UNIVERSITÉ DE GENÈVE

Archive ouverte UNIGE
https://archive-ouverte.unige.ch

Thèse 2025

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Convolutional Neural Network-Based Methods for Non-fiducial Electron Reconstruction and TeV Cosmic-Electron Flux Measurement with the DAMPE Satellite Experiment

Putti-Garcia, Enzo

How to cite

PUTTI-GARCIA, Enzo. Convolutional Neural Network-Based Methods for Non-fiducial Electron Reconstruction and TeV Cosmic-Electron Flux Measurement with the DAMPE Satellite Experiment. Thèse, 2025. doi: 10.13097/archive-ouverte/unige:190701

This publication URL: https://archive-ouverte.unige.ch/unige:190701
Publication DOI: 10.13097/archive-ouverte/unige:190701

© This document is protected by copyright. Please refer to copyright holder(s) for terms of use.

UNIVERSITÉ DE GENÈVE
Département de Physique
Nucléaire et Corpusculaire

FACULTÉ DES SCIENCES
Professeur Xin Wu

Convolutional Neural Network-Based Methods for Non-fiducial Electron Reconstruction and TeV Cosmic-Electron Flux Measurement with the DAMPE Satellite Experiment

THÈSE

présentée à la Faculté des Sciences de l'Université de Genève
pour obtenir le grade de Docteur ès Sciences, mention Physique

par

Enzo Putti-Garcia
de
Ferney-Voltaire, France

Thèse n°5942

GENÈVE
Atelier d'impression ReproMail
2025

UNIVERSITÉ DE GENÈVE

FACULTÉ DES SCIENCES

DOCTORAT ÈS SCIENCES, MENTION PHYSIQUE

Thèse de Monsieur Enzo PUTTI-GARCIA

intitulée :

«Convolutional Neural Network-Based Methods for Non-fiducial Electron Reconstruction and TeV Cosmic-Electron Flux Measurement with the DAMPE Satellite Experiment»

La Faculté des sciences, sur le préavis de

Monsieur X. WU, professeur ordinaire et directeur de thèse
Département de physique nucléaire et corpusculaire

Madame M. PANICCIA, docteure
Département de physique nucléaire et corpusculaire

Madame C. PERRINA, professeure
Laboratoire de particules cosmiques, Faculté des sciences de base, École polytechnique fédérale de Lausanne, Lausanne, Suisse

autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 4 septembre 2025

Thèse - 5942 -

La Doyenne

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Résumé

À l'œil nu, les merveilles de l'Univers semblent se limiter aux planètes, étoiles et galaxies. Cependant, une autre facette invisible existe—composée de particules énergétiques appelées rayons cosmiques. Bien que leur découverte ait pris plus d'un siècle, de nombreuses énigmes les concernant demeurent non résolues. Des questions subsistent sur leur origine ainsi que sur les mécanismes responsables de leurs énergies extrêmes. Cela n'a pourtant pas freinée les avancées scientifiques : ces dernières décennies ont vu le développement d'expériences de plus en plus précises pour explorer ces questions et repousser les frontières de notre compréhension au-delà de ce qui était auparavant imaginable.

Un mystère particulièrement intriguant concerne les électrons cosmiques, qui semblent être plus abondants à hautes énergies que ce que prévoient les modèles actuels. Des observations similaires ont été faites pour les positrons—les équivalents en antimatière des électrons. En raison de leurs pertes d'énergie importantes, notamment par rayonnement synchrotron et diffusion Compton inverse, ces particules doivent provenir de sources proches de la Terre, ou peut-être avoire une origine exotique telles que la matière noire.

Le satellite *Dark Matter Particle Explorer* (DAMPE), lancé en 2015, a été conçu pour répondre à ces questions. Il est équipé de quatre sous-détecteurs permettant d'identifier les particules incidentes et de reconstruire leur direction ainsi que leur énergie. En 2017, DAMPE a publié sa première mesure du flux combiné d'électrons et de positrons entre 24 GeV et 4.6 TeV, révélant une cassure spectrale proche de 0.9 TeV qui a suscité de vifs débats au sein de la communauté scientifique. Toutefois, l'extension de cette mesure au-delà de 10 TeV est cruciale pour comprendre le comportement des électrons à très haute énergie. Cette extension s'avère cependant difficile, car les électrons de haute énergie sont rares, nécessitant de grandes quantités de données pour atteindre une signification statistique suffisante. Dans ce travail, nous proposons d'augmenter l'acceptance du détecteur en incluant les événements entrants par les côtés. Bien que plus complexes à analyser, ces événements peuvent considérablement améliorer la statistique. Leur analyse exige toutefois des outils plus sophistiqués.

Pour relever ces défis, nous utilisons des techniques d'apprentissage profond, et plus précisément des réseaux de neurones convolutifs (CNN). Une application clé concerne

la reconstruction de l'énergie. En sélectionnant des événements avec un grand angle d'incidence (comme dans l'acceptance élargie), les fuites d'énergie deviennent plus importantes, entraînant une sous-estimation—souvent d'au moins 5%—de l'énergie réelle de la particule. Nous montrons que l'utilisation de CNNs permet une amélioration significative de la reconstruction de l'énergie dans ce régime.

Une autre difficulté majeure est la présence massive du bruit de fond dû aux protons, bien plus abondants que les électrons. Il est donc essentiel de séparer correctement les électrons des protons. Nous montrons que la variable utilisée jusqu'à maintenant dans la collaboration DAMPE, ζ, est moins adaptées pour cette tâche. Pour y remédier, nous avons développé deux classificateurs CNN dédiés, optimisés respectivement pour les basses et hautes énergies. Les deux modèles améliorent le rejet du fond d'un ordre de grandeur par rapport à ζ.

Grâce aux outils d'apprentissage profond développés dans ce travail, nous avons pu mesurer le flux d'électrons et de positrons de 100 GeV jusqu'à 3 TeV. Avant de réaliser cette mesure, les données réelles ont été soumises à une série de coupes de sélection—validées à l'aide de simulations Monte Carlo—afin d'éliminer les événements indésirables tels que les protons, les noyaux d'hélium et les noyaux plus lourds. Après application des coupes de préselection et des classificateurs, une efficacité supérieure à 80% a été atteinte à toutes les énergies.

Cette mesure démontre la possibilité d'étendre l'acceptance de DAMPE en incluant les événements non-fiduciels dans la mesure du flux dans l'intervalle d'énergie entre 100 GeV et 5 TeV. Le flux obtenu est cohérent avec les résultats publiés précédemment par DAMPE.

Abstract

To the naked eye, the wonders of the Universe appear to be limited to planets, stars, and galaxies. However, another invisible aspect exists—one made up of high-energy particles known as cosmic rays. Although their discovery took over a century, many mysteries surrounding them remain unresolved. Questions regarding their origin and the mechanisms responsible for their extreme energies are still open. Possible sources include supernova remnants (SNRs), pulsar wind nebulae (PWNe), dark matter interactions, and active galactic nuclei (AGN). Nevertheless, this has not hindered scientific progress. In recent decades, increasingly precise experiments have been developed to probe these questions and push the boundaries of our understanding beyond what was once thought possible.

A particularly intriguing mystery involves cosmic-ray electrons, which appear to be more abundant at high energies than expected. Similar observations have been made for positrons—the antimatter counterparts of electrons. Due to significant energy losses through synchrotron radiation and inverse Compton scattering, such particles must originate from nearby sources, or possibly from an exotic origin such as dark matter.

The satellite *Dark Matter Particle Explorer* (DAMPE), launched in 2015, was designed to help answer these questions. It is equipped with four sub-detectors to identify incoming particles and reconstruct their direction and energy. In 2017, DAMPE published its first measurement of the combined electron+positron flux from 24 GeV up to 4.6 TeV, revealing a spectral break at ~ 0.9 TeV that sparked extensive discussions within the physics community. However, extending this measurement beyond 10 TeV is crucial to understanding high-energy behaviour. This extension is challenging, as high-energy electrons are rare, requiring large datasets to achieve statistically meaningful results. In this work, we propose to include side-entering events, thereby increasing the effective acceptance, which, while more complex to analyse, can significantly improve statistics. These events demand more sophisticated analysis techniques.

To address these challenges, we employ deep learning techniques, specifically convolutional neural networks (CNNs). One key application is energy reconstruction. When selecting events with larger incidence angles (as in the expanded acceptance), energy leakage becomes more significant, leading to underestimation—often by at least 5%—of the

particle's true energy. We show that CNNs significantly improve energy reconstruction in this regime.

Another major difficulty is the overwhelming background of protons, which are far more abundant than electrons. It is therefore critical to accurately separate electrons from protons. We demonstrate that the current variable, ζ, is insufficient for this task. To address this, we developed two dedicated CNN classifiers optimized for low and high energy regimes. Both models improve background rejection by more than an order of magnitude compared to ζ.

Thanks to the development of these deep learning tools, we were able to measure the electron+positron flux from 100 GeV up to 3 TeV. Before performing this measurement, real data underwent a series of selection cuts—validated using Monte Carlo simulations—to remove unwanted events such as protons, helium nuclei, and heavier nucleis. After applying both preselection and classifier cuts, we achieved an efficiency above 80% across all energies.

This measurement demonstrates the potential to extend DAMPE's acceptance by including non-fiducial events in the flux measurement. The resulting flux is consistent with previously published DAMPE results.

Acknowledgements / Remerciements

I would like to express my deepest gratitude to Prof. Xin Wu, my supervisor, for his continuous guidance, support, and insightful discussions throughout my research journey. His expertise and mentorship have been invaluable in shaping this work.

A heartfelt thanks goes to Prof. Andrii Tykhonov for his advice and technical assistance. I am also very grateful to David Droz for his helpful feedback, support, and collaboration during various stages of the project—even when he was busy with his own thesis.

I would also like to thank the Geneva/EPFL DAMPE group for their support and camaraderie: Arshia Ruina, Andrii Kotenko, Jennifer Frieden, Mikhail Stolpovskiy, Paul Coppin, Manbing Li, and Hugo Boutin. The many activities and afternoon games we shared brought much-needed balance and enjoyment during these years.

I am deeply thankful to my family—Papa, Mama, Anaïs, and Elio—for their unwavering support and encouragement over the past four years. Their constant curiosity, patience, and belief in me have been a source of strength and motivation from beginning to end.

Last but not least, I would like to thank my friends for always being there and helping me disconnect from the PhD when I needed it. Their joyful energy and presence made a real difference. Special thanks to Laura Dupont, David Nagy, Anthony Joudrier, and Nicolas De Angelis. I am also very grateful to Jeanne and Joslain Bossy, who quite literally saved me by lending me a Swiss-French adapter at a critical moment. Finally, thank you to Claire Lepage for the wonderful hikes we shared over the years—each one represented a chance to leave the city behind and reconnect with nature.

Contents

Introduction **1**

1 Cosmic Rays **3**
1.1 The discovery of Cosmic Rays: . . . 4
1.2 Current Experiments . . . 6
1.2.1 Balloon Experiments . . . 7
1.2.2 Space-Based Experiments . . . 8
1.2.3 Ground-Based experiments . . . 12
1.3 Cosmic Rays . . . 14
1.4 Cosmic-Ray Acceleration Sites and Sources . . . 15
1.4.1 Galactic Sources . . . 17
1.4.2 Extra-Galatic sources . . . 18
1.5 Cosmic-Ray Production and Acceleration Mechanism . . . 18
1.5.1 Fermi's First Order Acceleration: Diffusive Shock Acceleration . . . 22
1.5.2 Beyond Fermi Mechanisms . . . 25
1.6 Cosmic-Ray Propagation . . . 25
1.7 Electrons and Positrons in Cosmic Rays . . . 27
1.7.1 Local Sources at the Origin of High Energies Electron+Positron Flux 29
1.7.2 Dark-Matter . . . 30

2 The Dark Matter Particle Explorer **35**
2.1 DAMPE Sub-detectors . . . 36
2.1.1 Plastic Scintillator Detector . . . 36
2.1.2 Silicon-tungsten Tracker-converter . . . 36
2.1.3 The BGO calorimeter . . . 38
2.1.4 Neutron Detector . . . 39
2.2 Data Acquisition and trigger . . . 40
2.3 Simulations . . . 41

3 Machine Learning Basics **46**
3.1 Machine learning applications 46
3.2 Convolutional Neural Networks 47

4 CNN method for energy reconstruction for electrons **54**
4.1 Non-fiducial Events 54
4.2 Classical Energy Reconstruction Method 55
4.3 CNN model 56
4.4 Results and Model validation 60
4.4.1 MC validation 60
4.4.2 MC-data validation 63

5 Convolutional Method for Identification of Non-fiducial Electrons with the DAMPE Experiment **73**
5.1 A Classical Discriminator for Protons and Electrons 74
5.2 Convolutional Neural Network 75
5.3 Results 78

6 Measurement of the Non-Fiducial Electron+Positron Flux **86**
6.1 Data and Livetime of DAMPE 87
6.1.1 Flight data 87
6.1.2 MonteCarlo Simulations 88
6.2 Event Selection 89
6.3 Electron Identification and Background Estimation 90
6.4 Flux Calculation 91
6.5 Statistical Analysis 97

Conclusion **103**

Nomenclature

AGN Active Galactic Nuclei

HE High Energy

LE Low Energy

ROC Receiver Operating Characteristic

GRB Gamma Ray burst

PWNe Pulsar Wind Nebulae

α Alpha radiation (Helium nuclei)

AMS-02 Alpha Magnetic Spectrometer (version 2)

β Beta radiation (electrons and positrons)

