.

[146] M. R. Krishnaswamy et al. "The Kolar gold fields neutrino experiment. 1. the interactions of cosmic ray neutrinos." *Proc. Roy. Soc. Lond. A* 323 (1971), pp. 489–509. DOI: 10.1098/rspa.1971.0119.

[147] F. Reines et al. "Evidence for high-energy cosmic ray neutrino interactions." *Phys. Rev. Lett.* 15 (1965), pp. 429–433. DOI: 10.1103/PhysRevLett.15.429.

[148] F. Ronga. ""Atmospheric Neutrino Results from the MACRO Experiment"." *The Ninth Marcel Grossmann Meeting*. Dec. 2002. DOI: 10.1142/9789812777386_0596.

[149] W. Rhode et al. "Limits on the flux of very high-energetic neutrinos with the Frejus detector." *Astropart. Phys.* 4 (1996), pp. 217–225. DOI: 10.1016/0927-6505(95)00038-0.

[150] Francis Halzen. "Neutrino astronomy: An update." *Astropart. Phys.* 53 (2014), pp. 166–174. DOI: 10.1016/j.astropartphys.2013.01.011.

[151] Francis Halzen and Spencer R. Klein. "IceCube: An Instrument for Neutrino Astronomy." *Rev. Sci. Instrum.* 81 (2010), p. 081101. DOI: 10.1063/1.3480478.

[152] M. Ahlers, L. A. Anchordoqui, M. C. Gonzalez-Garcia, F. Halzen, and S. Sarkar. "GZK Neutrinos after the Fermi-LAT Diffuse Photon Flux Measurement." *Astropart. Phys.* 34 (2010), pp. 106–115. DOI: 10.1016/j.astropartphys.2010.06.003.

[153] J Wdowczyk, W Tkaczyk, and A W Wolfendale. "Primary cosmic γ-rays above 10^{12} eV." *Journal of Physics A: General Physics* 5.9 (Sept. 1972), pp. 1419–1432. DOI: 10.1088/0305-4470/5/9/015.

[154] F. W. Stecker. "Ultrahigh energy photons, electrons and neutrinos, the microwave background, and the universal cosmic ray hypothesis." *Astrophys. Space Sci.* 20 (1973), pp. 47–57. DOI: 10.1007/BF00645585.

[155] Luis A. Anchordoqui, Haim Goldberg, Francis Halzen, and Thomas J. Weiler. "Galactic point sources of TeV antineutrinos." *Phys. Lett. B* 593 (2004), p. 42. DOI: 10.1016/j.physletb.2004.04.054.

[156] F. A. Aharonian et al. "High-energy particle acceleration in the shell of a supernova remnant." *Nature* 432 (2004), pp. 75–77. DOI: 10.1038/nature02960.

[157] J. Alvarez-Muñiz and F. Halzen. "Possible High-Energy Neutrinos from the Cosmic Accelerator RX J1713.7-3946." *The Astrophysical Journal* 576.1 (Sept. 2002), pp. L33–L36. DOI: 10.1086/342978.

[158] Mathieu Ribordy et al. "From AMANDA to IceCube." *Phys. Atom. Nucl.* 69 (2006), pp. 1899–1907. DOI: 10.1134/S1063778806110147.

[159] Thomas Gaisser and Francis Halzen. "IceCube." *Ann. Rev. Nucl. Part. Sci.* 64 (2014), pp. 101–123. DOI: 10.1146/annurev-nucl-102313-025321.

[160] Henry David Thoreau. *The Illustrated "A Week on the Concord and Merrimack Rivers"*. Princeton University Press, Dec. 1986. DOI: 10.1515/9781400857197.

[161] E. Andres et al. "AMANDA: Status, results and future." *8th International Workshop on Neutrino Telescopes*. Feb. 1999, pp. 63–79. arXiv: astro-ph/9906205.

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

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

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

[165] J Kiryluk, MV D'Agostino, SR Klein, C Song, and DR Williams. "IceCube performance with artificial light sources: The road to Cascade Analyses." *Proceedings of 30th International Cosmic Ray Conference — (ICRC2007)*, pp. 1233–1236. URL: https://bib-pubdb1.desy.de/record/83220.

[166] M. G. Aartsen et al. "Cosmic ray spectrum from 250 TeV to 10 PeV using IceTop." *Phys. Rev. D* 102 (2020), p. 122001. DOI: 10.1103/PhysRevD.102.122001.

[167] Ramesh Koirala. "Extension of the IceTop energy spectrum to 250 TeV and application of the constant intensity cut method to IceTop data." PhD thesis. University of Delaware, 2019. URL: https://udspace.udel.edu/handle/19716/24883.

[168] Abd al Karim Haj Ismail. "Low Energy Air Showers with IceTop." *Proceedings of 33rd International Cosmic Ray Conference — PoS(ICRC2013)*, p. 0674. URL: https://inspirehep.net/files/2305778183f925e169da763e6c89a57a.

[169] Katherine Rawlins. "A Function to Describe Attenuation of Cosmic Ray Air Shower Particles in Snow." *Proceedings of The 34th International Cosmic Ray Conference — PoS(ICRC2015)*. Vol. 236. 2016, p. 628. DOI: 10.22323/1.236.0628.

[170] Donald Shepard. "A two-dimensional interpolation function for irregularly-spaced data." ACM Press, 1968. DOI: 10.1145/800186.810616.

[171] A. Achterberg et al. "First Year Performance of The IceCube Neutrino Telescope." *Astropart. Phys.* 26 (2006), pp. 155–173. DOI: 10.1016/j.astropartphys.2006.06.007.

[172] Hermann Kolanoski. "IceCube–Astrophysics and Astroparticle Physics at the South Pole." *Proceedings of 32nd International Cosmic Ray Conference — PoS(ICRC2011)*. Vol. 11, p. 79. URL: https://arxiv.org/abs/1111.5188.

[173] Thomas K. Gaisser. "IceCube: Status and Results" (Aug. 2011). arXiv: 1108.1838 [astro-ph.HE].

[174] Stef Verpoest et al. "Testing Hadronic Interaction Models with Cosmic Ray Measurements at the IceCube Neutrino Observatory." *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 357. DOI: 10.22323/1.395.0357.

[175] Rasha Abbasi et al. "Observation of Cosmic Ray Anisotropy with Nine Years of IceCube Data." *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 320. DOI: 10.22323/1.395.0320.

[176] R. Abbasi et al. "Measurement of the Anisotropy of Cosmic Ray Arrival Directions with IceCube." *Astrophys. J. Lett.* 718 (2010), p. L194. DOI: 10.1088/2041-8205/718/2/L194.

[177] R. Abbasi et al. "Observation of Anisotropy in the Arrival Directions of Galactic Cosmic Rays at Multiple Angular Scales with IceCube." *Astrophys. J.* 740 (2011), p. 16. DOI: 10.1088/0004-637X/740/1/16.

[178] M. G. Aartsen et al. "Observation of Cosmic Ray Anisotropy with the IceTop Air Shower Array." *Astrophys. J.* 765 (2013), p. 55. DOI: 10.1088/0004-637X/765/1/55.

[179] R. Abbasi et al. "Observation of an Anisotropy in the Galactic Cosmic Ray arrival direction at 400 TeV with IceCube." *Astrophys. J.* 746 (2012), p. 33. DOI: 10.1088/0004-637X/746/1/33.

[180] M. G. Aartsen et al. "Searches for small-scale anisotropies from neutrino point sources with three years of IceCube data." *Astropart. Phys.* 66 (2015), pp. 39–52. DOI: 10.1016/j.astropartphys.2015.01.001.

[181] R. Abbasi et al. "Density of GeV muons in air showers measured with IceTop." *Phys. Rev. D* 106.3 (2022), p. 032010. DOI: 10.1103/PhysRevD.106.032010.

[182] M. G. Aartsen et al. "Search for Sources of High-Energy Neutrons with four Years of Data from the IceTop Detector." *Astrophys. J.* 830.2 (2016), p. 129. DOI: 10.3847/0004-637X/830/2/129.

[183] Hershal Pandya. "Search for PeV gamma rays and astrophysical neutrinos with IceTop and IceCube." PhD thesis. University of Delaware, 2019. URL: https://udspace.udel.edu/items/d94a07de-5a29-4e71-bd4e-f9606d4694ed.

[184] R. Abbasi et al. "Solar Energetic Particle Spectrum on 13 December 2006 Determined by IceTop." *Astrophys. J. Lett.* 689 (2008), pp. L65–L68. DOI: 10.1086/595679.

[185] Takao Kuwabara, J. W. Bieber, and R. Pyle. "Heliospheric physics with IceTop." *Proceedings of 30th International Cosmic Ray Conference — (ICRC2007)*. URL: https://articles.adsabs.harvard.edu/pdf/2008ICRC...30a.339K.

[186] R. Abbasi et al. "Observation of seasonal variations of the flux of high-energy atmospheric neutrinos with IceCube." *Eur. Phys. J. C* 83.9 (2023), p. 777. DOI: 10.1140/epjc/s10052-023-11679-5.

[187] Delia Tosi and Hershal Pandya. "IceTop as veto for IceCube: results." *PoS* ICRC2019 (2020), p. 445. DOI: 10.22323/1.358.0445.

[188] Delia Tosi and Hershal Pandya. "Performance of IceTop as a veto for IceCube." *PoS* ICRC2017 (2018), p. 967. DOI: 10.22323/1.301.0967.

[189] Delia Tosi and Hershal Pandya. "IceTop as veto for IceCube: results." *Proceedings of 36th International Cosmic Ray Conference — PoS(ICRC2019)*. Vol. 358. 2019, p. 445. DOI: 10.22323/1.358.0445.

[190] Najia Moureen Binte Amin. "Implementation of IceTop data in the IceCube Realtime Alert System." *J. Phys. Conf. Ser.* 2156.1 (2021), p. 012217. DOI: 10.1088/1742-6596/2156/1/012217.

[191] Rasha Abbasi et al. "Simulation Study of the Observed Radio Emission of Air Showers by the IceTop Surface Extension." *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 317. DOI: 10.22323/1.395.0317.

[192] P. B. Price, K. Woschnagg, and D. Chirkin. "Age vs depth of glacial ice at South Pole." *Geophysical Research Letters* 27.14 (July 2000), pp. 2129–2132. DOI: 10.1029/2000gl011351.

[193] R. C. Bay, R. A. Rohde, P. B. Price, and N. E. Bramall. "South Pole paleowind from automated synthesis of ice core records." *Journal of Geophysical Research* 115.D14 (July 2010). DOI: 10.1029/2009jd013741.

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

[195] "In situ estimation of ice crystal properties at the South Pole using LED calibration data from the IceCube Neutrino Observatory." *The Cryosphere* 18.1 (2024), pp. 75–102. DOI: 10.5194/tc-18-75-2024.

[196] Sebastian Euler. "Observation of oscillations of atmospheric neutrinos with the IceCube Neutrino Observatory." PhD thesis. RWTH Aachen, 2014. URL: https://publications.rwth-aachen.de/record/444944.

[197] R. Abbasi et al. "A muon-track reconstruction exploiting stochastic losses for large-scale Cherenkov detectors." *JINST* 16.08 (2021), P08034. DOI: 10.1088/1748-0221/16/08/P08034.

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

[199] John G Learned and Sandip Pakvasa. "Detecting ν_τ oscillations at PeV energies." *Astroparticle Physics* 3.3 (May 1995), pp. 267–274. DOI: 10.1016/0927-6505(94)00043-3.

[200] M. G. Aartsen et al. "Search for non-relativistic Magnetic Monopoles with IceCube." *Eur. Phys. J. C* 74.7 (2014), p. 2938. DOI: 10.1140/epjc/s10052-014-2938-8.

[201] Sam De Ridder. "Sensitivity of IceCube cosmic ray measurements to the hadronic interaction models." PhD thesis. Ghent University, 2019. URL: https://biblio.ugent.be/publication/8609789.

[202] R. Abbasi et al. "IceTop: The surface component of IceCube." *Nucl. Instrum. Meth. A* 700 (2013), pp. 188–220. DOI: 10.1016/j.nima.2012.10.067.

[203] Michael Beimforde. "Calibration of air shower signals in the IceTop detector using cosmic ray muons." Diploma thesis. Humboldt-Universität zu Berlin, Dec. 2006.

[204] Serap Tilav, Paolo Desiati, Takao Kuwabara, Dominick Rocco, Florian Rothmaier, Matt Simmons, and Henrike Wissing. "Atmospheric Variations as Observed by IceCube" (Jan. 2010). arXiv: 1001.0776 [astro-ph.HE].

[205] M. G. Aartsen et al. "IceCube-Gen2: the window to the extreme Universe." *J. Phys. G* 48.6 (2021), p. 060501. DOI: 10.1088/1361-6471/abbd48.

[206] National Academies of Sciences Engineering, Medicine, et al. *"Decadal survey on astronomy and astrophysics 2020"*. 2021. URL: https://www.nationalacademies.org/our-work/decadal-survey-on-astronomy-and-astrophysics-2020-astro2020.

[207] Markus Ackermann et al. "Astrophysics Uniquely Enabled by Observations of High-Energy Cosmic Neutrinos." *Bull. Am. Astron. Soc.* 51 (2019), p. 185. arXiv: 1903.04334 [astro-ph.HE].

[208] Markus Ackermann et al. "Fundamental Physics with High-Energy Cosmic Neutrinos." *Bull. Am. Astron. Soc.* 51 (2019), p. 215. eprint: 1903.04333.

[209] Frank G Schröder et al. "High-energy galactic cosmic rays (Astro2020 Science White Paper)" (2019). URL: https://arxiv.org/abs/1903.07713.

[210] Nissim Fraija et al. "Cosmic rays in the TeV to PeV primary energy range." *Astro2020: Decadal Survey on Astronomy and Astrophysics* 2020 (2019), p. 459. URL: https://www.snowmass21.org/docs/files/.

[211] Fred Sarazin et al. "What is the nature and origin of the highest-energy particles in the universe?" *Bull. Am. Astron. Soc* 51.3 (2019), p. 93. URL: https://arxiv.org/abs/1903.04063.

[212] Bindu Rani et al. "Multi-physics of AGN jets in the multi-messenger era" (2019). URL: https://arxiv.org/abs/1903.04504.

[213] Marcos Santander et al. "A unique messenger to probe active galactic nuclei: high-energy neutrinos." *Bulletin of the American Astronomical Society* 51.3 (2019), p. 228. URL: https://baas.aas.org/pub/2020n3i228/release/1.

[214] Tonia M Venters et al. "Energetic particles of cosmic accelerators I: galactic accelerators" (2019). URL: https://arxiv.org/abs/1903.04634.

[215] Pierre Cristofari et al. "Where are the PeVatrons?" *Bulletin of the American Astronomical Society* 51.3 (2019), p. 115. URL: https://ui.adsabs.harvard.edu/abs/2019BAAS...51c.115C/abstract.

[216] MU Nisa et al. "The Sun at GeV–TeV Energies: A New Laboratory for Astroparticle Physics" (2019). URL: https://arxiv.org/abs/1903.06349.

[217] Justin Vandenbroucke et al. "Multi-messenger and transient astrophysics with very-high-energy gamma rays." *Bull. Am. Astron. Soc* 51 (2019), p. 553. URL: https://ui.adsabs.harvard.edu/abs/2019BAAS...51c.553V/abstract.

[218] Jamie Holder et al. "Understanding the origin and impact of relativistic cosmic particles with very-high-energy gamma-rays." *Bulletin of the American Astronomical Society* (2019). URL: https://baas.aas.org/pub/2020n3i267/release/1.

[219] R. Ojha et al. "Neutrinos, Cosmic Rays and the MeV Band" (Mar. 2019). eprint: 1903.05765.

[220] Stefano Ciprini, Sara Buson, Chi C Cheung, and Carlotta Pittori. "Gravitationally lensed MeV gamma-ray blazars." *Astro2020: Decadal Survey on Astronomy and Astrophysics* 2020 (2019), p. 427. URL: https://ui.adsabs.harvard.edu/abs/2019astro2020T.427C/abstract.

[221] Eric Burns et al. "A summary of multimessenger science with neutron star mergers" (2019). URL: https://arxiv.org/abs/1903.03582.

[222] Vassiliki Kalogera et al. "The Yet-Unobserved Multi-Messenger Gravitational-Wave Universe" (2019). URL: https://arxiv.org/abs/1903.09224.

[223] KE Ford et al. "AGN (and other) astrophysics with Gravitational Wave Events" (2019). URL: https://arxiv.org/abs/1903.09529.

[224] Andreas Haungs. "A Scintillator and Radio Enhancement of the IceCube Surface Detector Array." *EPJ Web Conf.* 210 (2019), p. 06009. DOI: 10.1051/epjconf/201921006009.

[225] S. Shefali. "R&D and production of the scintillation detectors for the IceCube Surface Array Enhancement." *Proceedings of 27th European Cosmic Ray Symposium — PoS(ECRS)*. Vol. 423. 2023, p. 141. DOI: 10.22323/1.423.0141.

[226] "Development of a scintillation and radio hybrid detector array at the South Pole." *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 225. DOI: 10.22323/1.395.0225.

[227] Lew Classen. "The mDOM – a multi-PMT digital optical module for the IceCube-Gen2 neutrino telescope." PhD thesis. Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2017. URL: https://opus4.kobv.de/opus4-fau/frontdoor/index/index/docId/8678.

[228] Lew Classen and Alexander Kappes. "The multi-PMT optical module for the IceCube-Upgrade." *EPJ Web Conf.* 207 (2019), p. 06004. DOI: 10.1051/epjconf/201920706004.

[229] Aya Ishihara, Achim Stoessl, Satoshi Shimizu, and Shigeru Yoshida. "Overview and performance of the D-Egg optical sensor for IceCube-Gen2." *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Vol. 301. 2017, p. 1051. DOI: 10.22323/1.301.1051.

[230] Aya Ishihara. "The IceCube Upgrade - Design and Science Goals." *Proceedings of 36th International Cosmic Ray Conference — PoS(ICRC2019)*. Vol. 358. 2019, p. 1031. DOI: 10.22323/1.358.1031.

[231] Agnieszka Leszczyńska. "Potential of the IceTop Enhancement with a Scintillation Detector Array." PhD thesis. Karlsruher Institut für Technologie (KIT), 2021. URL: https://publikationen.bibliothek.kit.edu/1000131245.

[232] Matt Kauer, Thomas Huber, Delia Tosi, and Chris Wendt. "The Scintillator Upgrade of IceTop: Performance of the Prototype Array." *Proceedings of 36th International Cosmic Ray Conference — PoS(ICRC2019)*. Vol. 358. 2019, p. 309. DOI: 10.22323/1.358.0309.

[233] Alexander Aab et al. "Deep-learning based reconstruction of the shower maximum X_{max} using the water-Cherenkov detectors of the Pierre Auger Observatory." *JINST* 16.07 (2021), P07019. DOI: 10.1088/1748-0221/16/07/P07019.

[234] P. A. Bezyazeekov et al. "Reconstruction of cosmic ray air showers with Tunka-Rex data using template fitting of radio pulses." *Phys. Rev. D* 97.12 (2018), p. 122004. DOI: 10.1103/PhysRevD.97.122004.

[235] S. Buitink et al. "A large light-mass component of cosmic rays at 10^{17} - $10^{17.5}$ eV from radio observations." *Nature* 531 (2016), p. 70. DOI: 10.1038/nature16976.

[236] D. Kostunin et al. "Towards a cosmic-ray mass-composition study at Tunka Radio Extension." *EPJ Web Conf.* 135 (2017), p. 01005. DOI: 10.1051/epjconf/201713501005.

[237] William Hanlon. "Telescope Array 10 Year Composition." *Proceedings of 36th International Cosmic Ray Conference — PoS(ICRC2019)*. Vol. 358. 2019, p. 280. DOI: 10.22323/1.358.0280.

[238] N. Palmieri et al. "Reconstructing energy and X_{max} of cosmic ray air showers using the radio lateral distribution measured with LOPES." *AIP Conf. Proc.* 1535.1 (2013), pp. 89–93. DOI: 10.1063/1.4807527.

[239] Frank G. Schroeder et al. "The Surface Array planned for IceCube-Gen2." *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 407. DOI: 10.22323/1.395.0407.

[240] Aswathi V. Balagopal. "A quest for PeVatrons employing radio detection of extensive air Showers." PhD thesis. Karlsruher Institut für Technologie (KIT), 2019. URL: https://publikationen.bibliothek.kit.edu/1000091377.

[241] Frank G. Schröder. “Physics Potential of a Radio Surface Array at the South Pole.” *EPJ Web Conf.* 216 (2019), p. 01007. DOI: 10.1051/epjconf/201921601007.

[242] Rasha Abbasi et al. “First air-shower measurements with the prototype station of the IceCube surface enhancement.” *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 314. DOI: 10.22323/1.395.0314.

[243] Mark Weyrauch. “Simulation study for the IceCube-Gen2 surface array.” Master’s thesis. Karlsruher Institut für Technologie (KIT), 2022. DOI: 10.5445/IR/1000152578.

[244] Rasha Abbasi et al. “Sensitivity studies for the IceCube-Gen2 radio array.” *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 1183. DOI: 10.22323/1.395.1183.

[245] M. G. Aartsen et al. “First observation of PeV-energy neutrinos with IceCube.” *Phys. Rev. Lett.* 111 (2013), p. 021103. DOI: 10.1103/PhysRevLett.111.021103.

[246] “Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector.” *Science* 342.6161 (Nov. 2013). DOI: 10.1126/science.1242856.

[247] M. G. Aartsen et al. “Evidence for Astrophysical Muon Neutrinos from the Northern Sky with IceCube.” *Phys. Rev. Lett.* 115.8 (2015), p. 081102. DOI: 10.1103/PhysRevLett.115.081102.

[248] Mark G Aartsen et al. “Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data.” *Physical Review Letters* 113.10 (Sept. 2014). DOI: 10.1103/physrevlett.113.101101.

[249] M. G. Aartsen et al. “Detection of a particle shower at the Glashow resonance with IceCube.” *Nature* 591.7849 (2021), pp. 220–224. DOI: 10.1038/s41586-021-03256-1.

[250] M. G. Aartsen et al. “Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert.” *Science* 361.6398 (2018), pp. 147–151. DOI: 10.1126/science.aat2890.

[251] M. G. Aartsen et al. “Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A.” *Science* 361.6398 (2018). DOI: 10.1126/science.aat1378.

[252] R. Abbasi et al. “Evidence for neutrino emission from the nearby active galaxy NGC 1068.” *Science* 378.6619 (2022), pp. 538–543. DOI: 10.1126/science.abg3395.

[253] R. Abbasi et al. “Constraining High-energy Neutrino Emission from Supernovae with IceCube.” *Astrophys. J. Lett.* 949.1 (2023), p. L12. DOI: 10.3847/2041-8213/acd2c9.

[254] R. Abbasi et al. “First all-flavor search for transient neutrino emission using 3-years of IceCube DeepCore data.” *JCAP* 01 (2022), p. 027. DOI: 10.1088/1475-7516/2022/01/027.

[255] R. Abbasi et al. “A Search for Coincident Neutrino Emission from Fast Radio Bursts with Seven Years of IceCube Cascade Events.” *Astrophys. J.* 946.2 (2023), p. 80. DOI: 10.3847/1538-4357/acbea0.

[256] R. Abbasi et al. "Searching for High-energy Neutrino Emission from Galaxy Clusters with IceCube." *Astrophys. J. Lett.* 938.2 (2022), p. L11. DOI: 10.3847/2041-8213/ac966b.

[257] R. Abbasi et al. "A Search for Time-dependent Astrophysical Neutrino Emission with IceCube Data from 2012 to 2017." *Astrophys. J.* 911.1 (2021), p. 67. DOI: 10.3847/1538-4357/abe7e6.

[258] Axel Mellinger. "A Color All-Sky Panorama Image of the Milky Way." *Publ. Astron. Soc. Pac.* 121 (2009), pp. 1180–1187. DOI: 10.1086/648480.

[259] M. Ajello et al. "The Fourth Catalog of Active Galactic Nuclei Detected by the Fermi Large Area Telescope: Data Release 3." *Astrophys. J. Suppl.* 263.2 (2022), p. 24. DOI: 10.3847/1538-4365/ac9523.

[260] R. Abbasi et al. "Measurement of the high-energy all-flavor neutrino-nucleon cross section with IceCube." *Phys. Rev. D* 104 (2 2021), p. 022001. DOI: 10.1103/PhysRevD.104.022001.

[261] R. Abbasi et al. "Search for Unstable Sterile Neutrinos with the IceCube Neutrino Observatory." *Phys. Rev. Lett.* 129.15 (2022), p. 151801. DOI: 10.1103/PhysRevLett.129.151801.

[262] A. Albert et al. "Search for Spatial Correlations of Neutrinos with Ultra-high-energy Cosmic Rays." *Astrophys. J.* 934.2 (2022), p. 164. DOI: 10.3847/1538-4357/ac6def.

[263] R. Abbasi et al. "The Design and Performance of IceCube DeepCore." *Astropart. Phys.* 35 (2012), pp. 615–624. DOI: 10.1016/j.astropartphys.2012.01.004.

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

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

[266] James Vincent Mead. "Atmospheric neutrino oscillations in IceCube-DeepCore." *PoS* NOW2022 (2023), p. 007. DOI: 10.22323/1.421.0007.

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

[268] Martin Leuermann. "Testing the Neutrino Mass Ordering with IceCube DeepCore." PhD thesis. RWTH Aachen, 2018. DOI: 10.18154/RWTH-2018-231554.

[269] Gianfranco Bertone, Dan Hooper, and Joseph Silk. "Particle dark matter: Evidence, candidates and constraints." *Phys. Rept.* 405 (2005), pp. 279–390. DOI: 10.1016/j.physrep.2004.08.031.

[270] M. G. Aartsen et al. "Search for dark matter annihilations in the Sun with the 79-string IceCube detector." *Phys. Rev. Lett.* 110.13 (2013), p. 131302. DOI: 10.1103/PhysRevLett.110.131302.

[271] R. Abbasi et al. "Limits on a muon flux from neutralino annihilations in the Sun with the IceCube 22-string detector." *Phys. Rev. Lett.* 102 (2009), p. 201302. DOI: 10.1103/PhysRevLett.102.201302.

[272] R. Abbasi et al. "Search for neutrino lines from dark matter annihilation and decay with IceCube." *Phys. Rev. D* 108 (10 2023), p. 102004. DOI: 10.1103/PhysRevD.108.102004.

[273] R. Abbasi et al. "Search for GeV-scale dark matter annihilation in the Sun with IceCube DeepCore." *Phys. Rev. D* 105.6 (2022), p. 062004. DOI: 10.1103/PhysRevD.105.062004.

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

[275] Ivone F. M. Albuquerque and Spencer R. Klein. "Supersymmetric and Kaluza-Klein Particles Multiple Scattering in the Earth." *Phys. Rev. D* 80 (2009), p. 015015. DOI: 10.1103/PhysRevD.80.015015.

[276] John N. Bahcall and Peter Meszaros. "5 to 10 GeV neutrinos from gamma-ray burst fireballs." *Phys. Rev. Lett.* 85 (2000), pp. 1362–1365. DOI: 10.1103/PhysRevLett.85.1362.

[277] R. Abbasi et al. "Search for sub-TeV Neutrino Emission from Novae with IceCube-DeepCore." *Astrophys. J.* 953.2 (2023), p. 160. DOI: 10.3847/1538-4357/acdc1b.

[278] M. G. Aartsen et al. "Anisotropy in cosmic-ray arrival directions in the Southern Hemisphere based on six years of data from the IceCube detector." *Astrophys. J.* 826.2 (2016), p. 220. DOI: 10.3847/0004-637X/826/2/220.

[279] P. Abreu et al. "Search for First Harmonic Modulation in the Right Ascension Distribution of Cosmic Rays Detected at the Pierre Auger Observatory." *Astropart. Phys.* 34 (2011), pp. 627–639. DOI: 10.1016/j.astropartphys.2010.12.007.

[280] Andrea Chiavassa. "A study of the first harmonic of the large scale anisotropies with the KASCADE-Grande experiment." *Proceedings of The 34th International Cosmic Ray Conference — PoS(ICRC2015)*. Vol. 236. 2016, p. 281. DOI: 10.22323/1.236.0281.

[281] A. U. Abeysekara et al. "All-Sky Measurement of the Anisotropy of Cosmic Rays at 10 TeV and Mapping of the Local Interstellar Magnetic Field." *Astrophys. J.* 871.1 (2019), p. 96. DOI: 10.3847/1538-4357/aaf5cc.

[282] Paras Koundal et al. "Cosmic-Ray Composition analysis at IceCube using Graph Neural Networks." *Proceedings of 27th European Cosmic Ray Symposium — PoS(ECRS)*. Vol. 423. 2023, p. 085. DOI: 10.22323/1.423.0085.