CALET Calorimetric Electron Telescope

CHD Charge Detector

CMB Cosmic Microwave Background

CREAM Cosmic Ray Energetics and Mass

EAS Extensive Air Shower

ECAL Electromagnetic Calorimeter

FERMI-LAT Fermi Large Area Telescope

γ Gamma radiation (photons)

GBM Gamma-ray Burst Monitor

GZK cut off Greisen–Zatsepin–Kuzmin cut off; theoretical upper limit on the energy of cosmic rays due to interactions with the CMB

HERD High Energy cosmic-Radiation Detection

H.E.S.S. High Energy Stereoscopic System

IMC Imaging Calorimeter

ISS International Space Station

λ_{int} Nuclear interaction length (mean free path for nuclear reactions)

LAT Large Area Telescope

LHAASO Large High Altitude Air Shower Observatory

MD Muon Detector

BGO Bismuth Germanium Oxide calorimeter

CE Cross-Entropy Loss

CNN Convolutional Neural Network

CR Cosmic ray

CRE Cosmic ray electrons and positrons

DAQ Data acquisiton system

DL Deep Learning

DM Dark Matter

GD Gradient Descent

HET High Energy Trigger

LR Learning Rate

ML Machine Learning

MLP Multi-Layer Perceptron

MSE Mean Squared Error

$PDPU$ Payload Data Process Unit

PMU Payload Management Unit

PSD Plastic Scintillator Detector

$ReLU$	Rectified Linear Unit
SGD	Stochastic Gradient Descent
SNR	Supernova remnant
STK	Silicon-Tungsten Tracker-Converter
RICH	Ring-Imaging Cherenkov Detector
SAA	South Atlantic Anomaly
Si	Silicon
TASC	Total Absorption Calorimeter
TOF	Time-Of-Flight Detector
TRD	Transition Radiation Detector
VHE	Very High Energy
X_0	Radiation length (mean distance over which a high-energy electron loses all but $1/e$ of its energy)

Introduction

It is impressive to see how peaceful the Milky Way appears when observed on a clear night, far from city lights. However, cosmic rays tell a different story –one filled with violent and energetic events. These high-energy particles are thought to be produced and accelerated in extreme astrophysical phenomena such as supernova remnants and for the highest energies from extra-galactical sources like active galactic nuclei, among others. Their existence raises numerous questions: Where do they originate? How are they accelerated? What can they reveal about our galaxy and even beyond? Since scientists began exploring these questions in the early 20th century following the discovery of cosmic rays by Victor Hess, some answers have being obtained from of their origin and composition, understanding the acceleration mechanisms behind cosmic rays (CRs) and the diffusion effects.

In this thesis, we begin in Chapter 1 with an introduction to cosmic rays, discussing their historical discovery and the scientific progress made in understanding them. We then present three methods for detecting and studying these particles. Next, we explore the question of their origin, with a particular focus on electrons and positrons. We investigate the mechanisms responsible for their acceleration. Specifically, we discuss Fermi's second and first order acceleration processes, which describe how particles gain energy in astrophysical environments. Once accelerated, these particles escape their sources and propagate through space, shaping the observed spectrum – a crucial aspect we will also analyse. Finally, we discuss potential TeV-scale sources of electrons and positrons.

Chapter 2 introduces the Dark Matter Particle Explorer (DAMPE), an experiment launched in 2015. Its primary aim is the measurement of the electron+positron flux up to $\sim$ 10 TeV. Additionally, DAMPE precisely measures protons and heavier ions up to nickel, from 100 GeV to 100 TeV enabling comprehensive studies of cosmic radiation. We describe its four detectors and explain how they reconstruct key properties of particles, such as the absolute charge, direction, and energy.

In Chapter 3, we provide an introduction to Machine Learning (ML), a data analysis technique that has been widely adopted in the physics community. Chapter 4 focuses on

a new topology of events—non-fiducial events—which are a central topic of this thesis. In this chapter, we present an energy reconstruction method based on convolutional neural networks (CNNs) and compare its performance to classical approaches.

Chapter 5 explores the second application of CNNs: enhancing the separation power between electrons and protons beyond what is achievable with traditional classification methods. We develop a CNN-based classifier using shower images from DAMPE's calorimeter to improve particle identification. The developed tools, described in Chapters 4 and 5, are then used to measure the electron+positron flux, which we analyse in Chapter 6. This chapter also includes a detailed statistical analysis.

Finally, we conclude by summarizing the key findings of this thesis and discussing potential future directions.

Chapter 1

Cosmic Rays

The official discovery of Cosmic Rays (CRs) in 1912 by Victor Hess marked the birth of a new field: astroparticle physics. This breakthrough not only opened new avenues for studying the universe but also led to significant discoveries in particle physics. For instance, one of this discoveries is the positron –the antiparticle of the electron– was discovered by studying the track of a particle in a Cloud chamber experiment [1]. That being said, nearly two centuries passed between the initial observation of spontaneous electroscope discharge by Coulomb in 1785, Figure 1.1) shows the gold-leaf electroscope, developed in 1787, and the realisation that this phenomenon was caused by radiation. It was only in 1912 that Hess provided conclusive evidence that the discharge originated from particles arriving from outer space – now known as cosmic rays. The development of

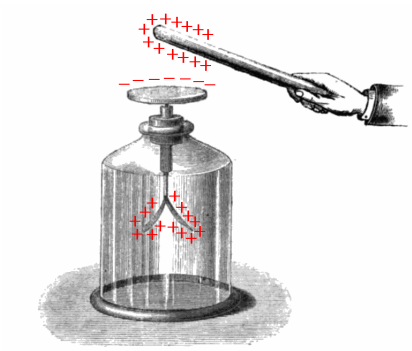


Figure 1.1: The induction charging of a gold-leaf electroscope is an early instrument used to detect the presence of electric charge on a body. The one shown in the figure is composed of two metallic foils at the bottom, connected to a metallic rod with at the top. The foils are encapsulated in a glass bottle to be protected from air friction. By Sylvanus P. Thompson [public domain], via Wikimedia Commons.

new detection technologies have significantly advanced our understanding of cosmic rays,

allowing us to probe their composition and properties in greater detail. In this chapter, we first explore the historical context of CR discovery. We then discuss the current methods used to study them, including ground-, balloon- and space-based experiments, highlighting advantages and limitations of each approach. Next, we examine the overall behaviour of cosmic rays, their acceleration mechanism, and how they are influenced during their propagation through space. Finally, we take a closer look at cosmic-ray electrons and positrons (CREs), discussing potential TeV-scale sources such as supernova remnants (SNRs), pulsars, and dark matter [2–4]

1.1 The discovery of Cosmic Rays:

In 1896, Henri Becquerel discovered radioactivity while studying uranium salts. He concluded that invisible radiation could pass through the paper wrapped around a photographic plate and cause it to react. The uranium salts led to the blackening of the photographic plate, concluding that some invisible radiation could pass through the paper and cause a reaction on the plate as if it was exposed to light. Shortly afterward, Pierre and Marie Curie discovered the Polonium and Radium, along with their radioactive decay. They found that the decay of these elements ionised the air, creating free charges that resulted in the discharge of an electroscope.

At the beginning of the 20th century, Charles Thomson Rees Wilson in the U.K., along with Elster and Geitel in Germany, improved the electroscope sensitivity by enhancing the insulation of the system with a closed vessel. This enable a precise measurement of the spontaneous discharge rate. It was assumed that the main contribution came from radioactive materials in the Earth's crust, and it was expected to decrease with height. Wilson suggested that the atmospheric ionization could be caused by very penetrating radiation with an extraterrestrial origin. However, through different measurement campaigns in tunnels, shielded from the atmosphere, did not support the idea. No reduction in ionization rate was observed. As a consequence, the hypothesis of an extraterrestrial origin was put aside.

By the same time, it was observed that this spontaneous discharged was also present in insulated environments and could even penetrate metal shields. This effect was difficult to attribute to α particles (^{4}He nuclei) or β particles (electrons and positrons). Instead, it was assumed that γ radiation (photons) was responsible for the ionisation, as it was the most penetrating form of radiation known at the time. This led to two main hypotheses for this radiation: it could come from an extraterrestrial source, or from Earth's crust. That being said, it was assumed that the main contribution should come from radioactive material present in Earth's crust, and it was expected to decrease with height.

To better understand the origin of the radiation, various experiments were conducted. Among those, Theodor Wulf built a portable electroscope (Figure 1.2) and began measuring radiation at different locations and heights. If the radiation was coming from Earth's crust, it would imply that radiation levels should decrease with increasing altitude. Wulf took his electroscope on the Eiffel tower (300 m) and observed a small reduction in radiation. However, the decrease was too slight to support the idea that the



Figure 1.2: Electroscope built by Theodor Wulf. By Alessandro De Angelis [public domain], via Wikimedia Commons.

radiation originated from the ground. While the results were inconclusive, Wulf believed that the unusual behaviour was still related to emissions from the Earth.

Despite Wulf's efforts to measure radiation at altitude, his results were inconclusive, leaving the origin of atmospheric radiation unclear. This prompted further investigations. Shortly after, two other scientists conducted separate studies to explore the source of the radiation: Domenico Pacini investigated ionization levels underwater, Albert Gockel ascended to 3000 feet above sea level in a balloon. On one hand, Pacini used the sea water as a shield from possible radiation coming from Earth's crust and observed a 20% decrease concluding this diminishment was due to radiation not coming from Earth. On the other hand, after the studies conducted by Wulf at the Eiffel tower, more data at high altitudes were needed. Hereafter, Gockel realised a flight and obtained no substantial decrease that could be related to an Earth origin for the radiation.

Even with all the results obtained, we had to wait until Victor Hess realised up to 5 flights at different altitudes including one at 5200 m. As shown by Figure 1.3, right above sea level the radiation decreases, around 1-2 km an increase of the radiation is observed, from heights of 4 km and above the ionization increases from 50-100% compared to sea level. Those observations made by Hess led him to conclude that the increase in ionisation had an extraterrestrial origin. A few years after, physicist Werner Kolôrster realised a flight that also confirmed Hess results. Only a few decades later Victor Hess became a Nobel laureate. Cosmic rays is the name given to this extraterrestrial radiation, which was initially thought to consist of gamma rays because of its high penetrating power.

During the second quarter of the 20th century the Dutch physicists Jacob Clay and Tettje Clay-Jolles found evidences of variation of the ionization level as a function of the latitude [6] while travelling between Indonesia and Netherlands. This suggest a dependence on Earth's magnetic field, hence, the observed radiation must have a electric charge. Further evidence came from Bruno Rossi's observation of the East–West asymmetry [7],

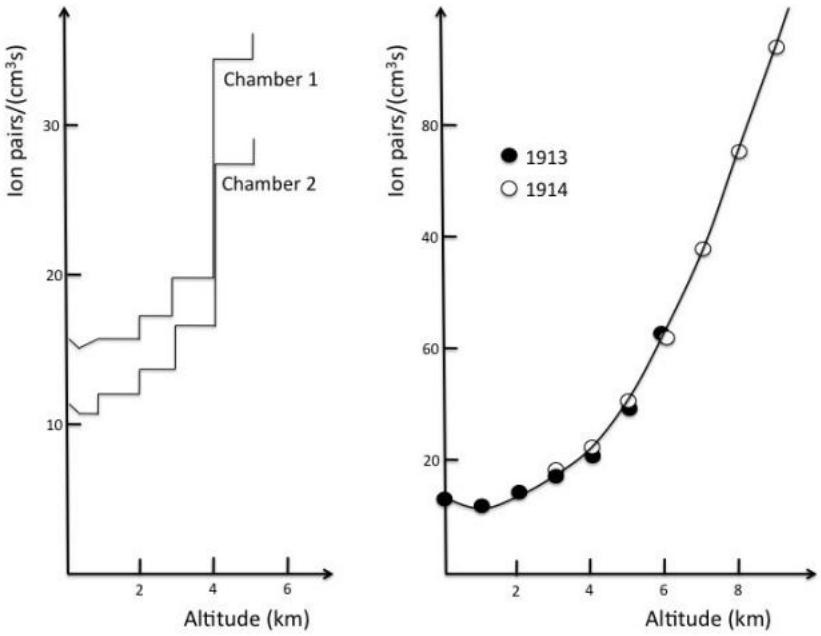


Figure 1.3: Plots representing the number of Ion pair as a function of the altitude. Left plot: Results from Hess ascent in 1912. Right plot: results from Kolhôrster ascent. Image by A. De Angelis [5]

which demonstrated that cosmic rays are predominantly composed of positively charged particles.

1.2 Current Experiments

Since the discovery of CRs, the detection technologies have evolved, leading to a better understanding of such particles. Charles Thomson Rees Wilson conceived a detector of supersaturated vapour of water or alcohol, called *Cloud chamber*. A particle crossing the medium, while ionising the vapour it will cause the vapour to condensate and form a trail where the particle has passed. The creation of such a detector combined with a magnetic field allowed the discovery of the positron [1] (the anti-particle of the electrons) from cosmic rays. Since then detector technology has advanced [8], allowing for more detailed studies about CRs in a way that electroscopes couldn't, such as the identification of the particle and its energy.

The study of CRs can be separated into two categories. The first being minimising or avoiding the presence of the atmosphere, which works as a shield. The second one is to detect indirectly CR by reconstructing the shower created by the interaction of a CR with the atmosphere. Balloon and space experiments allow direct measurement of such particles by sending detectors high in the atmosphere or by sending them into space.