[283] Paras Koundal and for IceCube Collaboration. "Composition Analysis of cosmic-rays at IceCube Observatory, using Graph Neural Networks." *Workshop on Machine Learning for Cosmic-Ray Air Showers (2022)*. 2022. eprint: https://zenodo.org/record/6439650.

[284] Paras Koundal, Matthias Plum, and Julian Saffer. "Study of mass composition of cosmic rays with IceTop and IceCube." *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 323. DOI: 10.22323/1.395.0323.

[285] Paras Koundal. “Graph Neural Networks and Application for Cosmic-Ray Analysis.” *Proceedings of The 5th International Workshop on Deep Learning in Computational Physics — PoS(DLCP2021)*. Vol. 410. 2021, p. 004. DOI: 10.22323/1.410.0004.

[286] Julian Saffer and For The IceCube Collaboration. “Composition of 100 TeV - 100 PeV Cosmic Rays with IceCube and IceTop using Boosted Decision Trees” (2022). DOI: 10.5281/ZENODO.6525146.

[287] Dennis Soldin. “Update on the Combined Analysis of Muon Measurements from Nine Air Shower Experiments.” *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 349. DOI: 10.22323/1.395.0349.

[288] M. G. Aartsen et al. “Search for PeV Gamma-Ray Emission from the Southern Hemisphere with 5 Years of Data from the IceCube Observatory.” *Astrophys. J.* 891 (Aug. 2019), p. 9. DOI: 10.3847/1538-4357/ab6d67.

[289] Federico Bontempo. “New approaches for gamma-hadron separation at the IceCube Neutrino Observatory.” *EPJ Web Conf.* 280 (2023), p. 04004. DOI: 10.1051/epjconf/202328004004.

[290] George W. Clark. “Arrival Directions of Cosmic-Ray Air Showers from the Northern Sky.” *Phys. Rev.* 108 (1957), pp. 450–457. DOI: 10.1103/PhysRev.108.450.

[291] M. G. Aartsen et al. “Observation of the cosmic-ray shadow of the Moon with IceCube.” *Phys. Rev. D* 89.10 (2014), p. 102004. DOI: 10.1103/PhysRevD.89.102004.

[292] M. Amenomori et al. “Cosmic ray deficit from the directions of the moon and the sun detected with the Tibet Air Shower Array.” *Phys. Rev. D* 47 (1993), pp. 2675–2681. DOI: 10.1103/PhysRevD.47.2675.

[293] B. Bartoli et al. “Observation of the Cosmic Ray Moon shadowing effect with ARGO-YBJ.” *Phys. Rev. D* 84 (2011), p. 022003. DOI: 10.1103/PhysRevD.84.022003.

[294] FW Samuelson, MILAGRO Collaboration, et al. “The cosmic ray Moon shadow seen by MILAGRO.” *Proceedings of 27th International Cosmic Ray Conference — PoS(ICRC2001)*. Vol. 2. 2001, p. 594. URL: https://ui.adsabs.harvard.edu/abs/2001ICRC....2..594S/abstract.

[295] M. G. Aartsen et al. “Measurements of the time-dependent cosmic-ray Sun shadow with seven years of IceCube data: Comparison with the Solar cycle and magnetic field models.” *Phys. Rev. D* 103.4 (2021), p. 042005. DOI: 10.1103/PhysRevD.103.042005.

[296] Donghwa Kang et al. “Studies of a muon-based mass sensitive parameter for the IceTop surface array.” *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 312. DOI: 10.22323/1.395.0312.

[297] Rasha Abbasi et al. “Discrimination of Muons for Mass Composition Studies of Inclined Air Showers Detected with IceTop.” *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 212. DOI: 10.22323/1.395.0212.

[298] Rasha Abbasi et al. "Hybrid cosmic ray measurements using the IceAct telescopes in coincidence with the IceCube and IceTop detectors." *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Vol. 395. 2021, p. 276. DOI: 10.22323/1.395.0276.

[299] Emily Dvorak, Xinhua Bai, Javier Gonzalez, and Dennis Soldin. "A Three-dimensional Reconstruction of Cosmic Ray Events in IceCube." *Proceedings of 36th International Cosmic Ray Conference — PoS(ICRC2019)*. Vol. 358. 2019, p. 244. DOI: 10.22323/1.358.0244.

[300] Wikipedia contributors. *ChatGPT — Wikipedia, The Free Encyclopedia*. [Online; accessed 10-April-2023]. 2023. URL: https://en.wikipedia.org/w/index.php?title=ChatGPT&oldid=1149034973.

[301] Robert Stein et al. "A tidal disruption event coincident with a high-energy neutrino." *Nature Astron.* 5.5 (2021), pp. 510–518. DOI: 10.1038/s41550-020-01295-8.

[302] Ann Finkbeiner. "Messengers from the Sky." *Scientific American* 318.5 (Apr. 2018), pp. 36–41. DOI: 10.1038/scientificamerican0518-36.

[303] R. Abbasi et al. "IceCube Search for Neutrinos Coincident with Gravitational Wave Events from LIGO/Virgo Run O3." *Astrophys. J.* 944.1 (2023), p. 80. DOI: 10.3847/1538-4357/aca5fc.

[304] R. Abbasi et al. "A Search for IceCube sub-TeV Neutrinos Correlated with Gravitational-Wave Events Detected By LIGO/Virgo" (Mar. 2023). arXiv: 2303.15970 [astro-ph.HE].

[305] Zhen Cao et al. "Ultrahigh-energy photons up to 1.4 petaelectronvolts from 12 γ-ray Galactic sources." *Nature* 594.7861 (2021), pp. 33–36. DOI: 10.1038/s41586-021-03498-z.

[306] R. Abbasi et al. "Searches for Neutrinos from Large High Altitude Air Shower Observatory Ultra-high-energy γ-Ray Sources Using the IceCube Neutrino Observatory." *Astrophys. J. Lett.* 945.1 (2023), p. L8. DOI: 10.3847/2041-8213/acb933.

[307] R. Abbasi et al. "Limits on Neutrino Emission from GRB 221009A from MeV to PeV Using the IceCube Neutrino Observatory." *Astrophys. J. Lett.* 946.1 (2023), p. L26. DOI: 10.3847/2041-8213/acc077.

[308] R. Abbasi et al. "Search for Astrophysical Neutrinos from 1FLE Blazars with IceCube." *Astrophys. J.* 938.1 (2022), p. 38. DOI: 10.3847/1538-4357/ac8de4.

[309] R. Abbasi et al. "All-flavor constraints on nonstandard neutrino interactions and generalized matter potential with three years of IceCube DeepCore data." *Phys. Rev. D* 104.7 (2021), p. 072006. DOI: 10.1103/PhysRevD.104.072006.

[310] R. Abbasi et al. "Search for quantum gravity using astrophysical neutrino flavour with IceCube." *Nature Phys.* 18.11 (2022), pp. 1287–1292. DOI: 10.1038/s41567-022-01762-1.

[311] Rasmus W. Rasmussen, Lukas Lechner, Markus Ackermann, Marek Kowalski, and Walter Winter. "Astrophysical neutrinos flavored with Beyond the Standard Model physics." *Phys. Rev. D* 96.8 (2017), p. 083018. DOI: 10.1103/PhysRevD.96.083018.

[312] Mauricio Bustamante and Sanjib Kumar Agarwalla. "Universe's Worth of Electrons to Probe Long-Range Interactions of High-Energy Astrophysical Neutrinos." *Phys. Rev. Lett.* 122.6 (2019), p. 061103. DOI: 10.1103/PhysRevLett.122.061103.

[313] Niki Klop and Shinichiro Ando. "Effects of a neutrino-dark energy coupling on oscillations of high-energy neutrinos." *Phys. Rev. D* 97.6 (2018), p. 063006. DOI: 10.1103/PhysRevD.97.063006.

[314] Yasaman Farzan and Sergio Palomares-Ruiz. "Flavor of cosmic neutrinos preserved by ultralight dark matter." *Phys. Rev. D* 99.5 (2019), p. 051702. DOI: 10.1103/PhysRevD.99.051702.

[315] Damiano F. G. Fiorillo, Gianpiero Mangano, Stefano Morisi, and Ofelia Pisanti. "IceCube constraints on violation of equivalence principle." *JCAP* 04 (2021), p. 079. DOI: 10.1088/1475-7516/2021/04/079.

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

[317] "South Pole glacial climate reconstruction from multi-borehole laser particulate stratigraphy." *J. Glaciol.* 59.218 (2013), pp. 1117–1128. DOI: 10.3189/2013JoG13J068.

[318] M. G. Aartsen et al. "Neutrino Interferometry for High-Precision Tests of Lorentz Symmetry with IceCube." *Nature Phys.* 14.9 (2018), pp. 961–966. DOI: 10.1038/s41567-018-0172-2.

[319] R. Abbasi et al. "Search for GeV neutrino emission during intense gamma-ray solar flares with the IceCube Neutrino Observatory." *Phys. Rev. D* 103.10 (2021), p. 102001. DOI: 10.1103/PhysRevD.103.102001.

[320] R. Abbasi et al. "Search for Relativistic Magnetic Monopoles with Eight Years of IceCube Data." *Phys. Rev. Lett.* 128.5 (2022), p. 051101. DOI: 10.1103/PhysRevLett.128.051101.

[321] P. Doll et al. "The Karlsruhe Cosmic Ray Project KASCADE" (Aug. 1990). URL: https://publikationen.bibliothek.kit.edu/270029106/18331842.

[322] Dieter Heck, Johannes Knapp, JN Capdevielle, G Schatz, T Thouw, et al. "CORSIKA: A Monte Carlo code to simulate extensive air showers." *Report fzka* 6019.11 (1998). URL: https://publikationen.bibliothek.kit.edu/270043064/3813660.

[323] Matthew A. Lazzara, George A. Weidner, Linda M. Keller, Jonathan E. Thom, and John J. Cassano. "Antarctic Automatic Weather Station Program: 30 Years of Polar Observation." Vol. 93. 10. American Meteorological Society, pp. 1519–1537. DOI: 10.1175/bams-d-11-00015.1.

[324] Thomas S. Pagano and Vivienne H. Payne. "The Atmospheric Infrared Sounder." *Handbook of Air Quality and Climate Change*. Springer Singapore, Dec. 2021, pp. 1–13. DOI: 10.1007/978-981-15-2527-8_64-1.

[325] Tim Huege. "CORSIKA 8 - the next-generation air shower simulation framework." *SciPost Physics Proceedings* 13 (Sept. 2023). DOI: 10.21468/scipostphysproc.13.028.

[326] Maximilian Reininghaus and Ralf Ulrich. "CORSIKA 8 – Towards a modern framework for the simulation of extensive air showers." Vol. 210. 2019, p. 02011. DOI: 10.1051/epjconf/201921002011.

[327] Torbjörn Sjöstrand. "The PYTHIA Event Generator: Past, Present and Future." *Comput. Phys. Commun.* 246 (2020), p. 106910. DOI: 10.1016/j.cpc.2019.106910.

[328] M. Bahr et al. "Herwig++ Physics and Manual." *Eur. Phys. J. C* 58 (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9.

[329] R. Engel. "Photoproduction within the two component dual parton model. 1. Amplitudes and cross-sections." *Z. Phys. C* 66 (1995), pp. 203–214. DOI: 10.1007/BF01496594.

[330] H Jung. "RAPGAP." *HERA and the LHC : A Workshop on the Implications of HERA for LHC Physics*. CERN, 2005. DOI: 10.5170/CERN-2005-014.601.

[331] Tanju Gleisberg, Stefan Hoeche, Frank Krauss, Andreas Schalicke, Steffen Schumann, and Jan-Christopher Winter. "SHERPA 1. alpha: A Proof of concept version." *JHEP* 02 (2004), p. 056. DOI: 10.1088/1126-6708/2004/02/056.

[332] S. Porteboeuf, T. Pierog, and K. Werner. "Producing Hard Processes Regarding the Complete Event: The EPOS Event Generator." *45th Rencontres de Moriond on QCD and High Energy Interactions*. 2010, pp. 135–140. arXiv: 1006.2967 [hep-ph].

[333] S. Baranov et al. "CASCADE3 A Monte Carlo event generator based on TMDs." *Eur. Phys. J. C* 81.5 (2021), p. 425. DOI: 10.1140/epjc/s10052-021-09203-8.

[334] Emil Avsar, Gosta Gustafson, and Leif Lonnblad. "Small-x dipole evolution beyond the large-N_c imit." *JHEP* 01 (2007), p. 012. DOI: 10.1088/1126-6708/2007/01/012.

[335] Miklos Gyulassy and Xin-Nian Wang. "HIJING 1.0: A Monte Carlo program for parton and particle production in high-energy hadronic and nuclear collisions." *Comput. Phys. Commun.* 83 (1994), p. 307. DOI: 10.1016/0010-4655(94)90057-4.

[336] S. Ostapchenko. "QGSJET-II: Towards reliable description of very high energy hadronic interactions." *Nucl. Phys. B Proc. Suppl.* 151 (2006), pp. 143–146. DOI: 10.1016/j.nuclphysbps.2005.07.026.

[337] Felix riehn, Ralph Engel, Anatoli Fedynitch, Thomas Gaisser, and Todor Stanev. "A new version of the event generator Sibyll." *Proceedings of The 34th International Cosmic Ray Conference — PoS(ICRC2015)*. Vol. 236. 2016, p. 558. DOI: 10.22323/1.236.0558.

[338] Stefan Roesler, Ralph Engel, and Johannes Ranft. "The Monte Carlo event generator DPMJET-III." Dec. 2000, pp. 1033–1038. DOI: 10.1007/978-3-642-18211-2_166.

[339] I. P. Lokhtin et al. "Heavy ion event generator HYDJET++ (HYDrodynamics plus JETs)." *Comput. Phys. Commun.* 180 (2009), pp. 779–799. DOI: 10.1016/j.cpc.2008.11.015.

[340] G Polesello et al. "Event generators for LHC." *Large Hadron Collider Workshop*. CERN, 1990. DOI: 10.5170/CERN-1990-010-V-2.130.

[341] W R Nelson, H Hirayama, and D W.O. Rogers. "EGS4 code system." Dec. 1985. URL: https://www.osti.gov/biblio/6137659.

[342] R. D. Parsons, C. Bleve, S. S. Ostapchenko, and J. Knapp. "Systematic uncertainties in air shower measurements from high-energy hadronic interaction models." *Astropart. Phys.* 34 (2011), pp. 832–839. DOI: 10.1016/j.astropartphys.2011.02.007.

[343] Ralf Ulrich, Ralph Engel, and Michael Unger. "Hadronic Multiparticle Production at Ultra-High Energies and Extensive Air Showers." *Phys. Rev. D* 83 (2011), p. 054026. DOI: 10.1103/PhysRevD.83.054026.

[344] J. Ranft. "The Dual parton model at cosmic ray energies." *Phys. Rev. D* 51 (1995), pp. 64–84. DOI: 10.1103/PhysRevD.51.64.

[345] Fritz W. Bopp, J. Ranft, R. Engel, and S. Roesler. "Antiparticle to Particle Production Ratios in Hadron-Hadron and d-Au Collisions in the DPMJET-III Monte Carlo." *Phys. Rev. C* 77 (2008), p. 014904. DOI: 10.1103/PhysRevC.77.014904.

[346] Klaus Werner, Fu-Ming Liu, and Tanguy Pierog. "Parton ladder splitting and the rapidity dependence of transverse momentum spectra in deuteron-gold collisions at RHIC." *Phys. Rev. C* 74 (2006), p. 044902. DOI: 10.1103/PhysRevC.74.044902.

[347] T. Pierog and K. Werner. "EPOS Model and Ultra High Energy Cosmic Rays." *Nucl. Phys. B Proc. Suppl.* 196 (2009), pp. 102–105. DOI: 10.1016/j.nuclphysbps.2009.09.017.

[348] T. Pierog, Iu. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner. "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider." *Phys. Rev. C* 92.3 (2015), p. 034906. DOI: 10.1103/PhysRevC.92.034906.

[349] T. Pierog, Iu. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner. "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider." *Phys. Rev. C* 92.3 (2015), p. 034906. DOI: 10.1103/PhysRevC.92.034906.

[350] AB Kaidalov and KA Ter-Martirosyan. *"Multiple Production of Hadrons at high energies in the model of quark-gluon strings"*. Tech. rep. Gosudarstvennyj Komitet po Ispolzovaniyu Atomnoj Ehnergii SSSR, 1983. URL: https://inis.iaea.org/search/search.aspx?orig_q=RN:15032104.

[351] NN Kalmykov and SS Ostapchenko. "The Nucleus-nucleus interaction, nuclear fragmentation, and fluctuations of extensive air showers." *Physics of Atomic Nuclei* 56.3 (1993), pp. 346–353. URL: https://inis.iaea.org/search/search.aspx?orig_q=RN:24066370.

[352] S. Ostapchenko. "Non-linear effects in high energy hadronic interactions" (Jan. 2005). arXiv: hep-ph/0501093.

[353] S. Ostapchenko. "On the re-summation of enhanced Pomeron diagrams." *Phys. Lett. B* 636 (2006), pp. 40–45. DOI: 10.1016/j.physletb.2006.03.026.

[354] Sergey Ostapchenko. "Status of QGSJET." *AIP Conf. Proc.* 928.1 (2007), pp. 118–125. DOI: 10.1063/1.2775904.

[355] Sergey Ostapchenko. "Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model." *Phys. Rev. D* 83 (2011), p. 014018. DOI: 10.1103/PhysRevD.83.014018.

[356] J. Engel, T. K. Gaisser, T. Stanev, and Paolo Lipari. "Nucleus-nucleus collisions and interpretation of cosmic ray cascades." *Phys. Rev. D* 46 (1992), pp. 5013–5025. DOI: 10.1103/PhysRevD.46.5013.

[357] R. S. Fletcher, T. K. Gaisser, Paolo Lipari, and Todor Stanev. "SIBYLL: An Event generator for simulation of high-energy cosmic ray cascades." *Phys. Rev. D* 50 (1994), pp. 5710–5731. DOI: 10.1103/PhysRevD.50.5710.

[358] Eun-Joo Ahn, Ralph Engel, Thomas K. Gaisser, Paolo Lipari, and Todor Stanev. "Cosmic ray interaction event generator SIBYLL 2.1." *Phys. Rev. D* 80 (2009), p. 094003. DOI: 10.1103/PhysRevD.80.094003.

[359] Anatoli Fedynitch, Felix Riehn, Ralph Engel, Thomas K. Gaisser, and Todor Stanev. "Hadronic interaction model SIBYLL 2.3c and inclusive lepton fluxes." *Phys. Rev. D* 100.10 (2019), p. 103018. DOI: 10.1103/PhysRevD.100.103018.

[360] Felix Riehn, Ralph Engel, Anatoli Fedynitch, Thomas K. Gaisser, and Todor Stanev. "The hadronic interaction model Sibyll 2.3c and muon production in extensive air-showers." *EPJ Web Conf.* 208 (2019), p. 11002. DOI: 10.1051/epjconf/201920811002.

[361] HS Fesefeldt. *"The Simulation of Hadronic Showers: Physics and Applications"*. Tech. rep. 1985. URL: http://cds.cern.ch/record/162911/.

[362] Alfredo Ferrari, Paola R. Sala, Alberto Fasso, and Johannes Ranft. "FLUKA: A multi-particle transport code (Program version 2005)" (Oct. 2005). DOI: 10.2172/877507.

[363] M Bleicher et al. "Relativistic hadron-hadron collisions in the ultra-relativistic quantum molecular dynamics model." *Journal of Physics G: Nuclear and Particle Physics* 25.9 (Sept. 1999), pp. 1859–1896. DOI: 10.1088/0954-3899/25/9/308.

[364] Sergey Ostapchenko. "Cosmic Ray Interaction Models: an Overview." *EPJ Web Conf.* 120 (2016), p. 04003. DOI: 10.1051/epjconf/201612004003.

[365] Tanguy Pierog. "Air Shower Simulation with a New Generation of post-LHC Hadronic Interaction Models in CORSIKA." *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Vol. 301. 2017, p. 1100. DOI: 10.22323/1.301.1100.

[366] A. Capella, U. Sukhatme, C-I Tan, and J. Tran Thanh Van. "Dual parton model." *Phys. Rept.* 236 (1994), pp. 225–329. DOI: 10.1016/0370-1573(94)90064-7.

[367] A. Capella and J. Tran Thanh Van. "Hadron - Nucleus Interactions and the Leading Particle Effect in a Dual Parton Model." *Z. Phys. C* 10 (1981), pp. 249–262. DOI: 10.1007/BF01549733.

[368] Hans-Uno Bengtsson and Torbjörn Sjöstrand. "The Lund Monte Carlo for hadronic processes — PYTHIA version 4.8." *Computer Physics Communications* 46.1 (July 1987), pp. 43–82. DOI: 10.1016/0010-4655(87)90036-1.

[369] L. Durand and H. Pi. "High-energy Nucleon Nucleus Scattering and Cosmic Ray Cross-sections." *Phys. Rev. D* 38 (1988), pp. 78–84. DOI: 10.1103/PhysRevD.38.78.

[370] T. K. Gaisser and F. Halzen. "Soft Hard Scattering in the TeV Range." *Phys. Rev. Lett.* 54 (1985), p. 1754. DOI: 10.1103/PhysRevLett.54.1754.

[371] G. Pancheri and Y. Srivastava. "Jets in minimum bias physics." *Physics Letters B* 159.1 (Sept. 1985), pp. 69–75. DOI: 10.1016/0370-2693(85)90121-2.

[372] Loyal Durand and Pi Hong. "QCD and Rising Total Cross-Sections." *Phys. Rev. Lett.* 58 (1987), pp. 303–306. DOI: 10.1103/PhysRevLett.58.303.

[373] R. J. Glauber and G. Matthiae. "High-energy scattering of protons by nuclei." *Nucl. Phys. B* 21 (1970), pp. 135–157. DOI: 10.1016/0550-3213(70)90511-0.

[374] Ralph Engel, Felix Riehn, Anatoli Fedynitch, Thomas K. Gaisser, and Todor Stanev. "The hadronic interaction model Sibyll – past, present and future." *EPJ Web Conf.* 145 (2017), p. 08001. DOI: 10.1051/epjconf/201614508001.

[375] Felix Riehn, Ralph Engel, Anatoli Fedynitch, Thomas K. Gaisser, and Todor Stanev. "Charm production in SIBYLL." *EPJ Web Conf.* 99 (2015), p. 12001. DOI: 10.1051/epjconf/20159912001.

[376] Ralph Engel. "Test of hadronic interaction models with data from the Pierre Auger Observatory." *30th International Cosmic Ray Conference*. Vol. 4. June 2007, pp. 385–388. arXiv: 0706.1921 [astro-ph].

[377] Alexander Aab et al. "Muons in Air Showers at the Pierre Auger Observatory: Mean Number in Highly Inclined Events." *Phys. Rev. D* 91.3 (2015), p. 032003. DOI: 10.1103/PhysRevD.91.032003.

[378] Alexander Aab et al. "Testing Hadronic Interactions at Ultrahigh Energies with Air Showers Measured by the Pierre Auger Observatory." *Phys. Rev. Lett.* 117.19 (2016), p. 192001. DOI: 10.1103/PhysRevLett.117.192001.

[379] Jakub Vícha. "Probing hadronic interaction models with the hybrid data of the Pierre Auger Observatory" (Sept. 2022). URL: https://scipost.org/10.21468/SciPostPhysProc.13.026.

[380] Stef Verpoest. "Studies of TeV muons in cosmic-ray air showers detected with IceTop and IceCube." PhD thesis. Ghent University, 2022. URL: https://biblio.ugent.be/publication/01GMWKFPF76N3GBC2CM7S7MN0X.

[381] Felix Riehn, Hans Dembinski, Anatoli Fedynitch, Ralph Engel, Thomas Gaisser, and Todor Stanev. "The hadronic interaction model Sibyll 2.3c and Feynman scaling." *Proceedings of 35th International Cosmic Ray Conference — PoS(ICRC2017)*. Vol. 301. 2017, p. 301. DOI: 10.22323/1.301.0301.

[382] H. J. Drescher, M. Hladik, S. Ostapchenko, and K. Werner. "A Unified treatment of high-energy interactions." *J. Phys. G* 25 (1999), pp. L91–L96. DOI: 10.1088/0954-3899/25/10/101.

[383] F. M. Liu, H. J. Drescher, S. Ostapchenko, T. Pierog, and K. Werner. "Consistent treatment of soft and hard processes in hadronic interactions." *J. Phys. G* 28 (2002), pp. 2597–2616. DOI: 10.1088/0954-3899/28/10/306.

[384] VN Gribov. "A Reggeon Diagram Technique." *Sov. Phys. JETP* 26.2 (1968), pp. 414–423. URL: https://ui.adsabs.harvard.edu/abs/1968JETP...26..414G.

[385] H.J. Drescher, M. Hladik, S. Ostapchenko, T. Pierog, and K. Werner. "Parton-based Gribov–Regge theory." *Physics Reports* 350.2-4 (Sept. 2001), pp. 93–289. DOI: 10.1016/s0370-1573(00)00122-8.

[386] M. Hladik, H. J. Drescher, S. Ostapchenko, T. Pierog, and K. Werner. "Self-Consistency Requirement in High-Energy Nuclear Scattering." *Phys. Rev. Lett.* 86 (2001), pp. 3506–3509. DOI: 10.1103/PhysRevLett.86.3506.

[387] A. B. Kaidalov and K. A. Ter-Martirosian. "Pomeron as Quark-Gluon Strings and Multiple Hadron Production at SPS Collider Energies." *Phys. Lett. B* 117 (1982), pp. 247–251. DOI: 10.1016/0370-2693(82)90556-1.

[388] N. N. Kalmykov, S. S. Ostapchenko, and A. I. Pavlov. "EAS and a quark - gluon string model with jets." *Bull. Russ. Acad. Sci. Phys.* 58 (1994), pp. 1966–1969.

[389] N. N. Kalmykov, S. S. Ostapchenko, and A. I. Pavlov. "Quark-Gluon String Model and EAS Simulation Problems at Ultra-High Energies." *Nucl. Phys. B Proc. Suppl.* 52 (1997), pp. 17–28. DOI: 10.1016/S0920-5632(96)00846-8.

[390] Felix Riehn, Ralph Engel, Anatoli Fedynitch, Thomas K. Gaisser, and Todor Stanev. "Hadronic interaction model Sibyll 2.3d and extensive air showers." *Phys. Rev. D* 102.6 (2020), p. 063002. DOI: 10.1103/PhysRevD.102.063002.

[391] Vardan Khachatryan et al. "Pseudorapidity distribution of charged hadrons in proton-proton collisions at $\sqrt{s} = 13$ TeV." *Phys. Lett. B* 751 (2015), pp. 143–163. DOI: 10.1016/j.physletb.2015.10.004.

[392] Tanguy Pierog. "Hadronic Interactions and Air Showers: Where Do We Stand?" *EPJ Web Conf.* 208 (2019), p. 02002. DOI: 10.1051/epjconf/201920802002.

[393] A. M. Hillas. "Shower simulation: Lessons from MOCCA." *Nucl. Phys. B Proc. Suppl.* 52 (1997), pp. 29–42. DOI: 10.1016/S0920-5632(96)00847-X.

[394] Matjaz Kobal. "Thinning of high-energy cosmic-ray air-showers." *26th International Cosmic Ray Conference*, vol. 1. 1999, p. 490. URL: https://ui.adsabs.harvard.edu/abs/1999ICRC....1..490K.

[395] Jim Madsen and Alessio Tamburro. *"Simulating UHE cosmic-ray showers in IceTop"*. IceCube - Internal Report, 2009.

[396] Pierre Billoir. "A sampling procedure to regenerate particles in a ground detector from a 'thinned' air shower simulation output." *Astropart. Phys.* 30 (2008), pp. 270–285. DOI: 10.1016/j.astropartphys.2008.10.002.

[397] Jeroen Van Houtte. "Air shower measurements with IceTop: The effect of (un) thinning simulation techniques." Master's thesis. Ghent University. URL: https://lib.ugent.be/catalog/rug01:002272587.

[398] S. Agostinelli et al. "Geant4—a simulation toolkit." *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[399] T Melzig. "Verbesserung der IceTop-Detektorsimulation unter Verwendung von Geant4." Master's thesis. Humboldt-Universität zu Berlin, 2011. URL: https://bib-pubdb1.desy.de/record/94371.

[400] Fabian Kislat, Thomas Melzig, Arne Van Overloop, and Tilo Waldenmaier. *"IceTop tank simulation using g4-tankresponse"*. IceCube - Internal Report, 2012.

[401] Arne Van Overloop. "Simulation of IceTop VEM calibration and the dependency on the snow layer." *Proceedings of 32nd International Cosmic Ray Conference — PoS(ICRC2011)*. Vol. 1. 2011, pp. 97–101. DOI: 10.7529/ICRC2011/V01/0899.