Such experiments are generally classified into two categories: calorimetric experiments, which use a calorimeter to estimate the energy of the traversing particle by measuring the deposited energy; and spectrometer experiments, which use a magnetic field in combination with a tracking system to determine the particle's momentum and charge sign based on the curvature of its trajectory. A caveat of such methods is the limitation in the size of such experiments that limits the observation at the highest energy. As a matter of facts for energies below 100 GeV we expect around one event per metre squared per second, while for PeV energies, we expect around 1 event per metre squared per year and above those energies we reach one event per kilometre squared per year. Thus implying the need for a large acceptance area, impossible to deploy for balloon and space experiments. Therefore, Earth becomes the best candidate, as it was understood that the atmosphere could be used as an active material for the detection of cosmic rays by-products, enabling indirect detection. They are known as ground-base experiments. Such experiments detect the secondary particles generated when a primary cosmic ray interacts with the Earth's atmosphere. Ground-based detectors—such as water Cherenkov detectors, scintillators, fluorescence and Cherenkov telescopes, and muon detectors—are deployed in large arrays to capture the resulting air showers. By analysing the distribution, timing, and type of secondary particles, these experiments can reconstruct the energy and incoming direction of the original cosmic ray. Notable examples include the Pierre Auger Observatory [9], H.E.S.S. [10], HAWC [11], the Cherenkov Telescope Array Observatory (CTAO) [12], and the Telescope Array [13].

1.2.1 Balloon Experiments

Balloon experiments were and are widely used due to the lower cost compared to space experiments. However, the fact that the atmosphere is still present, even if less dense, it can produce some interactions and trigger the beginning of shower. As a consequence, such experiments need to thoroughly understand those effects to be able to accurately measure CRs.

ATIC

The Advanced Thin Ionization Calorimeter (ATIC) [14] was a balloon experiment that investigated primary cosmic rays from 10^{10} to 10^{14} eV. Its payload is composed of a bismuth germanium oxide calorimeter (BGO) with 8 layers. A layer contains 40 BGO crystals of dimensions $2.5\,\text{cm} \times 2.5\,\text{cm} \times 25\,\text{cm}$. Each crystal connected to a Hamamatsu R5611 photomultiplier tube. Resulting in a depth of 17.9 X_0 [14]. On top of the BGO there is a graphite target ($\sim 0.75\lambda_{\text{int}}$ thick), which induces nuclear interactions for incident particles (for 53% of protons and a higher fraction of heavier nuclei) [15]; a matrix of Si pixels [16] is added to reconstruct the particle's charge; scintillator hodoscopes intercalated with the graphite layers to provide faster trigger and enhance track reconstruction. Figure 1.4 shows the structure of the ATIC experiment.

One of its main contribution was the observation of an excess of electrons in the 300-800 GeV energy range [17]. It also measured the fluxes of separated nuclei [18].

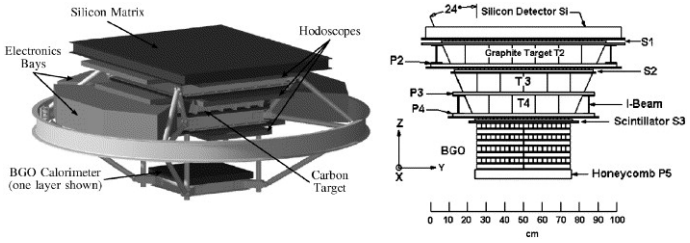


Figure 1.4: Schematic view of ATIC structure [15].

CREAM Experiment

The Cosmic Ray Energetics and Mass (CREAM) balloon experiment determined the composition of cosmic rays up to the 10^{15} eV [19]. The measurement of a particle's charge is done with a timing charge detector, silicon charge detector, and 4 layers of scintillating fiber hodoscopes. The energy is reconstructed by 2 methods, by a transition radiation detector and a tungsten calorimeter (Figure 1.5)

CREAM studied CRs spectra of C,O, Ne, Mg, Si and Fe up to $\sim 10^{15}$ eV [20] as well as proton and helium spectrum [21]

1.2.2 Space-Based Experiments

Space experiments allow for the direct measurement of cosmic rays, eliminating the uncertainties arising from having to model the atmospheric particle showering. Additionally, they allow for significantly longer measurement periods compared to balloon experiments, which ranges from days to weeks.

However, being launched into space imposes several constraints on the design of such experiments. One major limitation is payload mass, which restricts the size of detectors since excessively heavy instruments cannot be launched. Furthermore, the harsh environment during launch, characterized by intense vibrations and shocks, necessitates rigorous qualification. Another challenge is the South Atlantic Anomaly (SAA) [22], a region where Earth's inner Van Allen radiation belt comes unusually close to the surface, exposing instruments to high levels of radiation. On top of that such experiments are also constrained by power limitation and a harsh thermal environment, since heat cannot be evacuated easily.

DAMPE (2015)

The Dark Matter Particle Explorer (DAMPE), is a calorimeter-based experiment, launched in 2015. Since the work of this thesis is based on data obtained from DAMPE, Chapter 2 is dedicated to describe this instrument.

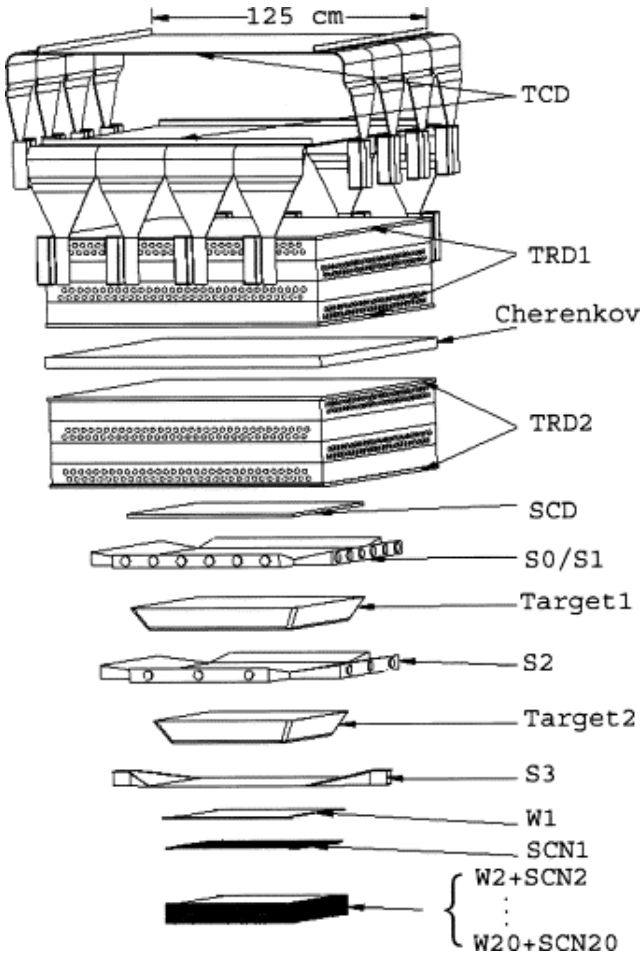


Figure 1.5: Schematic of the CREAM instrument configuration [19].

AMS-02 (2011)

The Alpha Magnetic Spectrometer (AMS-02) [23], is a space experiment located on the International Space Station (ISS) since 2011. As the name suggests, it is a spectrometer. For this a magnet [24] and a tracker are needed, which allows for the estimation of the momentum of the particle and the sign of the charge depending on which direction the particle bends. The experiment is composed of a silicon tracker [25], a transition radiation detector (TRD) [26], a time-of-flight detector (TOF) [27], a ring-imaging Cherenkov detector (RICH) [28] and an electromagnetic calorimeter (ECAL) [29]; see Figure 1.6 for a full view of the experiment.

The presence of the magnet allows AMS-02 to measure the positron flux as well as the positron fraction. Thus confirming and extending previous observations to higher energies [30–33]. The AMS-02 collaboration plans to upgrade the experiment by adding

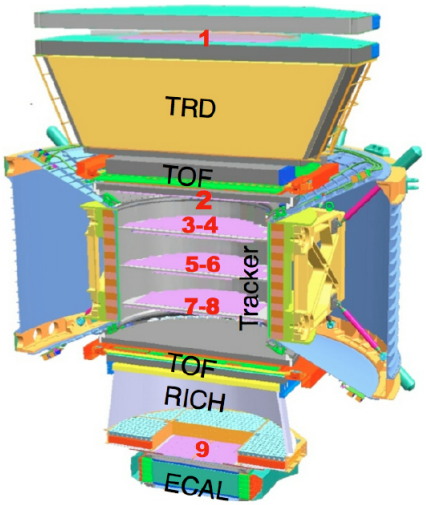


Figure 1.6: Representation of the AMS-02 detector layout [23].

an additional tracker layer, to allow to collect even more data [34].

AMS-02 has measured the positron spectrum up to 1 TeV [32], using three times the statistics of previous analyses [30, 35]. A complex energy dependence is observed, featuring an excess beginning at 25.2 GeV (relative to the lower-energy region), a sharp drop above 284 GeV, and a finite source-term energy cut off at 810 GeV, established with a significance greater than 4σ. These results point to two possible origins for the observed flux: either dark-matter interactions or contributions from astrophysical sources.

AMS-02 has also published a measurement of the electron spectrum from 0.5 GeV to 1.4 TeV [36]. These results exhibit behaviour distinct from that of the positron flux [32]. AMS-02 observes an excess in the electron spectrum starting at 42.1 GeV, relative to the lower-energy trend, although the nature of this excess differs from the positron excess observed above 25.2 GeV. Unlike the positron flux, which features an exponential energy cut off at 810 GeV, the electron flux shows no evidence of a cut off below 1.9 TeV at the 5σ confidence level. This indicates that most high-energy electrons likely originate from different sources than high-energy positrons.

Fermi (2008)

Fermi is a gamma-ray space telescope that studies gamma-rays in the energy range of 10 keV to 300 GeV. While its main purpose is the study of gamma-rays, the clever use of its detectors allows it to be used as a CR observatory by measuring the electron+positron flux [37].

Fermi is composed of the Large Area Telescope (LAT) [38], which is a pair-conversion telescope with a precision converter-tracker and calorimeter, each consisting of a 4×4 of identical towers. Each 40×40 cm^2 tower comprises a tracker, calorimeter and data acquisition module. The tracking detector consists of 18 xy layers of silicon strip

detectors.. It allows photon conversion through its 16 planes at the top of the tracker. These planes are interleaved with tungsten layers to facilitate photon conversion. The calorimeter is used to measure the deposited energy from the electromagnetic shower. The second instrument, the Gamma-ray Burst Monitor (GBM) [39] has a field of view several times larger than the LAT and provides spectral coverage of gamma-ray bursts that extends from the lower limit of the LAT down to 10 keV.

Fermi-LAT observed the CREs flux up to 2 TeV [37]. Another contribution was the study of the positron flux [40] by using Earth's magnetic field to separate positrons from electrons.

CALET (2015)

The Calorimetric Electron Telescope (CALET) [41,42] is also an experiment on board the ISS. The experiment was launched in 2015, it can measure the flux of cosmic-ray electrons+positrons up to 20 TeV, gamma-rays up to 10 TeV and nuclei with Z = 1 to 40 up to 10^3 TeV. CALET consists of a charge detector (CHD), a 3 radiation-length thick imaging calorimeter (IMC), and a 27 radiation-length thick total absorption calorimeter (TASC).

In 2017, CALET published its first measurement of the cosmic-ray electron (CRE) flux from 10.6 GeV up to 3 TeV [43]. The observed spectrum was consistent with a single power law with a spectral index of -3.152 ± 0.016. Possible structures observed above 100 GeV required further investigation, with improved statistics and refined analysis. The results were subsequently extended twice: in 2018, the measurement was extended up to 4.8 TeV [44], and in 2023, up to 7.5 TeV [45]. The spectrum becomes gradually harder below 30 GeV, in agreement with AMS-02 observations. However, between 300 GeV and 600 GeV, it appears significantly softer than the spectra measured by DAMPE [46] and Fermi-LAT [40]. At higher energies, the CALET spectrum exhibits a sharp spectral break around 1 TeV, with the spectral index steepening from -3.15 to -3.91. A broken power-law model fits the data in the 30 GeV–4.8 TeV range significantly better than a single power law, with a significance of 6.9σ, consistent with DAMPE's findings. This spectral break is compatible with expectations from radiative energy losses during propagation from distant sources, except in the highest energy bin. While CALET results are in good agreement with AMS-02 over much of the energy range, discrepancies are observed with DAMPE and Fermi-LAT in the 300–600 GeV region, despite an overall agreement with DAMPE above 1 TeV.

HERD

The High Energy cosmic-Radiation Detection facility (HERD) [47] is a future space-based mission planned for installation on the China space station around 2027. HERD is designed with a large geometric acceptance of 3 m^2sr at 200 GeV for electrons and positrons—ten times greater than DAMPE's acceptance of 0.3 m^2sr above 30 GeV. It is a calorimetric experiment with a total depth of 55 radiation lengths (X_0).

HERD will enable the extension of cosmic-ray nuclei measurements up to PeV energies. Another major objective of the mission is to study electrons and positrons at multi-TeV energies, thereby improving our current understanding of cosmic rays and their origins, and a gamma-ray program.

1.2.3 Ground-Based experiments

As discussed previously, at high energies the cosmic-ray (CR) flux becomes so low that space-based experiments are not expected to detect a significant number of events. Consequently, a complementary approach is required: the use of large-area detectors, on ground.

The atmosphere serves as a natural medium for the generation of particle showers. When a high-energy particle enters the atmosphere and interacts with its constituents, it initiates a cascade of secondary particles, known as an extensive air shower (EAS).

For leptons and photons, this interaction results in an *electromagnetic shower*, predominantly composed of electrons, positrons (from pair production), and photons. In contrast, when the primary particle is a proton or an ion, the resulting shower includes both an electromagnetic and a *hadronic component*. The hadronic component consists of particles such as pions, which further decay and contribute to the overall shower and a muonic component.

For a primary particle with initial energy E_0, the n^{th} hadronic component carries a fraction of the energy given by

$$E_{\mathrm{hadronic}} = \left(\frac{2}{3}\right)^n E_0, \tag{1.1}$$

while the remaining energy is transferred to the n^{th} electromagnetic component:

$$E_{\mathrm{EM}} = \left(1 - \left(\frac{2}{3}\right)^n\right) E_0, \tag{1.2}$$

as described by a simple model [48].

Various detection techniques are available to observe and analyse such showers, as illustrated in Figure 1.7. Furthermore, the morphology of the shower provides insights into the nature of the primary particle. Purely electromagnetic showers—such as those initiated by leptons or photons—tend to be narrower and shorter in longitudinal development compared to hadronic showers, as shown in Figure 1.8.

A first technique is to use scintillation counters, to obtain information about the arrival time and count the number of particles. Such information is correlated to the energy and the incoming direction of the primary particle. Time information allows reconstruction of the direction of the shower. Nonetheless such technique is limited by the inclination of the shower. Therefore, a second detection method is using Cherenkov light. Charged particles crossing a dielectric material with a speed greater than the speed of light in such material generate a cone of light [49]. Such effect can happen in the atmosphere and be detected by telescopes, as it is the case for H.E.S.S. experiment [50].

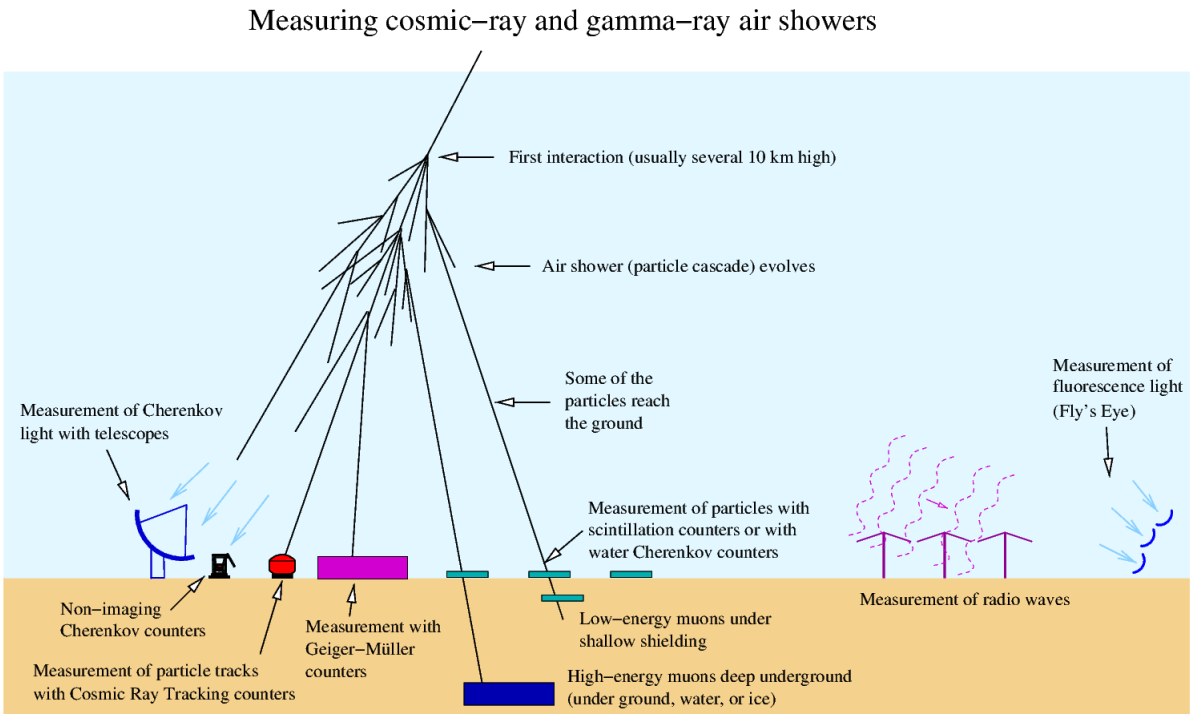
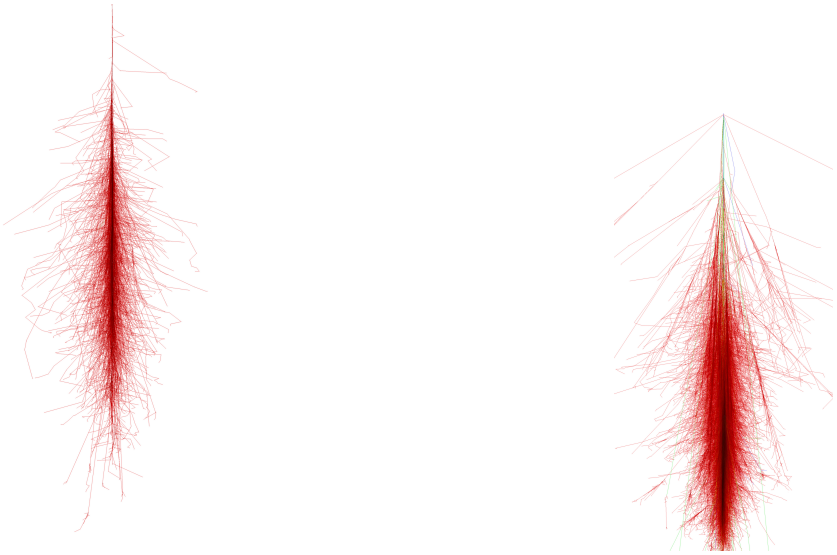


Figure 1.7: Representation of a shower development and different detection techniques. Image Author: Konrad Bernlöhr, from `https://www.mpi-hd.mpg.de/hfm/CosmicRay/ShowerDetection.html`.

However, using the atmosphere as a radiator of Cherenkov light results in small observation time, due to moon night and cloudy days. As a consequence, Cherenkov light can be produced by having water tanks, Pierre Auger observatory [51] is an example of experiments that uses this principle. Some other experiments combine different techniques, for example the LHAASO experiment [52] consists of 1.3 km^2 array of electromagnetic particle detectors and muon detectors (MD), a water Cherenkov detector array with a total active area of 78'000 m^2, 18 wide field-of-view air Cherenkov telescopes and a newly proposed electron-neutron detector array covering 10'000 m^2. The Cherenkov Telescope Array (CTA) [12, 53] follows a similar principle as H.E.S.S., while aiming to improve sensitivity through newer detection techniques. It will consist of two experimental sites: one in the Northern Hemisphere (Canary Islands, Spain) and one in the Southern Hemisphere (Atacama Desert, Chile). It will be composed of telescopes with different sizes. It will be able to measure CRs from 20 GeV up to 200 TeV.

Another major experiment is the Telescope Array (TA) [13], which observes cosmic rays with energies exceeding 10^{18} eV. The surface array spans over 300 square miles of desert and includes more than 500 scintillator detectors, each with an area of 3 m^2, arranged on a 1.2 km grid. Additionally, three fluorescence telescopes stations form a



(a) 50 GeV photon shower.

(b) 1 TeV proton shower.

Figure 1.8: Side view of a shower propagating through the atmosphere. Left figure: shower from a 50 GeV photon. Right plot: 1 TeV proton shower. From `https://www.iap.kit.edu/corsika/`.

30 km triangular array, each housing 12 to 14 telescopes. The TA has been operational in the high desert of Millard County, Utah, USA, since 2007. Pierre Auger Observatory [9] is also a ground-based experiment that uses fluorescence telescopes and water Cherenkov tanks.

1.3 Cosmic Rays

Since the discovery of cosmic rays in the 1910s, various experiments have shed light on their mysteries. Through these experiments, we have been able to understand their composition, about 99 % of the cosmic ray are nuclei, mainly protons (90 %) and helium (9 %). The electrons only represent 1 % of the proton, while positrons represents only 0.1%. Furthermore, cosmic rays have been categorised into two classes: primary cosmic rays, which originate directly from astrophysical sources, and secondary cosmic rays, which are produced through interactions between primary particles and the interstellar medium (ISM) or through particle decay. Experiments typically use energy flux as an observable, defined as the number of particles crossing a given surface and solid angle per unit of time and energy. The flux ϕ follows the equation:

$$\phi(E_1, E_2) = \frac{N(E_1, E_2)}{A \cdot T \cdot \Delta E}, \tag{1.3}$$

where $N(E_1, E_2)$ is the number of events detected by an experiment, E_1 and E_2 define the energy range (usually with $E_2 = E_1 + dE_1$ and $\Delta E = E_2 - E_1 = dE_1$), T is

the experiment's exposure time, and A is the acceptance. Acceptance is defined as the region of an experiment where particles can be detected, and it depends on geometrical factors.

Cosmic-ray fluxes follow a power-law spectrum of the form $E^{-\alpha}$, where α typically ranges between 2 and 3. Figure 1.9 presents measurements of the flux across different experiments and particle species. The energy range of cosmic-ray fluxes extends from a tens of MeV up to about 10^{12} GeV, indicating that particles are accelerated to energies far beyond what current human-made accelerators can achieve. Moreover, there is a hard cut off at 10^{19} eV could be Greizen-Zatsepin-Kuzmin effect (GZK cut off) [54, 55] if protons dominates at the highest energies. At very high energies (VHE), protons can interact with the cosmic microwave background (CMB) photons (γ_{CMB}, resulting in the production of lower-energy proton or neutron, and a pions that subsequently decays into photons or leptons.

$$p + \gamma_{CMB} \rightarrow \Delta^{+} \rightarrow p + \pi^{0} \tag{1.4}$$

$$p + \gamma_{CMB} \rightarrow \Delta^{+} \rightarrow n + \pi^{+} \tag{1.5}$$

Another possible option from sources if heavier nucleis dominate [56].
Moreover, we observe three distinct regions: energies below 5 PeV, where a change in the slope of the flux is observed (commonly referred as the knee); and around 5 EeV, known as the 'ankle'. The features suggest a possible change in the origin of the particles, with cosmic rays below 5 PeV likely originating within our galaxy.

1.4 Cosmic-Ray Acceleration Sites and Sources

Before introducing the mechanism behind the acceleration of CR in section 1.5 we discuss possible acceleration sources.

Sources of cosmic rays must be able to contain a particle that can be accelerated up to a maximum energy, a source should have a size R such that the magnetic field and the energy must be in a relation. The size should be of the order of the Larmor radius. If a particle escapes the acceleration site, it will no longer be able to gain energy. This relation is described by the so-called Hillas relation [58]. In 1984 Hillas developed a diagram that identifies potential sites by examining the relationship between the magnetic field, the size of the accelerator, and the maximum energy that a particle can be accelerated, defined by:

$$\frac{E}{\mathrm{PeV}} \simeq \beta \times Z \times \frac{B}{\mu\mathrm{G}} \times \frac{R}{\mathrm{pc}}, \tag{1.6}$$

where β is the acceleration efficiency of the accelerator and Z is the particle charge. This equation only indicates the maximum attainable energy; however, it is also necessary to have an acceleration mechanism.

Table 1.1 shows examples of maximal energy that can be attained at different sources.

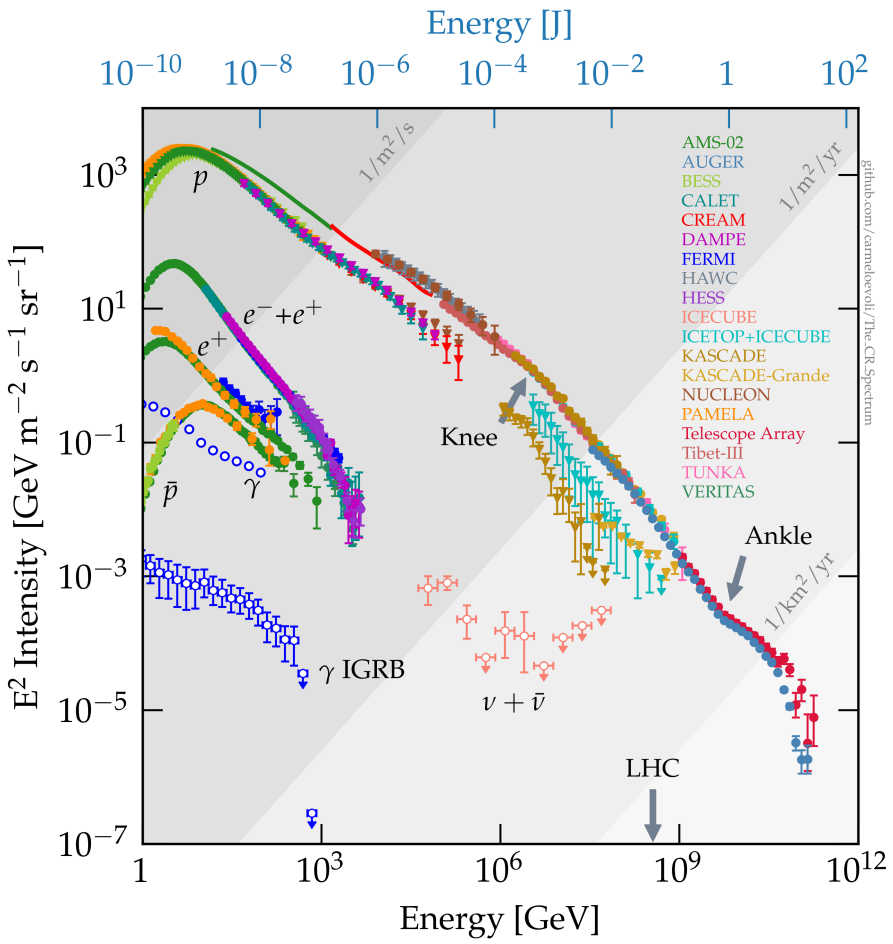


Figure 1.9: Plot of the differential energy flux, scaled by E^2 for different particle species. The different markers present the results of different experiments. The grey arrow at the bottom shows the collision energy of the experiment LHC in the fixed target frame [57].