[402] J. H. Koehne, K. Frantzen, M. Schmitz, T. Fuchs, W. Rhode, D. Chirkin, and J. Becker Tjus. "PROPOSAL: A tool for propagation of charged leptons." *Comput. Phys. Commun.* 184 (2013), pp. 2070–2090. DOI: 10.1016/j.cpc.2013.04.001.

[403] Johan Lundberg et al. "Light tracking for glaciers and oceans: Scattering and absorption in heterogeneous media with Photonics." *Nucl. Instrum. Meth. A* 581 (2007), pp. 619–631. DOI: 10.1016/j.nima.2007.07.143.

[404] Nathan Whitehorn, Jakob van Santen, and Sven Lafebre. "Penalized Splines for Smooth Representation of High-dimensional Monte Carlo Datasets." *Comput. Phys. Commun.* 184 (2013), pp. 2214–2220. DOI: 10.1016/j.cpc.2013.04.008.

[405] Shahid Hussain, Todor Stanev, and Serap Tilav. "Measurements of the Air Shower Parameters with IceTop." *Proceedings of 32nd International Cosmic Ray Conference — PoS(ICRC2011)*. Vol. 1. 2011, pp. 243–246. DOI: 10.7529/ICRC2011/V01/0336.

[406] H Rebel. "Studies of different LDFs for primary energy estimation and mass discrimination of cosmic rays by the EAS lateral charged particle distribution as observed by KASCADE-Grande" (2005). URL: https://ui.adsabs.harvard.edu/abs/2005ICRC....6..297R.

[407] D. Barnhill et al. "Measurement of the lateral distribution function of UHECR air showers with the Pierre Auger Observatory." *Proceedings of 29th International Cosmic Ray Conference — PoS(ICRC2005)*. July 2005. URL: https://adsabs.harvard.edu/full/2005ICRC....7..291B.

[408] Charles C. H. Jui. "Cosmic Ray in the Northern Hemisphere: Results from the Telescope Array Experiment." *J. Phys. Conf. Ser.* 404 (2012), p. 012037. DOI: 10.1088/1742-6596/404/1/012037.

[409] M. A. Lawrence, R. J. O. Reid, and A. A. Watson. "The Cosmic ray energy spectrum above $4\text{x}10^{17}$-eV as measured by the Haverah Park array." *J. Phys. G* 17 (1991), pp. 733–757. DOI: 10.1088/0954-3899/17/5/019.

[410] Tom Feusels. "Measurement of cosmic ray composition and energy spectrum between 1 PeV and 1 EeV with IceTop and IceCube." PhD thesis. Ghent University, 2013. URL: https://biblio.ugent.be/publication/4337238.

[411] Koichi Kamata and Jun Nishimura. "The Lateral and the Angular Structure Functions of Electron Showers." *Progress of Theoretical Physics Supplement* 6 (1958), pp. 93–155. DOI: 10.1143/ptps.6.93.

[412] Todor Stanev. *"High Energy Cosmic Rays"*. Springer Berlin Heidelberg, 2010. DOI: 10.1007/978-3-540-85148-6.

[413] F Kislat, S Klepser, and A Van Overloop. *"A lateral distribution function and fluctuation parametrisation for IceTop"*. IceCube - Internal Report, 2007.

[414] Karen Grace Andeen. "First Measurements of Cosmic Ray Composition from 1-50 PeV using New Techniques on Coincident Data from the IceCube Neutrino Observatory." PhD thesis. University of Wisconsin-Madison, 2011. URL: https://search.library.wisc.edu/catalog/9910112854202121.

[415] Stefan Klepser. "Reconstruction of extensive air showers and measurement of the cosmic ray energy spectrum in the range of 1 PeV - 80 PeV at the South Pole." Master's thesis. June 2008. DOI: 10.18452/15790.

[416] Mark Weyrauch, IceCube Collaboration, et al. "A Two Component Lateral Distribution Function for the Reconstruction of Air-Shower Events with IceCube." 2023. URL: https://publikationen.bibliothek.kit.edu/1000156235.

[417] Javier Gonzalez. "Inclined Cosmic Ray Air Showers in IceCube." *Proceedings of 33rd International Cosmic Ray Conference — PoS(ICRC2013)*, p. 0973. URL: https://ui.adsabs.harvard.edu/abs/2013ICRC...33.1776G.

[418] Fabian Kislat. "Study of charge and time fluctuations of signals in the IceTop detector." Master's thesis. Humboldt-Universität zu Berlin, 2007. URL: https://bib-pubdb1.desy.de/record/83102/.

[419] A. L. Samuel. ""Some Studies in Machine Learning Using the Game of Checkers"." *IBM Journal of Research and Development* 3.3 (1959), pp. 210–229. DOI: 10.1147/rd.33.0210.

[420] Yoshua Bengio, Aaron Courville, and Pascal Vincent. ""Representation Learning: A Review and New Perspectives"." *IEEE Trans. Pattern Anal. Mach. Intell.* (2013), 1798–1828. DOI: 10.1109/TPAMI.2013.50.

[421] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *Nature* 521.7553 (May 2015), pp. 436–444. DOI: 10.1038/nature14539.

[422] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks." *Commun. ACM* (2017), 84–90. DOI: 10.1145/3065386.

[423] Clement Farabet, Camille Couprie, Laurent Najman, and Yann LeCun. "Learning Hierarchical Features for Scene Labeling." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.8 (2013), pp. 1915–1929. DOI: 10.1109/TPAMI.2012.231.

[424] Geoffrey Hinton et al. "Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups." *IEEE Signal Processing Magazine* 29.6 (2012), pp. 82–97. DOI: 10.1109/MSP.2012.2205597.

[425] Tomáš Mikolov, Anoop Deoras, Daniel Povey, Lukáš Burget, and Jan Černocký. "Strategies for training large scale neural network language models." *IEEE Workshop on Automatic Speech Recognition and Understanding*. 2011, pp. 196–201. DOI: 10.1109/ASRU.2011.6163930.

[426] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. "Sequence to Sequence Learning with Neural Networks." *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*. 2014, 3104–3112. URL: https://dl.acm.org/doi/10.5555/2969033.2969173.

[427] Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. "Natural Language Processing (Almost) from Scratch." *J. Mach. Learn. Res.* 12 (2011), 2493–2537. URL: https://dl.acm.org/doi/10.5555/1953048.2078186.

[428] Mark Newman. *Networks*. Oxford University Press, July 2018. DOI: 10.1093/oso/9780198805090.001.0001.

[429] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. “Graph neural networks: A review of methods and applications.” *AI Open* 1 (), pp. 57–81. DOI: https://doi.org/10.1016/j.aiopen.2021.01.001.

[430] William L. Hamilton. *"Graph Representation Learning"*. 2019. URL: https://www.cs.mcgill.ca/~wlh/grl_book/.

[431] Y. Zhao, Q. Ke, F. Korn, J. Qi, and R. Zhang. “HexCNN: A Framework for Native Hexagonal Convolutional Neural Networks.” *2020 IEEE International Conference on Data Mining (ICDM)*. 2020, pp. 1424–1429. DOI: 10.1109/ICDM50108.2020.00188.

[432] Stefan Schubert, Peer Neubert, Johannes Pöschmann, and Peter Protzel. “Circular Convolutional Neural Networks for Panoramic Images and Laser Data.” *2019 IEEE Intelligent Vehicles Symposium (IV)*. 2019, pp. 653–660. DOI: 10.1109/IVS.2019.8813862.

[433] Abraham Kandel, Horst Bunke, and Mark Last, eds. *"Applied Graph Theory in Computer Vision and Pattern Recognition"*. Springer Berlin Heidelberg, 2007. DOI: 10.1007/978-3-540-68020-8.

[434] Farzad V. Farahani, Waldemar Karwowski, and Nichole R. Lighthall. “Application of Graph Theory for Identifying Connectivity Patterns in Human Brain Networks: A Systematic Review.” *Frontiers in Neuroscience* 13 (June 2019). DOI: 10.3389/fnins.2019.00585.

[435] Ferozuddin Riaz and Khidir M. Ali. “Applications of Graph Theory in Computer Science.” *Third International Conference on Computational Intelligence, Communication Systems and Networks*. 2011, pp. 142–145. DOI: 10.1109/CICSyN.2011.40.

[436] Mikaela Koutrouli, Evangelos Karatzas, David Paez-Espino, and Georgios A. Pavlopoulos. “A Guide to Conquer the Biological Network Era Using Graph Theory.” *Frontiers in Bioengineering and Biotechnology* 8 (Jan. 2020). DOI: 10.3389/fbioe.2020.00034.

[437] Bruno Scalzo Dees, Yao Lei Xu, Anthony G. Constantinides, and Danilo P. Mandic. “Graph Theory for Metro Traffic Modelling.” *International Joint Conference on Neural Networks (IJCNN)*. 2021, pp. 1–5. DOI: 10.1109/IJCNN52387.2021.9533420.

[438] Peter W Battaglia et al. “Relational inductive biases, deep learning, and graph networks” (2018). URL: https://research.google/pubs/pub47094/.

[439] Michael M. Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. “Geometric Deep Learning: Going beyond Euclidean data.” *IEEE Signal Processing Magazine* 34.4 (2017), pp. 18–42. DOI: 10.1109/MSP.2017.2693418.

[440] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. “Learning representations by back-propagating errors.” *Nature* 323.6088 (Oct. 1986), pp. 533–536. DOI: 10.1038/323533a0.

[441] Geoffrey Hinton. *"The Forward-Forward Algorithm: Some Preliminary Investigations"*. DOI: 10.48550/ARXIV.2212.13345.

[442] Weihua Hu et al. "Open Graph Benchmark: Datasets for Machine Learning on Graphs." *Proceedings of the 34th International Conference on Neural Information Processing Systems*. NIPS'20. 2020. URL: https://ogb.stanford.edu/.

[443] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. "ImageNet: A large-scale hierarchical image database." *IEEE Conference on Computer Vision and Pattern Recognition*. 2009, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.

[444] Matthias Fey and Jan Eric Lenssen. *"Fast Graph Representation Learning with PyTorch Geometric"*. DOI: 10.48550/ARXIV.1903.02428.

[445] Minjie Yu Wang. "Deep graph library: Towards efficient and scalable deep learning on graphs." *ICLR workshop on representation learning on graphs and manifolds*. 2019. URL: https://rlgm.github.io/papers/49.pdf.

[446] Oleksandr Ferludin et al. "TF-GNN: graph neural networks in TensorFlow" (2022). URL: https://doi.org/10.48550/arXiv.2207.03522.

[447] Petar Veličković. "Everything is connected: Graph neural networks." *Current Opinion in Structural Biology* 79 (Apr. 2023), p. 102538. DOI: 10.1016/j.sbi.2023.102538.

[448] Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. "Geometric deep learning: Grids, groups, graphs, geodesics, and gauges" (2021). URL: https://geometricdeeplearning.com/.

[449] Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. "Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering." *Advances in Neural Information Processing Systems*. Vol. 29. 2016. URL: https://proceedings.neurips.cc/paper_files/paper/2016/file/04df4d434d481c5bb723be1b6df1ee65-Paper.pdf.

[450] Chaitanya Joshi. "Transformers are graph neural networks." *The Gradient* 12 (2020). URL: https://thegradient.pub/transformers-are-graph-neural-networks/.

[451] John Jumper et al. "Highly accurate protein structure prediction with AlphaFold." *Nature* 596.7873 (July 2021), pp. 583–589. DOI: 10.1038/s41586-021-03819-2.

[452] William L. Hamilton, Rex Ying, and Jure Leskovec. "Inductive Representation Learning on Large Graphs." NIPS'17 (2017), 1025–1035. URL: https://dl.acm.org/doi/10.5555/3294771.3294869.

[453] Thomas N. Kipf and Max Welling. *Semi-Supervised Classification with Graph Convolutional Networks*. 2016. DOI: 10.48550/ARXIV.1609.02907.

[454] Hogun Park and Jennifer Neville. "Exploiting Interaction Links for Node Classification with Deep Graph Neural Networks." International Joint Conferences on Artificial Intelligence Organization, July 2019, pp. 3223–3230. DOI: 10.24963/ijcai.2019/447.

[455] Namyong Park, Andrey Kan, Xin Luna Dong, Tong Zhao, and Christos Faloutsos. "Estimating Node Importance in Knowledge Graphs Using Graph Neural Networks." *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2019, 596–606. DOI: 10.1145/3292500.3330855.

[456] Rex Ying, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L. Hamilton, and Jure Leskovec. "Graph Convolutional Neural Networks for Web-Scale Recommender Systems." *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2018, 974–983. DOI: 10.1145/3219819.3219890.

[457] Marinka Zitnik, Monica Agrawal, and Jure Leskovec. "Modeling polypharmacy side effects with graph convolutional networks." *Bioinformatics* 34.13 (June 2018), pp. i457–i466. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/bty294.

[458] Rui Wang, Bicheng Li, Shengwei Hu, Wenqian Du, and Min Zhang. "Knowledge Graph Embedding via Graph Attenuated Attention Networks." *IEEE Access* 8 (2020), pp. 5212–5224. DOI: 10.1109/ACCESS.2019.2963367.

[459] Bo Wang, Tao Shen, Guodong Long, Tianyi Zhou, Ying Wang, and Yi Chang. "Structure-Augmented Text Representation Learning for Efficient Knowledge Graph Completion." *Proceedings of the Web Conference 2021*. Association for Computing Machinery, 2021, 1737–1748. DOI: 10.1145/3442381.3450043.

[460] Monica Agrawal, Marinka Zitnik, and Jure Leskovec. "Large-scale analysis of disease pathways in the human interactome." World Scientific. 2018, pp. 111–122. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5731453/.

[461] Shashank Pandit, Duen Horng Chau, Samuel Wang, and Christos Faloutsos. "Netprobe: A Fast and Scalable System for Fraud Detection in Online Auction Networks." *Proceedings of the 16th International Conference on World Wide Web*. 2007, 201–210. DOI: 10.1145/1242572.1242600.

[462] Chun Wang, Shirui Pan, Ruiqi Hu, Guodong Long, Jing Jiang, and Chengqi Zhang. "Attributed Graph Clustering: A Deep Attentional Embedding Approach" (July 2019), pp. 3670–3676. DOI: 10.24963/ijcai.2019/509.

[463] Xiaotong Zhang, Han Liu, Qimai Li, and Xiao-Ming Wu. "Attributed Graph Clustering via Adaptive Graph Convolution" (July 2019), pp. 4327–4333. DOI: 10.24963/ijcai.2019/601.