Source	Magnetic field	Size [pc]	Maximum energy [eV]
SNR	$30\,\mu$G	1	3×10^{16}
Active Galactic Nuclei	$300\,\mu$G	10^4	$> 10^{21}$
Gamma-ray burst	10^9 G	10^{-9}	0.2×10^{21}

Table 1.1: Order of magnitudes of radius and magnetic fields in accelerations sites, and the maximum attainable energy [5].

1.4.1 Galactic Sources

Supernova Remnants and Pulsar Wind Nebulae

The observation of very high-energy (VHE) gamma-ray emission within our galaxy points towards Supernova remnants (SNRs) as sites for particle acceleration [59]. However works like [60] points toward other scenarios to explain the full CR spectrum. SNRs are the remnants of massive stars that have reached the end of their fusion process. Most massive stars can fuse elements up to iron [61, 62] before undergoing a supernova explosion. This process results in the ejection of material from the outer layers of the star, leaving behind a dense compact object at its centre. If the residual core's mass is below approximately 1.4 M$\odot$, a white dwarf is formed. For masses above this threshold and below $2 - 3$ M$\odot$, a neutron star remains, and cores with a mass above 4 $M_\odot$, it will collapse into a black hole. Such supernova explosions are among the most energetic events in the Universe, capable of releasing up to 10^{53} erg in a few seconds, primarily in the form of neutrinos. Moreover, because the ejected matter reaches supersonic velocities, a shock wave is generated. This shock wave interacts with the ISM and serves as a mechanism for accelerating cosmic rays (see Section 1.5 for more details on acceleration mechanisms).
While the outer layers of the star are ejected, its core collapses into a neutron star [63]—a compact object with a radius no larger than a typical city (10-20 km) but with an extreme density of approximately 5×10^{17} kg·m^{-3}. Due to the conservation of angular momentum, when a star collapses, its rotational speed increases significantly, often leading to millisecond-period rotations. Additionally, neutron stars can have extremely strong magnetic fields. To estimate the magnitude of such a field, we consider the conservation of magnetic flux during contraction:

$$BR^2 = B_{\mathrm{NS}} R_{\mathrm{NS}}^2 \Rightarrow B_{\mathrm{NS}} = B \frac{R^2}{R_{\mathrm{NS}}^2}, \tag{1.7}$$

where B and R are the magnetic field and radius of the initial star, B_{NS} and R_{NS} are the magnetic field and radius of a neutron star. A typical star has a magnetic field of approximately $B = 1$ kG, and using the Sun's radius $R \sim 7 \times 10^5$ km, we obtain a neutron star magnetic field on the order of 10^{12} G. The first high-energy pulsars are highly magnetized rotating neutron star emitting beams of electromagnetic radiation out of their magnetic poles– observed were the Crab and Vela pulsars in the 1970s [64]. Remarkably, both were also observed to emit VHE photons. SNRs are classified into three categories: shell-type SNRs, pulsar wind nebulae (PWNe), and a hybrid type that exhibits characteristics of both.

Shell SNR

As discussed in the previous paragraph, the magnetic field at the surface of a neutron star can be extremely strong. However, the magnetic field carried by the ejected matter interacting with the ISM can be significantly weaker, ranging from 10 μG to 1 mG [65].

Despite this weakening, the gradual decrease in magnetic field strength and velocity occurs over a timescale of up to $\sim 10^5$ years, which is still sufficient for cosmic ray acceleration.

Pulsar Wind Nebulae

PWNe are SNRs that contain a pulsar at their core. The evolution of a PWN occurs in three phases. Initially, the pulsar's wind expands supersonically into the surrounding medium composed of the ejected supernova material, lasting for thousands of years. Subsequently, a reverse shock wave is generated due to interactions with the surrounding medium, travelling back toward the pulsar over a timescale of $\sim 10^4$ years. This interaction causes the PWN to oscillate between expansion and contraction. Throughout this process, the PWN loses energy primarily through synchrotron radiation losses. The recent detection of the Geminga pulsar wind nebula (PWN) by HAWC [66].

1.4.2 Extra-Galatic sources

As mentioned in section 1.4.1, SNR can produce CRs up to a few PeV. As shown in Figure 1.9, the observed CRs flux goes beyond this limits, implying a change in the source for CRs. This change arise because a particle can only be accelerated if its contained in an acceleration with a size of the size of the Larmor radius; once it exceeds this size, the particle escapes and no longer gains energy. As a consequence above PeV energies, a galaxy cannot contain such particles. Two possible sites for such high-energy events are active galactic nuclei (AGN) [67] and gamma-ray burst (GRB) [68].

Active Galactic Nuclei

At the centre of some galaxies, it can be observed a very bright and active region with emissions in the full electromagnetic spectrum. They are composed of massive black holes with a mass between 10^5 and up to 10^{10} solar masses. Due to the extreme conditions found in AGN, such systems are expected to be a source of ultra high energy (UHE) CRs. Matter falling into the black hole will conserve angular momentum and will form a rotating accretion disk. A part of the matter will form collimated jets. As shown in Figure 1.10, depending on the angle between the AGN and the observer, different behaviours will be observed, from radio galaxies to blazars.

1.5 Cosmic-Ray Production and Acceleration Mechanism

From the XXth century up to now the development of experimental techniques has allowed to study with more details the CRs, such as its composition, subatomic to stable atomic particles as well as its spectrum. Therefore any model describing the acceleration mechanism of the particles should account for the following conditions [69]:

- The spectrum should follow a power law $dN(E) \propto E^{-\gamma} dE$, where $\gamma \approx 2 - 3$.

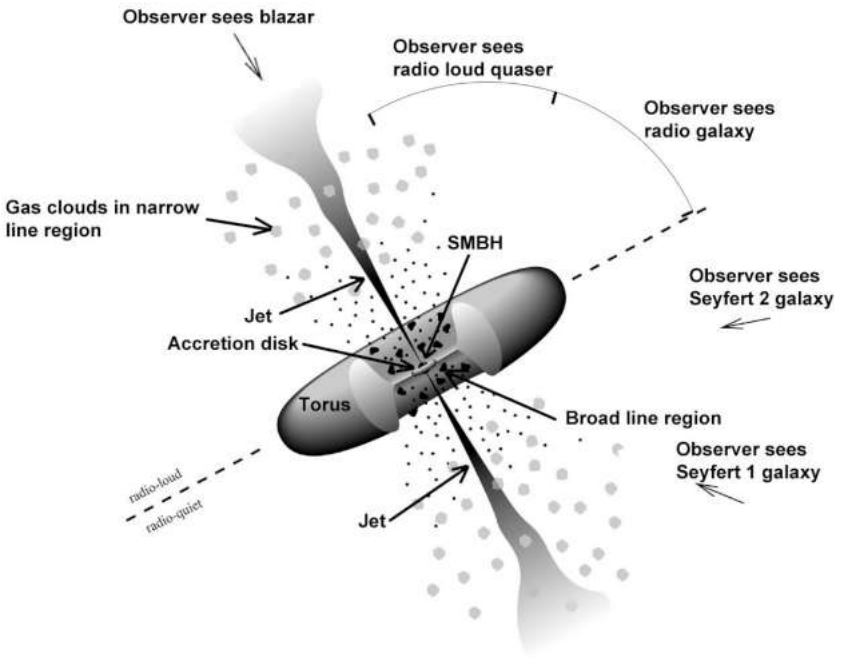


Figure 1.10: Schematic view of an AGN, with the different components. Depending on the angle between the observer and the plane of the AGN, we observe different behaviours. Image from `https://fermi.gsfc.nasa.gov/science/eteu/agn/`.

- It should be able to explain the energy of particles up to 10^{20} eV.
- And in the process the composition of cosmic rays should be close the cosmic abundance of the elements.

Fermi's Second Order Acceleration Mechanism

Fermi's mechanism [70] was developed to describe a stochastic model explaining how particles could be accelerated by colliding with clouds in the interstellar medium. In Fermi's original development, charged particles with mass m are deflected by elastic collisions with a cloud. The cloud has a mass $M >> m$ and a velocity V in a random direction, such that if a particle remains within the region for a time τ_{escape}, a power-law spectrum can be obtained.

Considering a particle with energy E, incoming momentum angle θ_1 and final momentum angle θ_2 –where θ_1 (θ_2) angle is between the initial (final) direction of the particle and the normal to the surface of the cloud. We can analyse the initial energy of the particle in the cloud's reference frame:

$$E_1' = \gamma E_1 - \gamma\vec{\beta}\cdot\vec{p}, \tag{1.8}$$

where $\beta = V/c$, $\vec{p}$ is the momentum vector of the incoming particle. $\vec{\beta}\cdot\vec{p} = \beta p \cdot cos(\theta_1)$. Moreover, we neglected the mass of the particle such that $p = E_1$. Now, E_1' becomes:

$$E_1' = \gamma E_1(1 - \beta cos(\theta_1)) \tag{1.9}$$

Still in the frame of the cloud, since it's mass $M >> m$, the energy of the particle after collision is $E_2' = E_1'$. Now if we go back to the lab frame, the energy of the particle becomes:

$$E_2 = \gamma E_2'(1 + \beta cos(\theta_2)) \tag{1.10}$$

Such that the relative change of energy is:

$$\frac{\Delta E}{E} = \frac{E_2 - E_1}{E_1} = \frac{1 - \beta cos(\theta_1) + \beta cos(\theta_2) - \beta^2 cos(\theta_1)cos(\theta_2)}{1 - \beta^2} - 1 \tag{1.11}$$

When looking at the average change in energy, the final angle θ_2 as no prefer direction ($\langle cos(\theta_2) = 0$), such that the average energy change is:

$$\langle\frac{\Delta E}{E}\rangle = \frac{1 - \beta\langle cos(\theta_1)\rangle}{1 - \beta^2} \tag{1.12}$$

One can show that the probability of a collision depends on the particle speed and the direction of the particle, such that:

$$P \approx 1 - \beta cos(\theta_1) \tag{1.13}$$

Thus the average increase energy if $\langle\frac{\Delta E}{E}\rangle \simeq \frac{4}{3}\beta^2$

As we can see the increase in energy is second-order in β, hence the name Second Order Acceleration Mechanism. It results in a exponential gain of energy since we expect the same fractional increase for each collision. We define the mean free path L along the magnetic field, such that the time between collisions is $\frac{L}{c \times cos(\phi)}$, averaging over $cos(\phi)$ over the pitch angle ϕ, such that the average time between collisions is $2L/c$. Therefore the rate of energy gain is:

$$\frac{dE}{dt} = \frac{2}{3} \cdot \frac{c\beta^2}{L} E = \alpha E \tag{1.14}$$

Now by plugging the rate of energy gain into the diffusion-loss equation (Eq 1.41):

$$\frac{dN}{dt}(E) = \frac{d}{dE}[b(E)N(E)] + Q(E,t) + D\nabla^2 N(E), \tag{1.15}$$

where $N(E)$ is the number density of a given species and is a function of energy, $b(E)$ is the energy loss term, $Q(E,t)$ is the source term (see 1.6 for more details), D is the diffusion coefficient.
We will look at the case of steady-state solution with no diffusion so $dN/dt = 0$ and $D\nabla^2 N = 0$. We will also assume that there is no source. The energy loss term $b(E) = -dE/dt = \alpha E$. Therefore Eq 1.15 becomes:

$$-\frac{d[\alpha E N(E)]}{dE} - \frac{N(E)}{\tau_{\text{escape}}} = 0, \tag{1.16}$$

where τ_{escape} is the time it takes for a particle to escape the source. We rearrange the equation:

$$\frac{dN(E)}{dE} = -(1 + \frac{1}{\alpha \tau_{\text{escape}}}) \frac{N(E)}{E} \tag{1.17}$$

By solving the differential equation, we get $N(E) = A \times E^{-\gamma}$, where A is a normalisation constant and $\gamma = 1 + \frac{1}{\alpha \tau_{\text{escape}}}$. The measured value of γ for the electron+positron is around 3, as observed by different experiments like CALET [41] ($\gamma = 3.15 \pm 0.01$ with a change of spectral index of $\Delta\gamma = -0.77 \pm 0.22$ at 1 TeV), DAMPE [46] (broken power-law fit yields $\gamma_1 = 3.09 \pm 0.01$ and $\gamma_2 = 3.92 \pm 0.20$ with a break at 914 GeV), AMS-02 [23] ($\gamma = 3.28^{+0.014}_{-0.016}$ and a change of spectral index of $\Delta\gamma = 0.094 \pm 0.014$ at 42 GeV).

This final equation describes the power-law spectrum predicted by Fermi's mechanism. Nonetheless, this description encounters a few problems:

- The interstellar cloud speed is well below the speed of light $\frac{V}{c} \leqslant 10^{-4}$. Additionally, the mean free path for the scattering of particles with the interstellar medium is of the order of 0.1 pc [69], resulting in only a few collisions per year– insufficient to explain the observed flux.

- The injection problem: since electrons lose energy through ionisation losses, the acceleration mechanism must either be fast enough to overcome the losses or inject electrons with energy greater than that corresponding to the maximum of energy loss rate.

- The theory is unable to explain why the value of γ should be $\sim 2-3$

1.5.1 Fermi's First Order Acceleration: Diffusive Shock Acceleration

The acceleration mechanism described in Section 1.5 can be improved by taking into account shock waves, this mechanism is also known as Fermi's First order Acceleration Mechanism [71]. This acceleration mechanism has been of interest for modern CRs physics. A key feature of such theory is the first order dependence of the shock spread contrary to the previous described model. We define two constants β and P: $E = E_0\beta$ is the average energy of the particle after one collision and P is the probability that the particle remains within the acceleration after one collision. After k collisions, there are $N = N_0P^k$ particles with energies $E = E_0\beta^k$. By eliminating k from both quantities, we get:

$$\frac{ln(N/N_0)}{ln(E/E_0)} = \frac{ln(P)}{ln(\beta)}, \tag{1.18}$$

such that:

$$\frac{N}{N_0} = (\frac{E}{E_0})^{\frac{ln(P)}{ln(\beta)}}, \tag{1.19}$$

where N is the number of particles with energy greater or equal to E. By rewriting Eq 1.19, we obtain:

$$N(E)dE = \text{constant} \cdot E^{-1+\frac{ln(P)}{ln(\beta)}}dE \tag{1.20}$$

Moreover we recover a power law energy distribution by the derivation of the Fermi's second order acceleration mechanism. From there, we can see that β is related to α (the gain of energy) and P to τ_{escape} (the containment time).

Let's now go through a more systematic calculation of this acceleration mechanism. For example let's take as a shock wave the material ejected in supernova explosions which can attain velocities of order 10^4 km s^{-1} [69]. Therefore a strong shock wave travels at supersonic velocity $U \gg c_{\text{sound}}$, where c_{sound} is the sound celerity in the ambient medium. We define the rest frame of the shock wave such that the upstream gas flows into the shock front with speed $v_1 = U$ and leaves the shock with a downstream speed v_2 (Figures 1.11 a and b.). The conservation of the mass through the shock is conserved by the equation of continuity, so:

$$\rho_1 v_1 = \rho_1 U = \rho_2 v_2 \tag{1.21}$$

In the case on strong shock, the pressure ratio between the upstream and downstream is $\frac{\rho_2}{\rho_1} = \frac{\gamma+1}{\gamma-1}$, where γ is the ratio of specific heat capacities of the gas. For example, a monoatomic or fully ionised gas $\gamma = 5/3$, $\frac{\rho_2}{\rho_1} = 4$ and $v_2 = (1/4)v_1$. Resulting in the downstream speed being 4 times slower than the upstream one.

If we consider a particle ahead of the shock, the multiple scatterings within the medium ensure the particle's distribution is isotropic. Hence, the particle sees the shock front moving at a speed $\Delta U = U - \frac{1}{4}U = \frac{3}{4}U$ relative to the upstream as shown by Figure 1.11.c. If now we look at particles moving from downstream to the shock wave from the reference frame of the downstream (Figure 1.11.d), the relative velocity is again $\frac{3}{4}U$.

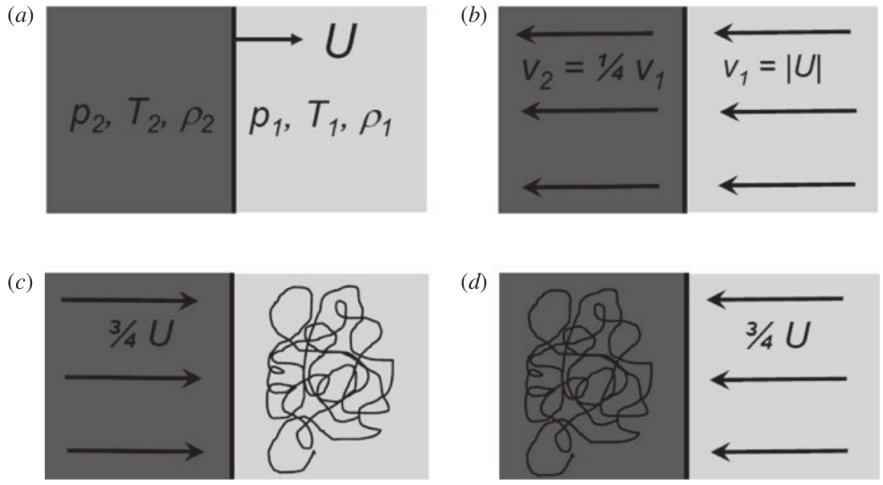


Figure 1.11: Interaction of a particle with a strong shock wave. (a) Shock waves propagating at speed U, separating the medium into two regions with different pressure and temperature conditions. (b) Flow of interstellar gas surrounding the shock front. (c) Flow of gas in the frame in which the upstream is stationary and the speed distribution is isotropic. (d) Flow of gas in the frame in which the downstream is stationary and the speed distribution is isotropic. Image from [69].

Thus resulting in another head-on collisions. The overall result is that, independently of the incoming region of the particle, it will undergo the same process, leading to a small gain of energy in both directions with no losses. Contrary to Fermi second order mechanism, this time the collisions are all head-on and the particle gains energy.
The description showed above can be used to quantitatively determine the parameters introduced earlier, β and P, for a complete acceleration cycle. Let's first look at the energy gain for a particle travelling from the upstream to the downstream. For such a particle, the downstream approaches with a speed $V = \frac{3}{4}U$, we perform a Lorentz transformation such that the particle's energy when it passes into the downstream space is:

$$E' = \gamma(E + p_x V), \tag{1.22}$$

where p_x is the momentum in the direction normal to the shock wave. We suppose a non-relativistic shock wave, $V \ll c$. Nonetheless, the particle is relativistic such that $E = pc$ and $p_x = \frac{E}{c}cos(\theta)$. The change and relative change of energy becomes:

$$\Delta E = pVcos(\theta) \text{ and } \frac{\Delta E}{E} = \frac{V}{c}cos(\theta) \tag{1.23}$$

Following the same idea as previously described, the probability for a particle to arrive to the shock wave in the angle range $[\theta, \theta + d\theta]$ is proportional to $sin(\theta)d\theta$. The rate of approach to the shock front is proportional to the x component of the velocity,

$c\cos(\theta)$. Therefore the probability for a particle to cross the shock wave is proportional to $sin(\theta)cos(\theta)d\theta$. As a consequence, the probability becomes:

$$p(\theta) = 2sin(\theta)cos(\theta)d\theta, \tag{1.24}$$

where θ is in the range $[0, \pi/2]$.

Taking this into account, the average gain in energy when crossing the shock wave is:

$$\left\langle \frac{\Delta E}{E} \right\rangle = \frac{V}{c} \int_0^{\pi/2} 2cos^2(\theta)sin(\theta)d\theta = \frac{2}{3}\frac{V}{c} \tag{1.25}$$

When the particle goes from the upstream to the downstream then goes back through the shock front, the particles energy increases by $\frac{2}{3}\frac{V}{c}$. Therefore, the resulting energy gain for one trip through the shock front and back is :

$$\left\langle \frac{\Delta E}{E} \right\rangle = \frac{4}{3}\frac{V}{c} \tag{1.26}$$

Consequently,

$$\beta = \frac{E}{E_0} = 1 + \frac{4V}{3c} \tag{1.27}$$

We can now calculate P the probability that the particle remains within the acceleration after one collision. We will follow the work made by [72], the number of particles crossing the shock from either direction is $\frac{1}{4}Nc$, with N the number density of particles. However, downstream particles leaves the acceleration site progressively at a rate $NV = \frac{1}{4}NU$. Therefore the decrease of particles per time is $\frac{1}{4}NU/\frac{1}{4}Nc = U/c$. As previously assumed, the shock is non-relativistic thus resulting in $P = 1 - (U/c)$:

$$ln(P) = ln(1 - \frac{U}{c}) \simeq -\frac{U}{c} \text{ and } ln(\beta) = ln(1 + \frac{4V}{3c}) \simeq \frac{4V}{3c} = \frac{U}{c} \tag{1.28}$$

Therefore

$$\frac{ln(P)}{ln(\beta)} = -1 \tag{1.29}$$

Plugging this results into 1.20, we get $N(E) \propto E^{-2}$. This result explains how to get a power ≈ -2 for the fluxes at the source. However, the observed power law by experiment is different from -2. That being said, the observed fluxes differed from prediction due to the propagation of CRs, and, as it will be discussed in Section 1.6, electrons and positrons suffer from radiative energy losses. The conditions require strong shock waves, that can be produced by supernova remnants or active galactic nuclei, for example.

1.5.2 Beyond Fermi Mechanisms

The aforementioned mechanisms describing particle acceleration capture the essential physics involved in the process. That being said, these models often neglect important details—such as the fact that transitions between regions can occur on timescales shorter than the particle's mean free path. Additionally, the presence of magnetic fields can alter the plasma viscosity, thereby affecting energy transport. On top of that, relativistic effects are also not taken into account. One of this works is the non-linear diffusive shock acceleration [73]. Moreover, Fermi's acceleration mechanisms don't take into account the influence that the cosmic rays can have on the environment [74].

1.6 Cosmic-Ray Propagation

As discussed in Section 1.5, sources of CRs accelerate particles following a power law spectrum that being said the power law at the spectrum differs from the one observed by experiments. Since the work of this thesis focuses on electron+positron flux, we will focus on such particles. Nonetheless, during their journey through interstellar space CREs will lose energy, mainly through synchrotron radiation loss and inverse Compton scattering [75]. Bremsstrahlung losses are small for CREs with energies above 100 MeV [76]. Thus resulting in spectrum different at the source. As mentioned before, electrons lose energy mainly through synchrotron radiation and inverse Compton scattering with the CMB photons, resulting in a loss of :

$$\frac{dE}{dt} = -\frac{32\pi c}{9}\left(\frac{e^2}{mc^2}\right)^2 \frac{c}{(mc)^2}\left(\omega_{ph} + \frac{H^2}{8\pi}\right)E^2, \tag{1.30}$$

where m is the electron mass, e is the electron charge, c is the speed of light in the vacuum, ω_{ph} is the photon energy density, H is the magnetic field strength and E is the energy of the particle. Eq 1.30 can be rewritten as

$$\frac{dE}{dt} = b(E) = -\beta E^2 \tag{1.31}$$

where $b(E)$ is the average rate at which electrons lose energy.

From Equations 1.30 and 1.31 it can be observed that the losses of relativistic electrons increases with increasing energy. By integrating Eq 1.31, we can find that the energy of a particle has the following time dependence:

$$E(t) = \frac{E_0}{1 + \beta E_0 t}, \tag{1.32}$$

where E_0 is the initial energy of the particle, β is define in Eq 1.31. Therefore, the time it takes a particle to loose half its energy is $\tau(E_0) = 1/(\beta E_0)$. As a consequence, the average distance crossed by a particle is $\lambda \approx \sqrt{2D\tau}$, where D is the diffusion coefficient. While if we take into account the energy dependence for the diffusion coefficient we

get [77]:

$$\lambda(E, E_0) = \left(\int_0^{\tau(E)} D(E')dE' \right)^{1/2} = \left(\int_{E_0}^{E} \frac{D(E')}{b(E')} dE' \right)^{1/2} \tag{1.33}$$

By taking the following conditions for the galaxy: $D \approx 10^{29}$ cm^{2}s^{-1}; $\omega_{\text{ph}} \approx 1$ eV/cm^3 and $H \approx 4 \cdot 10^{-10}$ T [77]. It results that a 1 GeV electron has an average path length $\lambda \approx 3.7$ kpc. Therefore, an electron produced at distance greater than the size of the galaxy ($R_{\text{galaxy}} \approx 10$ kpc) cannot reach the Earth. This implies that electrons with for energies greater than 1 GeV and for which the path length is smaller than the radius of the galaxy, the spectrum at injection will differ considerably from the one observed with our detectors even in space.

As we can see here, we supposed the diffusion coefficient to be independent of the energy. Nonetheless, already a simple supposition shows already the dependence of the observed flux on the scale of propagation for CREs, on how the sources are distributed.

Let's derive the diffusion-loss equation (Eq 1.15), already introduced in Section 1.5. To do so, we consider particles injected, by a given source, at a rate $Q(E, t)$ and an elementary volume dV. Electrons going through that volume are subject to energy gains and losses, defined by:

$$-\frac{dE}{dt} = b(E), \tag{1.34}$$

where $b(E)$ is the average rate at which electrons gain/lose energy.
At an instant t, there are $N(E)\Delta E$ particles with an energy in the range [E,$E + \Delta E$]. After a Δt has passed, these particles are replaced by particles that had energies between E' and $E' + \Delta E'$ at time t, where:

$$E' = E + b(E)\Delta t \text{ and } E' + \Delta E' = E + \Delta E + b(E + \Delta E)\Delta t \tag{1.35}$$

We can perform a Taylor expansion for small values of ΔE and replace E' (Eq 1.35) with its definition:

$$E' + \Delta E' = E + \Delta E + [b(E)\Delta t + \frac{db(E)}{dE}\Delta E \Delta t] \tag{1.36}$$

Hence:

$$\Delta E' = \Delta E + \frac{db(E)}{dE}\Delta E \Delta t \tag{1.37}$$

We also know that the number of events that gain or lose energy is:

$$\Delta N \Delta E = N\left(E', t\right) \Delta E' - N(E, t)\Delta E \tag{1.38}$$

In Eq 1.38, we replace $\Delta E'$ (Eq 1.37) and E' (Eq 1.35) by their definition and we do a Taylor expansion for small $b(E)\Delta t$.