[464] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. "Neural Message Passing for Quantum Chemistry." *Proceedings of the 34th International Conference on Machine Learning - Volume 70*. 2017, 1263–1272. URL: https://dl.acm.org/doi/10.5555/3305381.3305512.

[465] Yujia Li, Chenjie Gu, Thomas Dullien, Oriol Vinyals, and Pushmeet Kohli. "Graph matching networks for learning the similarity of graph structured objects." *International conference on machine learning*. 2019, pp. 3835–3845. URL: https://proceedings.mlr.press/v97/li19d.html.

[466] Jonathan M. Stokes et al. "A Deep Learning Approach to Antibiotic Discovery." *Cell* 180.4 (2020), 688–702.e13. DOI: https://doi.org/10.1016/j.cell.2020.01.021.

[467] Jia Li, Yu Rong, Hong Cheng, Helen Meng, Wenbing Huang, and Junzhou Huang. "Semi-Supervised Graph Classification: A Hierarchical Graph Perspective." *The World Wide Web Conference*. 2019, 972–982. DOI: 10.1145/3308558.3313461.

[468] Fan-Yun Sun, Jordan Hoffmann, Vikas Verma, and Jian Tang. "InfoGraph: Unsupervised and Semi-supervised Graph-Level Representation Learning via Mutual Information Maximization" (2019). URL: https://github.com/sunfanyunn/InfoGraph.

[469] David Buterez, Jon Paul Janet, Steven J. Kiddle, Dino Oglic, and Pietro Liò. "Graph Neural Networks with Adaptive Readouts" (2022). DOI: 10.48550/ARXIV.2211.04952.

[470] Kai Han, Yunhe Wang, Jianyuan Guo, Yehui Tang, and Enhua Wu. "Vision GNN: An Image is Worth Graph of Nodes" (2022). eprint: 2206.00272.

[471] Xu andothers Ma. "Image as Set of Points" (2023). URL: https://github.com/ma-xu/Context-Cluster.

[472] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep Residual Learning for Image Recognition." *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

[473] Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio. "On the Difficulty of Training Recurrent Neural Networks" (2013), III–1310–III–1318. URL: https://dl.acm.org/doi/10.5555/3042817.3043083.

[474] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting." *J. Mach. Learn. Res.* 15.1 (2014), 1929–1958. URL: https://dl.acm.org/doi/10.5555/2627435.2670313.

[475] Brady Neal. *"On the Bias-Variance Tradeoff: Textbooks Need an Update"*. DOI: 10.48550/ARXIV.1912.08286.

[476] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. *Explaining and Harnessing Adversarial Examples*. 2015. eprint: 1412.6572.

[477] Eran Malach and Shai Shalev-Shwartz. "Decoupling 'When to Update' from 'How to Update'." NIPS'17 (2017), 961–971. URL: https://dl.acm.org/doi/10.5555/3294771.3294863.

[478] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. "Multilayer feedforward networks are universal approximators." *Neural Networks* 2.5 (1989), pp. 359–366. DOI: https://doi.org/10.1016/0893-6080(89)90020-8.

[479] Wikipedia contributors. *Time complexity — Wikipedia, The Free Encyclopedia*. [Online; accessed 6-November-2022]. 2022. URL: https://en.wikipedia.org/w/index.php?title=Time_complexity&oldid=1119120154.

[480] Wikipedia contributors. *NP-completeness — Wikipedia, The Free Encyclopedia*. [Online; accessed 6-November-2022]. 2022. URL: https://en.wikipedia.org/w/index.php?title=NP-completeness&oldid=1115151033.

[481] Boris Weisfeiler and Andrei Leman. "The reduction of a graph to canonical form and the algebra which appears therein." *NTI, Series* 2.9 (1968), pp. 12–16. URL: https://api.semanticscholar.org/CorpusID:49579538.

[482] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. *"How Powerful are Graph Neural Networks?"* 2019. URL: https://openreview.net/forum?id=ryGs6iA5Km.

[483] Christopher Morris et al. “Weisfeiler and Leman Go Neural: Higher-Order Graph Neural Networks.” AAAI Press, 2019. DOI: 10.1609/aaai.v33i01.33014602.

[484] Haggai Maron, Heli Ben-Hamu, Nadav Shamir, and Yaron Lipman. “Invariant and Equivariant Graph Networks.” DOI: 10.48550/ARXIV.1812.09902.

[485] R. Sato. *"A Survey on The Expressive Power of Graph Neural Networks"*. 2020. URL: https://api.semanticscholar.org/CorpusID:212634030.

[486] Vijay Prakash Dwivedi, Chaitanya K Joshi, Thomas Laurent, Yoshua Bengio, and Xavier Bresson. “Benchmarking graph neural networks” (2020). URL: https://github.com/graphdeeplearning/benchmarking-gnns.

[487] Giorgos Bouritsas et al. “Improving Graph Neural Network Expressivity via Subgraph Isomorphism Counting.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45.1 (2023), pp. 657–668. DOI: 10.1109/TPAMI.2022.3154319.

[488] Dumitru Erhan, Yoshua Bengio, Aaron Courville, and Pascal Vincent. “Visualizing higher-layer features of a deep network.” *University of Montreal* 1341.3 (2009), p. 1. URL: https://api.semanticscholar.org/CorpusID:15127402.

[489] Matthew D. Zeiler and Rob Fergus. “Visualizing and Understanding Convolutional Networks.” *Computer Vision – ECCV 2014*. 2014, pp. 818–833. DOI: 10.1007/978-3-319-10590-1_53.

[490] Bolei Zhou, David Bau, Aude Oliva, and Antonio Torralba. “Interpreting Deep Visual Representations via Network Dissection.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41.9 (2019), pp. 2131–2145. DOI: 10.1109/TPAMI.2018.2858759.

[491] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. *"Object Detectors Emerge in Deep Scene CNNs"*. DOI: 10.48550/ARXIV.1412.6856.

[492] Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals. “Understanding Deep Learning (Still) Requires Rethinking Generalization.” *Commun. ACM* 64.3 (2021), 107–115. DOI: 10.1145/3446776.

[493] Behnam Neyshabur, Ryota Tomioka, and Nathan Srebro. *"In Search of the Real Inductive Bias: On the Role of Implicit Regularization in Deep Learning"*. DOI: 10.48550/ARXIV.1412.6614.

[494] Benjamin D. Haeffele and René Vidal. “Global Optimality in Neural Network Training.” *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017, pp. 4390–4398. DOI: 10.1109/CVPR.2017.467.

[495] Andre Araujo, Wade Norris, and Jack Sim. “Computing Receptive Fields of Convolutional Neural Networks.” *Distill* (2019). DOI: 10.23915/distill.00021.

[496] Fisher Yu and Vladlen Koltun. *"Multi-Scale Context Aggregation by Dilated Convolutions"*. DOI: 10.48550/ARXIV.1511.07122.

[497] Kenta Oono and Taiji Suzuki. *"Graph Neural Networks Exponentially Lose Expressive Power for Node Classification"*. DOI: 10.48550/ARXIV.1905.10947.

[498] Qimai Li, Zhichao Han, and Xiao-Ming Wu. “Deeper Insights into Graph Convolutional Networks for Semi-Supervised Learning.” AAAI Press, 2018. URL: https://dl.acm.org/doi/10.5555/3504035.3504468.

[499] Deli Chen, Yankai Lin, Wei Li, Peng Li, Jie Zhou, and Xu Sun. "Measuring and Relieving the Over-Smoothing Problem for Graph Neural Networks from the Topological View." Vol. 34. 04. 2020, pp. 3438–3445. DOI: 10.1609/aaai.v34i04.5747.

[500] C. Lee Giles, Kurt D. Bollacker, and Steve Lawrence. "CiteSeer: An Automatic Citation Indexing System." *Proceedings of the Third ACM Conference on Digital Libraries*. 1998, 89–98. DOI: 10.1145/276675.276685.

[501] Yu Rong, Wenbing Huang, Tingyang Xu, and Junzhou Huang. "DropEdge: Towards Deep Graph Convolutional Networks on Node Classification." DOI: 10.48550/ARXIV.1907.10903.

[502] Lingxiao Zhao and Leman Akoglu. "PairNorm: Tackling Oversmoothing in GNNs." DOI: 10.48550/ARXIV.1909.12223.

[503] Kuangqi Zhou et al. "Understanding and Resolving Performance Degradation in Deep Graph Convolutional Networks." Association for Computing Machinery, 2021, 2728–2737. DOI: 10.1145/3459637.3482488.

[504] Kaixiong Zhou, Xiao Huang, Yuening Li, Daochen Zha, Rui Chen, and Xia Hu. "Towards deeper graph neural networks with differentiable group normalization." *Advances in neural information processing systems* 33 (2020), pp. 4917–4928. URL: https://dl.acm.org/doi/abs/10.5555/3495724.3496137.

[505] Keyulu Xu et al. "Representation Learning on Graphs with Jumping Knowledge Networks." *Proceedings of the 35th International Conference on Machine Learning*. Vol. 80. Proceedings of Machine Learning Research. 2018, pp. 5453–5462. URL: https://proceedings.mlr.press/v80/xu18c.html.

[506] S. Gong, M. Bahri, M. M. Bronstein, and S. Zafeiriou. "Geometrically Principled Connections in Graph Neural Networks." *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020, pp. 11412–11421. DOI: 10.1109/CVPR42600.2020.01143.

[507] Guohao Li, Matthias Müller, Ali Thabet, and Bernard Ghanem. "DeepGCNs: Can GCNs Go As Deep As CNNs?" *IEEE/CVF International Conference on Computer Vision (ICCV)*. 2019, pp. 9266–9275. DOI: 10.1109/ICCV.2019.00936.

[508] Uri Alon and Eran Yahav. "On the bottleneck of graph neural networks and its practical implications." URL: https://openreview.net/forum?id=i80OPhOCVH2.

[509] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. "Densely Connected Convolutional Networks." *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017, pp. 2261–2269. DOI: 10.1109/CVPR.2017.243.

[510] C. Szegedy et al. "Going deeper with convolutions." *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, 2015, pp. 1–9. DOI: 10.1109/CVPR.2015.7298594.

[511] Sergey Ioffe and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *International conference on machine learning*. 2015, pp. 448–456. URL: https://dl.acm.org/doi/10.5555/3045118.3045167.

[512] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. "Deep Sparse Rectifier Neural Networks." Vol. 15. Proceedings of Machine Learning Research. 2011, pp. 315–323. URL: https://proceedings.mlr.press/v15/glorot11a.html.

[513] Y. LeCun and others. "Backpropagation Applied to Handwritten Zip Code Recognition." *Neural Computation* 1.4 (1989), pp. 541–551. DOI: 10.1162/neco.1989.1.4.541.

[514] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition." *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324. DOI: 10.1109/5.726791.

[515] Rupesh Kumar Srivastava, Klaus Greff, and Jürgen Schmidhuber. "Training Very Deep Networks." *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*. 2015, 2377–2385. URL: https://dl.acm.org/doi/10.5555/2969442.2969505.

[516] Gao Huang, Yu Sun, Zhuang Liu, Daniel Sedra, and Kilian Q. Weinberger. "Deep Networks with Stochastic Depth." *Computer Vision – ECCV 2016*. 2016, pp. 646–661. DOI: 10.1007/978-3-319-46493-0_39.

[517] Gustav Larsson, Michael Maire, and Gregory Shakhnarovich. *"FractalNet: Ultra-Deep Neural Networks without Residuals"*. DOI: 10.48550/ARXIV.1605.07648.

[518] Andreas Veit, Michael Wilber, and Serge Belongie. "Residual Networks Behave like Ensembles of Relatively Shallow Networks." *Proceedings of the 30th International Conference on Neural Information Processing Systems*. 2016, 550–558. URL: https://dl.acm.org/doi/10.5555/3157096.3157158.

[519] Wenting Zhao et al. "When Work Matters: Transforming Classical Network Structures to Graph CNN" (2018). DOI: 10.48550/ARXIV.1807.02653.

[520] Min Lin, Qiang Chen, and Shuicheng Yan. *"Network In Network"*. DOI: 10.48550/ARXIV.1312.4400.

[521] Thao Nguyen, Maithra Raghu, and Simon Kornblith. *"Do Wide and Deep Networks Learn the Same Things? Uncovering How Neural Network Representations Vary with Width and Depth"*. DOI: 10.48550/ARXIV.2010.15327.

[522] IceCube Collaboration and Julian Saffer. *Low-Energy Cosmic Ray Composition with IceCube and IceTop*. Vortrag gehalten auf DPG-Frühjahrstagung der Sektion Materie und Kosmos (SMuK 2022). 2022.

[523] Jessie Micallef. "Using convolutional neural networks to reconstruct energy of GeV scale IceCube neutrinos." *JINST* 16.09 (2021), p. C09019. DOI: 10.1088/1748-0221/16/09/C09019.

[524] S. Yu. "Direction reconstruction using a CNN for GeV-scale neutrinos in IceCube." *JINST* 16.11 (2021), p. C11001. DOI: 10.1088/1748-0221/16/11/C11001.

[525] Jonas Uhrig et al. *"Sparsity Invariant CNNs"*. DOI: 10.48550/ARXIV.1708.06500.

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

[527] Nicholas Choma et al. *"Graph Neural Networks for IceCube Signal Classification"*. DOI: 10.48550/ARXIV.1809.06166.

[528] Alfred Lord Tennyson. *Ulysses*. 1950. URL: https://www.poetryfoundation.org/poems/45392/ulysses.

[529] Dennis Soldin, Roxanne Turcotte-Tardif, and Frank Schröder. "The Surface Array Enhancement of the IceCube Neutrino Observatory." *TeV Particle Astrophysics conference (TeVPA 2022)*. 2022. URL: https://publikationen.bibliothek.kit.edu/1000152108.

[530] Andres Medina. "Reconstruction of X_{max} and Energy from 3–100 PeV Using 5 Years of Data from IceTop and IceCube and Its Applications." PhD thesis. The Ohio State University, 2021. URL: http://rave.ohiolink.edu/etdc/view?acc_num=osu1617101771880381.

[531] Lorenzo Cazon, Ruben Conceição, and Felix Riehn. "Probing the energy spectrum of hadrons in proton air interactions at ultrahigh energies through the fluctuations of the muon content of extensive air showers." *Phys. Lett. B* 784 (2018), pp. 68–76. DOI: 10.1016/j.physletb.2018.07.026.

[532] 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." *Phys. Rev. D* 86 (2012), p. 114024. DOI: 10.1103/PhysRevD.86.114024.