Therefore, we get $N(E',t) = N(E,t) + \frac{\partial N(E,t)}{\partial E} b(E)\Delta t$. Then Eq 1.38 becomes:

$$\Delta N \Delta E = [N(E,t)\frac{\partial b(E)}{E} + \frac{\partial N(E,t)}{\partial E} b(E)]\Delta E \Delta t \tag{1.39}$$

Finally we get:

$$\frac{dN}{dt} = \frac{d}{dE}[b(E)N]. \tag{1.40}$$

We get the evolution of the particle spectrum in a given volume subject to only gains and losses. Moreover, particles can be injected by sources (see Section 1.7.1 for more details) at a rate Q(E,t) per unit of volume. On top of that we have to take into account particles entering and leaving volume dV by diffusion, which depends upon the gradient of he density. Including those effects, Eq 1.40 becomes:

$$\frac{dN}{dt} = \frac{d}{dE}[b(E)N] + Q(E,t) + D\nabla^2 N(E) \tag{1.41}$$

Eq 1.41 is the so called diffusion-loss equation that allows to describe the energy particle spectrum.

More generally other terms can be added to the equation, for example by including terms for radioactive decay, spallation and so on. For example for ions (i) the diffusion-loss equation becomes [69]:

$$\frac{\partial N_i}{\partial t} = D\nabla^2 N_i + \frac{\partial}{\partial E}[b(E)N_i] + Q_i - (n\beta_i c\sigma_i^{\text{spall}} + \frac{1}{\tau_i \gamma_i})N_i + \sum_{j>i}(n\beta_j c\sigma_{ji}^{\text{spall}} + \frac{1}{\gamma_i \tau_i^{\text{decay}}})N_j, \tag{1.42}$$

where,

- First term correspond to the diffusion and convection effects.
- The second term describes the changes of the spectrum due to energy losses. With $b(E)$ represents the gains and losses of the particle.
- The last two terms take into account the effects from spallation and decay of heavier elements.

1.7 Electrons and Positrons in Cosmic Rays

The electron and positron cosmic rays (CREs), from this point forward, we will refer to electrons and positrons simply as electrons, constitute less than 1% and 0.1% of proton cosmic rays. Respectively to their relatively low mass of 511 keV [78], electrons experience significant energy losses from synchrotron radiation in the presence of magnetic fields, as well as from inverse Compton scattering. These energy loss mechanisms make electrons and positrons exceedingly rare at multi-TeV energies. This limits the distances

at which sources can be found.

In recent decades, significant progress has been made in understanding the electron spectrum. The measured electron flux shows distinctly different magnitude and energy dependence compared to that of the positron flux. Figure 1.12 shows AMS-02 measurement of each individual spectrums up to 1.4 TeV. The comparison of the spectral index for both species by AMS-02 [36] shows also a different behaviour (Figure 1.13, thus pointing toward different sources for electrons and positrons.

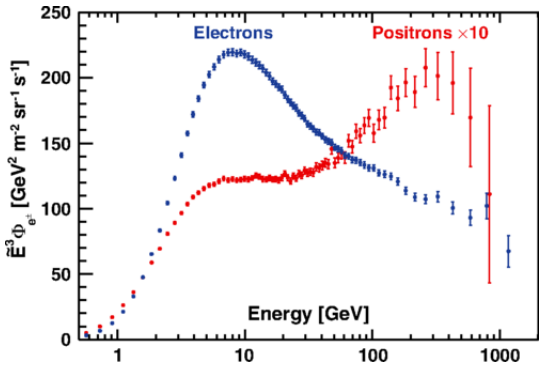


Figure 1.12: The AMS electron (blue data points) and positron (red data points, multiplied by 10) spectra . For display purposes the electron data point at $\sim$ 830 GeV is slightly shifted horizontally to avoid overlap with the positron point [36].

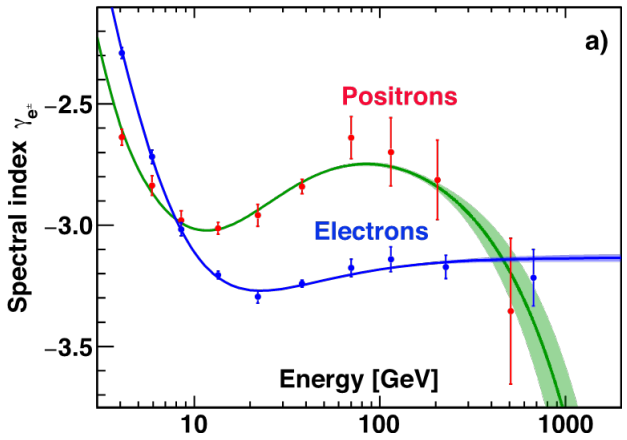


Figure 1.13: The spectral indices of the electron flux γ_{e^-} and of the positron flux γ_{e^+} as a function of energy. The green band represents the 68 % confidence level interval of a fit to the positron flux. The blue band represents the 68 % confidence level interval of a fit to the electron flux. [36].

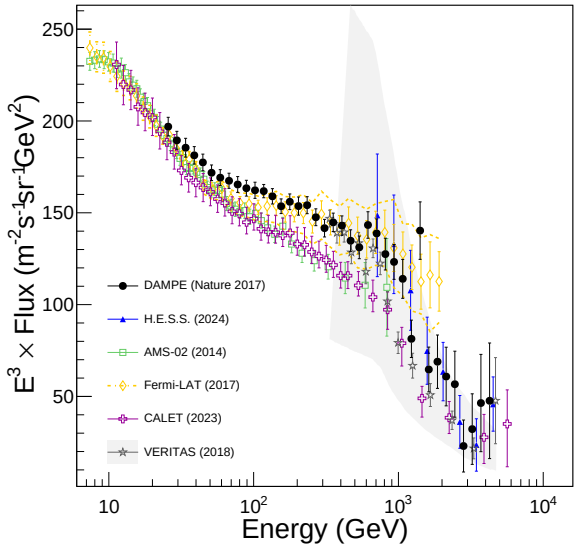


Figure 1.14: electron+positron flux measured by DAMPE and comparison with other experiments [46]. H.E.S.S [80], AMS-02 [81], Fermi-LAT [82]

On top of the measurement of individual spectra, different experiments have measured the electrons+positron flux as shown in Figure 1.14. Observations by AMS-02 [79] have revealed a slight change in slope in the CREs spectrum around 40 GeV, while DAMPE [46] and CALET [45] have detected a spectral break at TeV energies through direct measurements. Indirect measurements from H.E.S.S. [80] further support these findings. Additionally, AMS-02 [32], as shown in Figure 1.15, has reported an increase in the positron spectrum from $\sim$ 40 GeV to $\sim$ 300 GeV, previously observed by Fermi-LAT up to $\sim$ 200 GeV [40]. In addition to a sharp drop above 300 GeV compatible with an energy cutoff at the TeV [23].

1.7.1 Local Sources at the Origin of High Energies Electron+Positron Flux

In section 1.4, we discussed potential sources for cosmic-ray acceleration to very high energy (VHE). SNRs are well known to be a major source for VHE electrons. Works from [88–90] already showed the possible contribution from SNRs in the electron spectra at TeV energies. Today, the main question is the likelihood of this process occurring. Hence, the question is: what are the characteristics of such sources (distance, age and critical energy)? Observations by major experiments suggest that possible contributions from local sources [2–4] could contribute to the observed spectrum at around 10 TeV, with some earlier contributions at 40 GeV, as shown in Figure 1.16, where Monogem contributes in the range of 100 to 1000 GeV. Work from [91] explain that PWNe contribution could

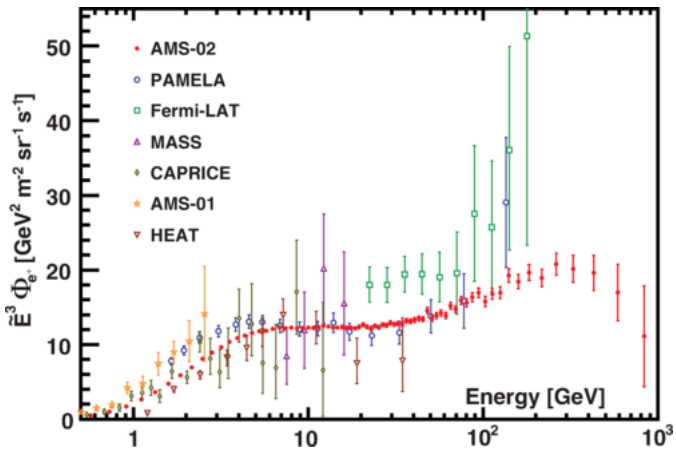


Figure 1.15: Positron flux measured by AMS-02 [32] together with earlier measurements from PAMELA [83], Fermi-LAT [40], MASS [84], CAPRICE [85], AMS-01 [86], and HEAT [87].

explain the 40 GeV break. On the other hand, the main contribution above 1 TeV seems to originate from Vela SNR. As it is pointed out by [2], there is some conditions that need to be taken into account. First, the greater the distance from the source, the more diffuse the electron will be. hence the observation. A particular source will fade due to the losses already mentioned. As a consequence a source should be no further than 2-3 kpc [2–4]. The second point is the age of the SNR [2] showed that the age of SNR should not be greater than 100 kyr. Studies [3, 4] suggest also that Vela SNR could be a potential source of electrons.

1.7.2 Dark-Matter

Fritz Zwicky studied the redshift of various galaxies in the Coma Cluster and noticed a significant scatter in the apparent velocities of eight galaxies within this cluster [92]. Similar observations were made by Hubble and Humason [93]. Zwicky went further by estimating a velocity dispersion of 80 km/s, while the observed value was approximately 10^3 km/s. This discrepancy led Zwicky to conclude the need for additional unseen mass, which he termed Dark Matter (DM).

In the 1970s, while studying the rotation curves of various galaxies, Vera Rubin observed [94] deviations from model predictions (Figure 1.17). To account for the observed rotation curves, additional mass was required. However, this missing mass was not detected through spectral studies, in fact it was conclude that it did not emit any electromagnetic signal nor have an hadronic origin. Consequently, the existence of a new type of matter was proposed—one that interacts gravitationally and possibly weakly with standard matter [95].

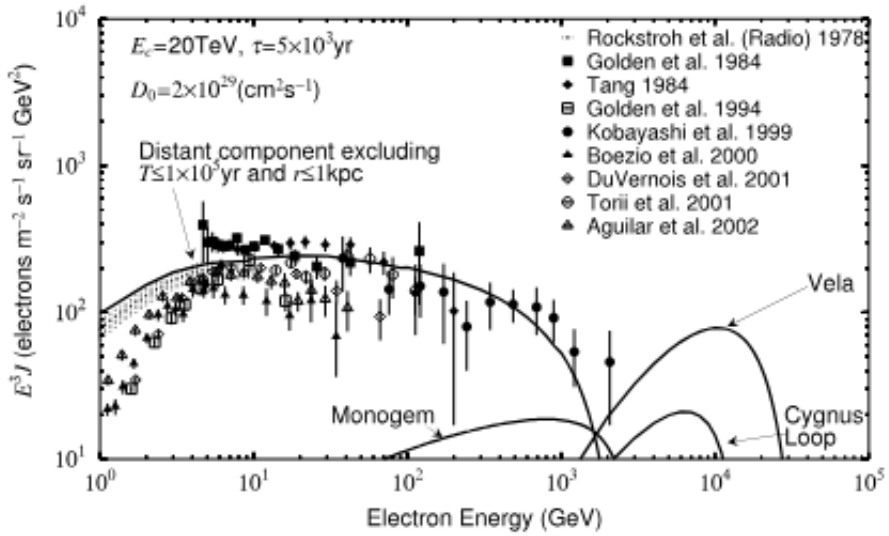


Figure 1.16: Example of expected spectra of electrons with an energy cut off of E_c = 20 TeV, escape time after SNR formation $\tau = 5 \times 10^3$ yr [2].

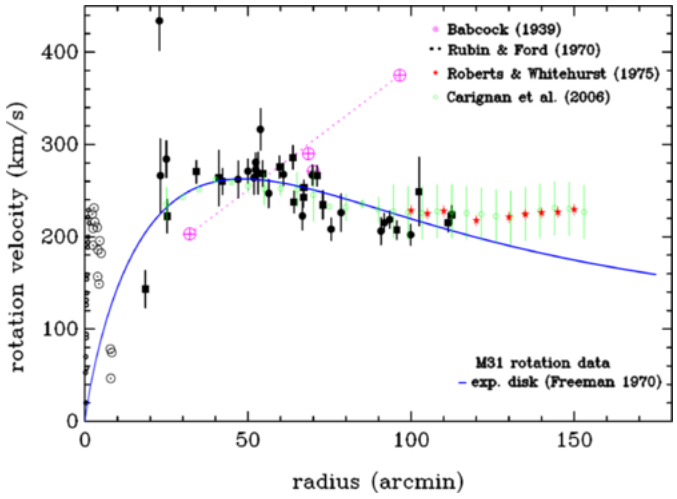


Figure 1.17: Rotation curve data for M31 [95].

Dark Matter Search

Current evidence for DM comes from gravitational observations that lead to its detection (rotations curves of galaxies, gravitational lensing). A gravitational lens is matter, such as a cluster of galaxies or DM, that bends light from a distant source as it travels toward an observer. By observing where the observed source is located, information

about the mass and distribution of the system that causes the bending can be obtained. If the source of such bending cannot be observed thanks to the electromagnetic waves it implies the existence of DM. Nonetheless, it is hypothesised that DM could be observed from a particle standpoint, since we can expect to have a weak coupling to it [96]. That being said, there exists other hypotheses to replace DM, one of them if primordial black holes, for instance see Chapter 27 from [97]. However, DM can span many orders of magnitude in terms of masses, from sub eV [98] to a few solar masses [99]. The search for DM can go into three different perspectives as shown in Figure 1.18.