[533] T. K. Gaisser and M. Honda. "Flux of atmospheric neutrinos." *Ann. Rev. Nucl. Part. Sci.* 52 (2002), pp. 153–199. DOI: 10.1146/annurev.nucl.52.050102.090645.

[534] W. D. Apel et al. "KASCADE-Grande measurements of energy spectra for elemental groups of cosmic rays." *Astropart. Phys.* 47 (2013), pp. 54–66. DOI: 10.1016/j.astropartphys.2013.06.004.

[535] S. F. Berezhnev et al. "Tunka-133: Primary Cosmic Ray Mass Composition in the Energy Range $6 \cdot 10^{15} - 10^{18}$ eV." *32nd International Cosmic Ray Conference*. Vol. 1. 2011, pp. 197–200. DOI: 10.7529/ICRC2011/V01/0184.

[536] Keitaro Fujita. "Cosmic ray mass composition measurement with the TALE hybrid detector." Vol. 283. 2023, p. 02009. DOI: 10.1051/epjconf/202328302009.

[537] Stanislav Knurenko and Igor Petrov. "Mass composition of cosmic rays above 0.1 EeV by the Yakutsk array data." *Adv. Space Res.* 64.12 (2019), pp. 2570–2577. DOI: 10.1016/j.asr.2019.07.019.

[538] Elke Van Den Broucke. "Chemische samenstelling van kosmische straling." Master's thesis. Ghent University. URL: https://lib.ugent.be/catalog/rug01:001805430.

[539] D. Heck and R. Engel. "The EHISTORY Option of the Air-Shower Simulation Program CORSIKA." Dec. 2009. URL: https://publikationen.bibliothek.kit.edu/270078292/3815642.

[540] *"On Language: On the Diversity of Human Language Construction and its Influence on the Mental Development of the Human Species"*. Cambridge University Press, 1999. URL: https://www.cambridge.org/9780521667722.

[541] Noam Chomsky. *"Aspects of the Theory of Syntax"*. Vol. 11. MIT press, 2014. URL: https://www.jstor.org/stable/j.ctt17kk81z.

[542] S. Thoudam, J. P. Rachen, A. van Vliet, A. Achterberg, S. Buitink, H. Falcke, and J. R. Hörandel. "Cosmic-ray energy spectrum and composition up to the ankle: the case for a second Galactic component." *Astron. Astrophys.* 595 (2016), A33. DOI: 10.1051/0004-6361/201628894.

[543] Jörg R. Hörandel. "The Composition of Cosmic Rays at the Knee." Vol. 1516. 1. 2013, pp. 185–194. DOI: 10.1063/1.4792566.

[544] Jörg R. Hörandel. "On the knee in the energy spectrum of cosmic rays." *Astroparticle Physics* 19.2 (May 2003), pp. 193–220. DOI: 10.1016/s0927-6505(02)00198-6.

[545] Domenik Ehlert, Foteini Oikonomou, and Michael Unger. "Curious case of the maximum rigidity distribution of cosmic-ray accelerators." *Phys. Rev. D* 107.10 (2023), p. 103045. DOI: 10.1103/PhysRevD.107.103045.

[546] Silvia Mollerach and Esteban Roulet. "Extragalactic cosmic rays diffusing from two populations of sources." *Phys. Rev. D* 101.10 (2020), p. 103024. DOI: 10.1103/PhysRevD.101.103024.

[547] Michael Unger, Glennys R. Farrar, and Luis A. Anchordoqui. "Origin of the ankle in the ultrahigh energy cosmic ray spectrum, and of the extragalactic protons below it." *Phys. Rev. D* 92.12 (2015), p. 123001. DOI: 10.1103/PhysRevD.92.123001.

[548] Noemie Globus, Denis Allard, and Etienne Parizot. "A complete model of the cosmic ray spectrum and composition across the Galactic to extragalactic transition." *Phys. Rev. D* 92.2 (2015), p. 021302. DOI: 10.1103/PhysRevD.92.021302.

[549] A. D. Supanitsky, A. Cobos, and A. Etchegoyen. "Origin of the light cosmic ray component below the ankle." *Phys. Rev. D* 98.10 (2018), p. 103016. DOI: 10.1103/PhysRevD.98.103016.

[550] A. A. Watson. "Further evidence for an increase of the mean mass of the highest-energy cosmic-rays with energy." *JHEAp* 33 (2022), pp. 14–19. DOI: 10.1016/j.jheap.2021.11.001.

[551] Katherine Rawlins. "Measuring the Composition of Cosmic Rays with the SPASE and AMANDA Detectors." PhD thesis. University of Wisconsin-Madison, 2001. URL: https://search.library.wisc.edu/catalog/999927248602121.

[552] K. S. Caballero Mora, M. Roth, Ioana C. Maris, and T. Schmidt. "Composition studies using the surface detector of the Pierre Auger Observatory." Vol. 116. 2008, p. 012007. DOI: 10.1088/1742-6596/116/1/012007.

[553] Alexander Aab et al. "Inferences on mass composition and tests of hadronic interactions from 0.3 to 100 EeV using the water-Cherenkov detectors of the Pierre Auger Observatory." *Phys. Rev. D* 96.12 (2017), p. 122003. DOI: 10.1103/PhysRevD.96.122003.

[554] A. D. Supanitsky, A. Etchegoyen, G. Medina-Tanco, I. Allekotte, M. Gomez Berisso, and Maria Clementina Medina. "Underground Muon Counters as a Tool for Composition Analyses." *Astropart. Phys.* 29 (2008), pp. 461–470. DOI: 10.1016/j.astropartphys.2008.05.003.

[555] A. D. Supanitsky, G. Medina-Tanco, and A. Etchegoyen. "On the possibility of primary identification of individual cosmic ray showers." *Astropart. Phys.* 31 (2009), pp. 116–127. DOI: 10.1016/j.astropartphys.2008.12.003.

[556] N. Hawthorne and G.P. Lathrop. *The Marble Faun: Or, The Romance of Monte Beni*. Houghton, Mifflin, 1883. URL: https://www.gutenberg.org/files/2181/2181-h/2181-h.htm.

[557] Martin Larocca, Frederic Sauvage, Faris M. Sbahi, Guillaume Verdon, Patrick J. Coles, and M. Cerezo. "Group-Invariant Quantum Machine Learning." *PRX Quantum* 3.3 (2022), p. 030341. DOI: 10.1103/PRXQuantum.3.030341.

[558] G. Cybenko. "Approximation by superpositions of a sigmoidal function." *Mathematics of Control, Signals, and Systems* 2.4 (Dec. 1989), pp. 303–314. DOI: 10.1007/bf02551274.

[559] Kurt Hornik. "Approximation capabilities of multilayer feedforward networks." *Neural Networks* (1991), pp. 251–257. DOI: 10.1016/0893-6080(91)90009-t.

[560] A.R. Barron. "Universal approximation bounds for superpositions of a sigmoidal function." *IEEE Transactions on Information Theory* 39.3 (1993), pp. 930–945. DOI: 10.1109/18.256500.

[561] Moshe Leshno, Vladimir Ya. Lin, Allan Pinkus, and Shimon Schocken. "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function." *Neural Networks* 6.6 (Jan. 1993), pp. 861–867. DOI: 10.1016/s0893-6080(05)80131-5.

[562] V.E Maiorov. "On Best Approximation by Ridge Functions." *Journal of Approximation Theory* 99.1 (July 1999), pp. 68–94. DOI: 10.1006/jath.1998.3304.

[563] Allan Pinkus. "Approximation theory of the MLP model in neural networks." *Acta numerica* 8 (1999), pp. 143–195. DOI: 10.1017/S0962492900002919.

[564] Tom M. Mitchell. "The Need for Biases in Learning Generalizations" (1980), pp. 184–191. URL: http://www.cs.nott.ac.uk/~bsl/G52HPA/articles/Mitchell:80a.pdf.

[565] Jürgen Schmidhuber. "Deep learning in neural networks: An overview." *Neural Networks* 61 (Jan. 2015), pp. 85–117. DOI: 10.1016/j.neunet.2014.09.003.

[566] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. URL: http://www.deeplearningbook.org.

[567] Sašo Džeroski, Luc De Raedt, and Kurt Driessens. "Relational reinforcement learning." *Machine Learning* (2001), pp. 7–52. DOI: 10.1023/a:1007694015589.

[568] Lise Getoor and Ben Taskar. *Introduction to statistical relational learning*. MIT press, 2007. URL: https://mitpress.mit.edu/9780262538688/introduction-to-statistical-relational-learning/.

[569] SS Nishana and Subu Surendran. "Graph embedding and dimensionality reduction - A survey." *International Journal of Computer Science & Engineering Technology (IJCSET)* (2013), pp. 29–34. URL: https://ijcset.com/docs/IJCSET13-04-01-021.pdf.

[570] Yufei Han and Yun Shen. "Partially Supervised Graph Embedding for Positive Unlabelled Feature Selection." *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*. 2016, 1548–1554. URL: https://dl.acm.org/doi/10.5555/3060832.3060837.

[571] Ming Yin, Junbin Gao, and Zhouchen Lin. "Laplacian Regularized Low-Rank Representation and Its Applications." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38.3 (2016), pp. 504–517. DOI: 10.1109/TPAMI.2015.2462360.

[572] Mengfan Tang, Feiping Nie, and Ramesh Jain. "Capped Lp-Norm Graph Embedding for Photo Clustering." *Proceedings of the 24th ACM International Conference on Multimedia*. 2016, 431–435. DOI: 10.1145/2964284.2967257.

[573] Mukund Balasubramanian and Eric L. Schwartz. "The Isomap Algorithm and Topological Stability." *Science* 295.5552 (2002), pp. 7–7. DOI: 10.1126/science.295.5552.7a.

[574] Xiang Ren, Wenqi He, Meng Qu, Clare R Voss, Heng Ji, and Jiawei Han. "Label noise reduction in entity typing by heterogeneous partial-label embedding." *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. 2016, pp. 1825–1834. URL: https://github.com/INK-USC/PLE.

[575] Isaac Henrion, Johann Brehmer, Joan Bruna, Kyunghyun Cho, Kyle Cranmer, Gilles Louppe, and Gaspar Rochette. "Neural message passing for jet physics" (2017). URL: https://dl4physicalsciences.github.io/files/nips_dlps_2017_29.pdf.

[576] Liang Sun, Shuiwang Ji, and Jieping Ye. "Hypergraph Spectral Learning for Multi-Label Classification." *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2008, 668–676. DOI: 10.1145/1401890.1401971.

[577] Shah Rukh Qasim, Jan Kieseler, Yutaro Iiyama, and Maurizio Pierini. "Learning representations of irregular particle-detector geometry with distance-weighted graph networks." *The European Physical Journal C* 79.7 (July 2019). DOI: 10.1140/epjc/s10052-019-7113-9.

[578] Jonathan Shlomi, Peter Battaglia, and Jean-Roch Vlimant. "Graph neural networks in particle physics." *Machine Learning: Science and Technology* 2.2 (2020), p. 021001. DOI: 10.1088/2632-2153/abbf9a.

[579] Teresa Bister, Martin Erdmann, Jonas Glombitza, Niklas Langner, Josina Schulte, and Marcus Wirtz. "Identification of patterns in cosmic-ray arrival directions using dynamic graph convolutional neural networks." *Astropart. Phys.* 126 (2021), p. 102527. DOI: 10.1016/j.astropartphys.2020.102527.

[580] Zewen Li, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects." *IEEE Transactions on Neural Networks and Learning Systems* (2021), pp. 1–21. DOI: 10.1109/TNNLS.2021.3084827.

[581] Oriol Vinyals, Samy Bengio, and Manjunath Kudlur. "Order Matters: Sequence to sequence for sets." DOI: 10.48550/ARXIV.1511.06391.

[582] Jinheon Baek, Minki Kang, and Sung Ju Hwang. "Accurate Learning of Graph Representations with Graph Multiset Pooling." DOI: 10.48550/ARXIV.2102.11533.

[583] E. Hemingway and C. Baker. *Ernest Hemingway Selected Letters 1917-1961*. Scribner Classics. Scribner, 2003. URL: https://books.google.de/books?id=SgKeQfmimJEC.

[584] Gary Marcus. "Deep Learning: A Critical Appraisal." DOI: 10.48550/ARXIV.1801.00631.

[585] Been Kim, Emily Reif, Martin Wattenberg, Samy Bengio, and Michael C. Mozer. "Neural Networks Trained on Natural Scenes Exhibit Gestalt Closure." *Computational Brain &amp Behavior* (Apr. 2021), pp. 251–263. DOI: 10.1007/s42113-021-00100-7.

[586] George Stockman and Linda G. Shapiro. *Computer Vision*. Prentice Hall PTR, 2001. URL: https://dl.acm.org/doi/book/10.5555/558008.

[587] JRR Toliken. *The Lord of the Ring: Fellowship of the Ring*. 1954. URL: https://tolkiengateway.net/wiki/Fellowship_of_the_Ring.

[588] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. 2nd. 1994. URL: https://dl.acm.org/doi/10.5555/562056.

[589] Miller McPherson, Lynn Smith-Lovin, and James M Cook. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27.1 (2001), pp. 415–444. DOI: 10.1146/annurev.soc.27.1.415.

[590] Derek Lim et al. *Large scale learning on non-homophilous graphs: New benchmarks and strong simple methods*. 2021. DOI: 10.48550/ARXIV.2110.14446.

[591] Jiong Zhu et al. "Beyond homophily in graph neural networks: Current limitations and effective designs." *Advances in Neural Information Processing Systems* 33 (2020), pp. 7793–7804. URL: https://dl.acm.org/doi/10.5555/3495724.3496377.

[592] Hongbin Pei, Bingzhe Wei, Kevin Chen-Chuan Chang, Yu Lei, and Bo Yang. "Geom-GCN: Geometric Graph Convolutional Networks." 2020. URL: https://github.com/bingzhewei/geom-gcn.

[593] Yao Ma, Xiaorui Liu, Neil Shah, and Jiliang Tang. "Is Homophily a Necessity for Graph Neural Networks?" 2021. DOI: 10.48550/ARXIV.2106.06134.

[594] Wendong Bi, Lun Du, Qiang Fu, Yanlin Wang, Shi Han, and Dongmei Zhang. "Make Heterophily Graphs Better Fit GNN: A Graph Rewiring Approach." 2022. DOI: 10.48550/ARXIV.2209.08264.

[595] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. "Graph attention networks." 2017. URL: https://petar-v.com/GAT/.

[596] Sami Abu-El-Haija et al. "MixHop: Higher-Order Graph Convolutional Architectures via Sparsified Neighborhood Mixing." *The Thirty-sixth International Conference on Machine Learning (ICML)*. 2019. URL: http://proceedings.mlr.press/v97/abu-el-haija19a/abu-el-haija19a.pdf.

[597] Haowen Deng. “Learned 3D Local Features for Rigid Pose Estimation on Point Clouds.” PhD thesis. Technische Universität München, 2020. URL: https://mediatum.ub.tum.de/doc/1549078/f0ldehlt264vo8qnfkab9qngf.Deng-Haowen.pdf.

[598] Haowen Deng, Tolga Birdal, and Slobodan Ilic. “PPFNet: Global Context Aware Local Features for Robust 3D Point Matching.” *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018, pp. 195–205. DOI: 10.1109/CVPR.2018.00028.

[599] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation.” *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 652–660. URL: https://github.com/charlesq34/pointnet.

[600] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. “Surface reconstruction from unorganized points.” *Proceedings of the 19th annual conference on computer graphics and interactive techniques*. 1992, pp. 71–78. URL: https://graphics.pixar.com/library/Reconstruction/.

[601] Bertram Drost, Markus Ulrich, Nassir Navab, and Slobodan Ilic. “Model globally, match locally: Efficient and robust 3D object recognition.” *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2010, pp. 998–1005. DOI: 10.1109/CVPR.2010.5540108.

[602] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. “Fast Point Feature Histograms (FPFH) for 3D registration.” *IEEE International Conference on Robotics and Automation*. 2009, pp. 3212–3217. DOI: 10.1109/ROBOT.2009.5152473.

[603] A.S. Mian, M. Bennamoun, and R. Owens. “Three-Dimensional Model-Based Object Recognition and Segmentation in Cluttered Scenes.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28.10 (2006), pp. 1584–1601. DOI: 10.1109/TPAMI.2006.213.

[604] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. “Attention is All you Need.” 30 (2017). URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

[605] Felix Wu, Angela Fan, Alexei Baevski, Yann N. Dauphin, and Michael Auli. “Pay Less Attention with Lightweight and Dynamic Convolutions.” 2019. DOI: 10.48550/ARXIV.1901.10430.

[606] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding” (June 2019), pp. 4171–4186. DOI: 10.18653/v1/N19-1423.

[607] Zihang Dai, Zhilin Yang, Yiming Yang, Jaime Carbonell, Quoc V Le, and Ruslan Salakhutdinov. “Transformer-xl: Attentive language models beyond a fixed-length context.” 2019. URL: https://github.com/kimiyoung/transformer-xl.

[608] Han Hu, Zheng Zhang, Zhenda Xie, and Stephen Lin. “Local Relation Networks for Image Recognition.” *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. 2019, pp. 3463–3472. DOI: 10.1109/ICCV.2019.00356.

[609] Prajit Ramachandran, Niki Parmar, Ashish Vaswani, Irwan Bello, Anselm Levskaya, and Jon Shlens. "Stand-Alone Self-Attention in Vision Models." 32 (2019). URL: https://proceedings.neurips.cc/paper_files/paper/2019/file/3416a75f4cea9109507cacd8e2f2aefc-Paper.pdf.

[610] Hengshuang Zhao, Jiaya Jia, and Vladlen Koltun. "Exploring Self-Attention for Image Recognition." *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020, pp. 10073–10082. DOI: 10.1109/CVPR42600.2020.01009.

[611] Alec Radford, Karthik Narasimhan, Tim Salimans, Ilya Sutskever, et al. *Improving language understanding by generative pre-training*. 2018. URL: https://openai.com/research/language-unsupervised.

[612] Long Ouyang et al. "Training language models to follow instructions with human feedback." *Advances in Neural Information Processing Systems* 35 (2022), pp. 27730–27744. URL: https://github.com/openai/following-instructions-human-feedback.

[613] Barry Schwartz. *Google: BERT now used on almost every English query*. [Online; accessed 11-April-2023]. URL: https://searchengineland.com/google-bert-used-on-almost-every-english-query-342193.

[614] Roger Monti. *Google's BERT Rolls Out Worldwide*. [Online; accessed 11-April-2023]. Sept. 12, 2019. URL: https://www.searchenginejournal.com/google-bert-rolls-out-worldwide/339359/.

[615] Volodymyr Mnih, Nicolas Heess, Alex Graves, and koray kavukcuoglu. "Recurrent Models of Visual Attention." *Advances in Neural Information Processing Systems*. Vol. 27. 2014. URL: https://proceedings.neurips.cc/paper_files/paper/2014/file/09c6c3783b4a70054da74f2538ed47c6-Paper.pdf.

[616] Thang Luong, Hieu Pham, and Christopher D. Manning. "Effective Approaches to Attention-based Neural Machine Translation" (Sept. 2015), pp. 1412–1421. DOI: 10.18653/v1/D15-1166.

[617] J. M. Henderson and G. L. Pierce. "Eye movements during scene viewing: Evidence for mixed control of fixation durations." *Psychonomic Bulletin &amp Review* 15.3 (June 2008), pp. 566–573. DOI: 10.3758/pbr.15.3.566.

[618] John M. Henderson and Andrew Hollingworth. "Eye Movements During Scene Viewing: An Overview" (1998), pp. 269–293. DOI: https://doi.org/10.1016/B978-008043361-5/50013-4.

[619] Saumya Jetley, Nicholas A. Lord, Namhoon Lee, and Philip H. S. Torr. "Learn To Pay Attention." 2018. DOI: 10.48550/ARXIV.1804.02391.

[620] J. Schmidhuber. "Reducing the Ratio Between Learning Complexity and Number of Time Varying Variables in Fully Recurrent Nets." *ICANN '93*. Springer London, 1993, pp. 460–463. DOI: 10.1007/978-1-4471-2063-6_110.

[621] Imanol Schlag, Kazuki Irie, and Jürgen Schmidhuber. "Linear transformers are secretly fast weight programmers." *International Conference on Machine Learning*. PMLR. 2021, pp. 9355–9366. URL: https://github.com/ischlag/fast-weight-transformers.

[622] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. "Neural Machine Translation by Jointly Learning to Align and Translate." 2014. DOI: 10.48550/ARXIV.1409.0473.

[623] Hengshuang Zhao, Li Jiang, Jiaya Jia, Philip HS Torr, and Vladlen Koltun. "Point transformer." *Proceedings of the IEEE/CVF international conference on computer vision*. 2021, pp. 16259–16268. URL: https://github.com/Pointcept/Pointcept.

[624] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, and Justin M. Solomon. "Dynamic Graph CNN for Learning on Point Clouds." Vol. 38. 5. Oct. 2019, pp. 1–12. URL: https://doi.org/10.1145/3326362.

[625] Chuang Liu, Yibing Zhan, Jia Wu, Chang Li, Bo Du, Wenbin Hu, Tongliang Liu, and Dacheng Tao. "Graph Pooling for Graph Neural Networks: Progress, Challenges, and Opportunities" (Aug. 2023), pp. 6712–6722. DOI: 10.24963/ijcai.2023/752.

[626] David K Duvenaud et al. "Convolutional networks on graphs for learning molecular fingerprints." Vol. 28. 2015. URL: https://github.com/HIPS/neural-fingerprint.

[627] Matthew M. Botvinick. "Hierarchical models of behavior and prefrontal function." *Trends in Cognitive Sciences* 12.5 (2008), pp. 201–208. DOI: https://doi.org/10.1016/j.tics.2008.02.009.

[628] Joshua B. Tenenbaum, Charles Kemp, Thomas L. Griffiths, and Noah D. Goodman. "How to Grow a Mind: Statistics, Structure, and Abstraction." *Science* 331.6022 (2011), pp. 1279–1285. DOI: 10.1126/science.1192788.

[629] Ilya Sutskever, James Martens, George Dahl, and Geoffrey Hinton. "On the Importance of Initialization and Momentum in Deep Learning." JMLR.org, 2013, III–1139–III–1147. URL: https://dl.acm.org/doi/10.5555/3042817.3043064.

[630] Michael A. Lones. "How to avoid machine learning pitfalls: a guide for academic researchers." 2021. DOI: 10.48550/ARXIV.2108.02497.

[631] Matthias Plum. "Cosmic ray composition study using machine learning at the IceCube Neutrino Observatory." *PoS* ICRC2019 (2020), p. 394. DOI: 10.22323/1.358.0394.

[632] Tianqi Chen and Carlos Guestrin. "XGBoost." ACM, Aug. 2016. DOI: 10.1145/2939672.2939785.

[633] Isabelle Guyon and André Elisseeff. "An Introduction to Variable and Feature Selection." *J. Mach. Learn. Res.* (2003), 1157–1182. ISSN: 1532-4435. URL: https://dl.acm.org/doi/10.5555/944919.944968.

[634] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer New York, 2009. DOI: 10.1007/978-0-387-84858-7.

[635] Kyle Cranmer. "Kernel estimation in high-energy physics." *Computer Physics Communications* 136.3 (2001), pp. 198–207. DOI: 10.1016/S0010-4655(00)00243-5.

[636] Roger Barlow and Christine Beeston. "Fitting using finite Monte Carlo samples." *Computer Physics Communications* 77.2 (1993), pp. 219–228. DOI: 10.1016/0010-4655(93)90005-W.

[637] F. James and M. Roos. "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations." *Comput. Phys. Commun.* 10 (1975), pp. 343–367. DOI: 10.1016/0010-4655(75)90039-9.

[638] J.S. Conway. "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra" (2011). DOI: 10.5170/CERN-2011-006.115.

[639] C. A. Argüelles, A. Schneider, and T. Yuan. "A binned likelihood for stochastic models." *Journal of High Energy Physics* 2019.6 (June 2019). DOI: 10.1007/jhep06(2019)030.

[640] Steve Baker and Robert D. Cousins. "Clarification of the use of CHI-square and likelihood functions in fits to histograms." *Nuclear Instruments and Methods in Physics Research* 221.2 (1984), pp. 437–442. DOI: 10.1016/0167-5087(84)90016-4.

[641] Z. I. Botev, J. F. Grotowski, and D. P. Kroese. "Kernel density estimation via diffusion." *The Annals of Statistics* 38.5 (2010), pp. 2916 –2957. DOI: 10.1214/10-AOS799.

[642] Jonas Eschle, Albert Puig Navarro, Rafael Silva Coutinho, and Nicola Serra. "zfit: Scalable pythonic fitting." *SoftwareX* 11 (2020), p. 100508. ISSN: 2352-7110. DOI: https://doi.org/10.1016/j.softx.2020.100508.

[643] Hans Dembinski and Piti Ongmongkolkul et al. "scikit-hep/iminuit" (2020). DOI: 10.5281/zenodo.3949207.

[644] S.S.A. Ashram. *Shrimad Bhagavad Gita (Hindi)*. Sant Shri Asharamji Ashram. URL: https://books.google.de/books?id=Yc_AAAAAQBAJ.

[645] Alberto Daniel Supanitsky. "Determination of the Cosmic-Ray Chemical Composition: Open Issues and Prospects." *Galaxies* 10.3 (2022), p. 75. DOI: 10.3390/galaxies10030075.

[646] LG Sveshnikova et al. "Cosmic Ray Spectrum Above the Knee Measured by the Tunka-133 Experiment: Special Features and Possible Interpretations." *International Cosmic Ray Conference*. Vol. 33. 2013, p. 1538. URL: https://ui.adsabs.harvard.edu/abs/2013ICRC...33.1538S.

[647] Frank J. Massey. "The Kolmogorov-Smirnov Test for Goodness of Fit." *Journal of the American Statistical Association* 46.253 (Mar. 1951), pp. 68–78. DOI: 10.1080/01621459.1951.10500769.

[648] Faiz Ahmad Faiz, Khwaja Tariq Mahmood, and Amar Varma. *Selected Poems of Faiz Ahmad Faiz: With Original Urdu Text, Roman and Hindi Transliteration and Poetical Translation Into English*. Star Publ., 2010. URL: https://search.library.wisc.edu/catalog/999935256402121.

[649] Dmitry Chirkin and Wolfgang Rhode. "Muon Monte Carlo: A High-precision tool for muon propagation through matter" (July 2004). arXiv: hep-ph/0407075. URL: http://icecube.wisc.edu/~dima/work/MUONPR/.

[650] Dmitry Chirkin. *dCORSIKA update*. IceCube - Internal Report, 2002.

[651] Jonathan Eisch. "A Multi-Variate Fit to the Chemical Composition of the Cosmic-Ray Spectrum." PhD thesis. The University of Wisconsin-Madison, 2014. URL: https://search.library.wisc.edu/catalog/9910217855102121.

[652] Ewa M. Holt, Frank G. Schröder, and Andreas Haungs. "Enhancing the cosmic-ray mass sensitivity of air-shower arrays by combining radio and muon detectors." *Eur. Phys. J. C* 79.5 (2019), p. 371. DOI: 10.1140/epjc/s10052-019-6859-4.

[653] Alexander Aab et al. "The Pierre Auger Observatory Upgrade - Preliminary Design Report." Apr. 2016. DOI: 10.48550/ARXIV.1604.03637.

[654] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Springer New York, 2002. DOI: 10.1007/978-0-387-21738-3.

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

[656] Suvrit Sra, Dongmin Kim, Inderjit Dhillon, and Bernhard Schölkopf. "A new non-monotonic algorithm for PET image reconstruction." *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*. 2009, pp. 2500–2502. DOI: 10.1109/NSSMIC.2009.5402060.

[657] Patrick Hallen. "On the Measurement of High-Energy Tau Neutrinos with IceCube." Master's thesis. RWTH Aachen, 2013. URL: http://tinyurl.com/yckx9zcz.

[658] MG Aartsen et al. "Characterization of the atmospheric muon flux in IceCube." *Astroparticle Physics* 78 (2016), pp. 1–27. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2016.01.006.

[659] Alex Sherstinsky. "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network." *Physica D: Nonlinear Phenomena* 404 (2020), p. 132306. ISSN: 0167-2789. DOI: https://doi.org/10.1016/j.physd.2019.132306.

DECLARATION

Erklärung der selbständigen Anfertigung meiner Dissertationsschrift

Hiermit versichere ich, dass ich die Dissertationsschrift mit dem Titel

Elemental Composition of Cosmic Rays: Analysis of IceCube data using Graph Neural Networks

selbständig und ohne unerlaubte fremde Hilfe verfasst habe. Dabei habe ich keine anderen, als die von mir angegebenen Hilfsmittel benutzt.

Ort und Datum

Paras Koundal