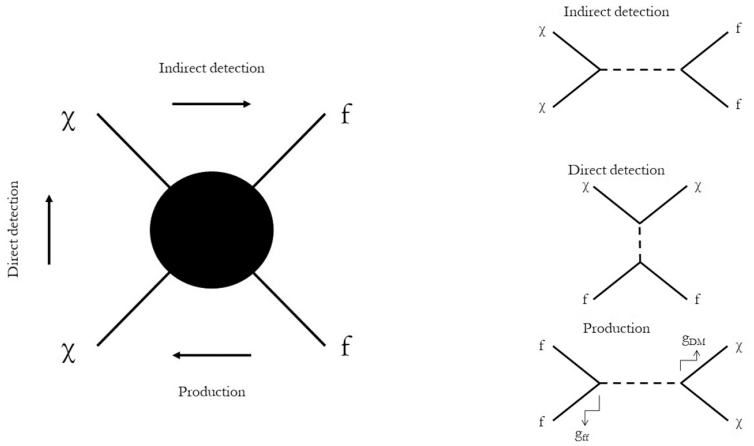


Figure 1.18: Overview of three types of searches for dark matter particles (represented by χ), all based on interactions with standard model particles (represented by f). The three Feynman diagrams on the right are specifications of the diagram on the left, where the arrows on the left diagram indicate the direction in which the diagram should be read. On the right, the dashed lines represent the mediator, and g_{ff} and g_{DM} are the coupling strengths at the standard model and dark matter vertex, respectively [100].

Production

A way to search for dark matter (DM) is by producing it in Standard Model particle collisions. Colliders on Earth, such as the LHC, can be used to collide protons, potentially generating DM particles in the process. However, such searches are limited by the energy accessible to the collider, meaning that only DM particles with masses in the MeV to TeV range can be probed. By placing a detector at the collision point, one can analyse missing transverse energy as a potential signature of DM production. More detailed description of the current status of DM searches at colliders can be found in [101].

Direct research

Otherwise, since dark matter is present in the galaxy and that Earth is travelling through it, we can build an experiment with very sensitive detectors and very low back-

ground, such that we can expect to detect the collision of dark matter with atoms composing the detector. An example is the XENONnT experiment [102], which uses a double phase of liquid/gas using Xenon a noble gas, and its precursor XENON1T experiment reported the observation of an excess [103].

While the search continues, up to now there is no confirmed evidence for the detection of DM and only constrains have been set on. On possible dark matter candidate are weakly interacting massive particles (WIMP), figure 1.19 shows the constrains on the WIMP-nucleon cross section as a function of the WIMP mass.

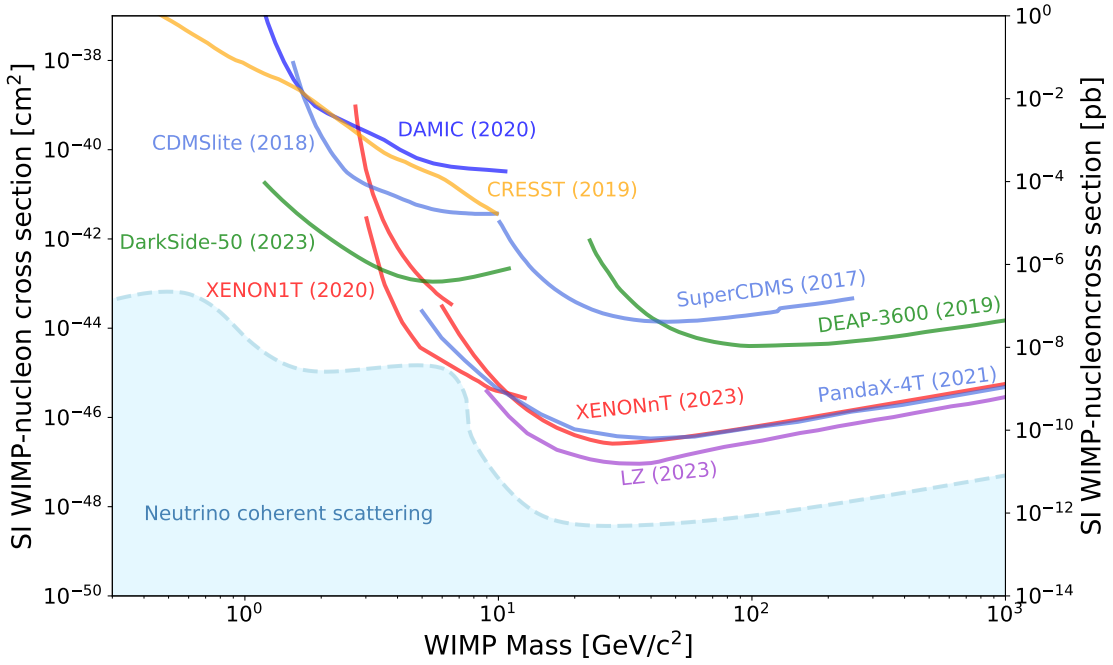


Figure 1.19: Upper limits on the self interacting DM-nucleon cross section as a function of DM mass. Plot from Chapter 27 of [97].

Indirect research

Through indirect searches, dark matter particles can be detected via the observation of cosmic-ray spectra. Models predict that DM can decay or annihilate into Standard Model (SM) particles, which, in turn, may be detected by cosmic ray experiments.

As previously mentioned, one mass range for DM candidates lies in the MeV-TeV range (WIMPs). Consequently, if CRs are produced by such particles, they should also exhibit energies within this range. That being said, many detection channels exist, depending on the underlying DM models.

One approach is the gamma-ray line search, in which experiments look for an approximately monochromatic spectral feature in gamma-ray observations. Since gamma

rays do not diffuse through the galaxy, experiments can trace them back to their sources. Observatories such as Fermi-LAT [104] and DAMPE [105] are involved in such searches.

Another possible channel are charged CRs more precisely anti-matter particles like positron or anti-protons [106], with the caveat that we don't have the possibility to point back to their origin and their spectrum is affected by its propagation. This could allow to compare model for diffusion with and without a contribution from DM. Notably, the production of positrons through dedicated channels from DM annihilation [107]. The observation by AMS-02 of an excess of positrons [23], as shown in Figure 1.20 has sparked interest as being possibly produced by DM annihilation [108]. This measurement shows the comparison of the contribution from a DM model [109] and a classical contribution.

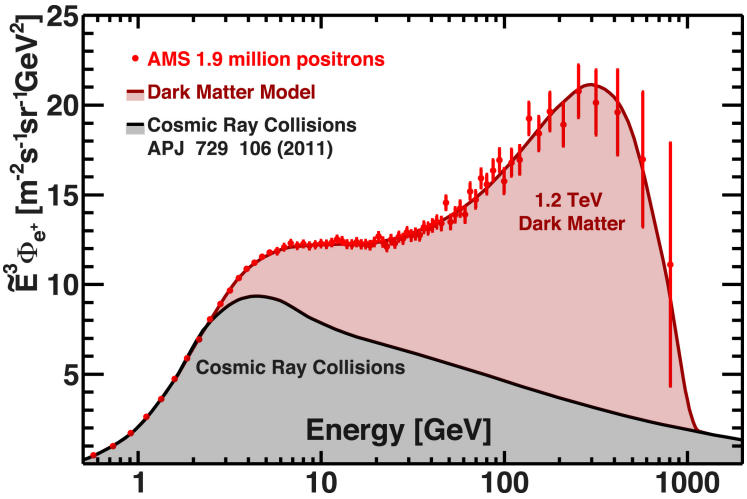


Figure 1.20: Comparison of the AMS positron flux (red data points) with predictions of a dark matter model [109] (red line) based on a DM particle with a mass of 1.2 TeV [23] and classical cosmic-ray contribution (black line)

In addition, observations of a hard cut off in the CRE spectrum have been reported by H.E.S.S. [80]and DAMPE [46]. That been said, explanation from astrophysical and dark models are able to justify such behaviour. Possible dark matter models for the positron excess is in disagreement with constraints from gamma-ray searches which have constrained the cross section to be smaller compared to leptonic channels.

Up to now, the existence dark matter particle has not been confirmed or denied. As discussed in Section 1.7.1 and here motivates the extension of the CRE flux to higher energies.

Chapter 2

The Dark Matter Particle Explorer

Since the discovery of Cosmic Rays (CRs) almost a century ago, our understanding of these high-energy particles has significantly improved, thanks to advancements in detection techniques. These developments have allowed us to push the boundaries of CRs research, as discussed in Chapter 1.

Chapter 1 highlighted how CRs were first detected using the electroscope. However, a deeper understanding required the development and implementation of more precise techniques. It highlighted three different observation methods and discusses both historical and ongoing experiments, providing a comprehensive overview of CR studies.

In this chapter, we focus on the Dark Matter Particle Explorer (DAMPE) experiment (Figure 2.1), which collected the data analysed in this study. DAMPE is a space-based experiment designed to investigate high-energy CRs, gamma rays and potential DM signatures. It was launched in December 2015 from China [110].

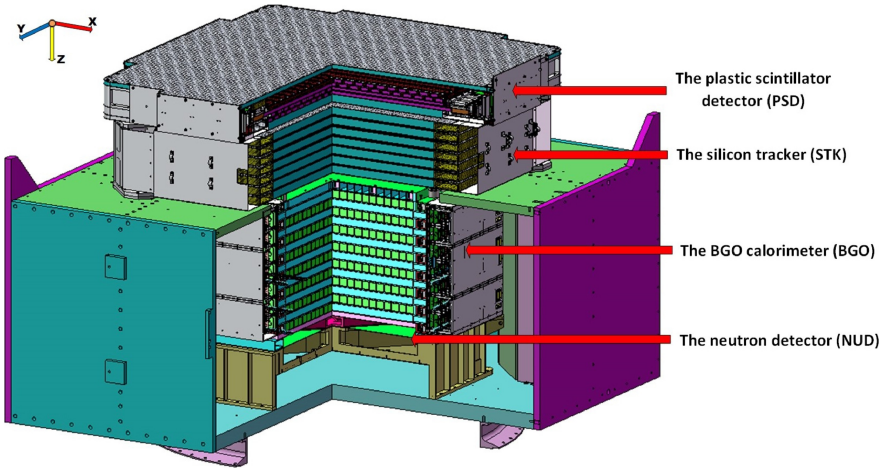


Figure 2.1: Schematic representation of DAMPE and its layout. Figure from [110].

Its primary aim is the measurement of the electron+positron fluxes up to ~ 10 TeV. Additionally, DAMPE provides precise measurement measures of protons and heavier ions from 40 GeV up to the several tens of TeV [111–114] providing insights into their acceleration and propagation mechanisms. By observing high-energy gamma rays it also contributes to the study of astrophysical sources such as pulsars and supernova remnants [115]. Overall, DAMPE enhances our understanding of the origin of CRs, their interaction in the galaxy, and the fundamental physics governing these high-energy particles.

2.1 DAMPE Sub-detectors

The DAMPE payload consists of four sub-detectors, namely the Plastic Scintillator Detector (PSD) [116], the Silicon-Tungsten Tracker-Converter (STK) [117], the Bismuth Germanium Oxide (BGO) calorimeter [118], and the Neutron Detector (NUD) [119], as shown in Figure 2.1. These sub-detectors work together to precisely identify incoming CRs while providing accurate measurements of their charge, energy, and direction.

2.1.1 Plastic Scintillator Detector

The Plastic Scintillator Detector (PSD) is the first detector to be encountered by CRs. Its main focus is: to measure the charge of incident particles from 1 to 26 and to serve as a veto for discriminating gamma rays from charged particles [120]. The detector is composed of two double-layers of plastic scintillators, see Figure 2.2. Each layer is composed of 41 plastic scintillators bars of size $884 \times 28\,(25) \times 10$ mm^3 (for bars on the edge). As shown in Figure 2.3, the top layers are aligned with the x-axis and the bottom ones with the y-axis of the DAMPE coordinates system. The scintillating light is recovered using Hamamatsu R4443 Photo Multipliers-Tubes (PMTs) at the two ends. A total of 8 dynodes with double dynode read-out for each PSD, such that we can cover a high-dynamic range.

The full sub-detector has an effective area of 820×820 mm^2 and a resolution below 25% for $Z = 1$ particles, an efficiency of ≥ 0.95 for each module and an overall efficiency ≥ 0.9975 for charged particles [110].

From the Bethe-Bloch formula [97], the energy deposited is proportional to the charge squared of the incoming particle Z^2. However, as described in [121], some effects have to be taken into account before charge reconstruction, among which light attenuation and quenching effects. We show, in Figure 2.4, the charge reconstruction of PSD for 2 years of on-orbit data.

2.1.2 Silicon-tungsten Tracker-converter

The second sub-detector is the silicon-tungsten tracker-converter (STK) [117], it has the tasks of reconstructing a particle's track, charge measurement and photon detection by its conversion into electron+positron pair.

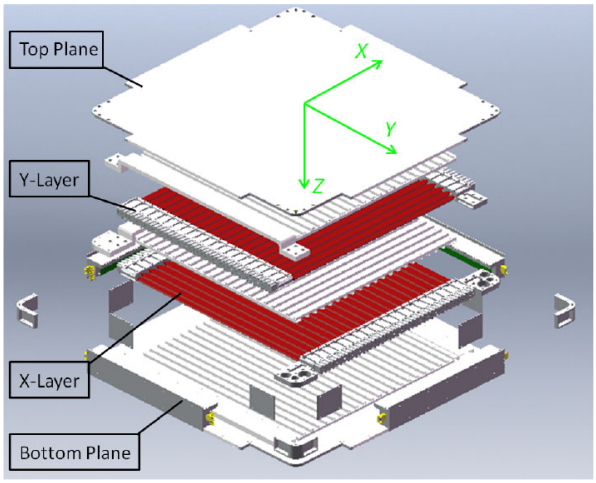


Figure 2.2: Full view of the PSD. Figure from [120]

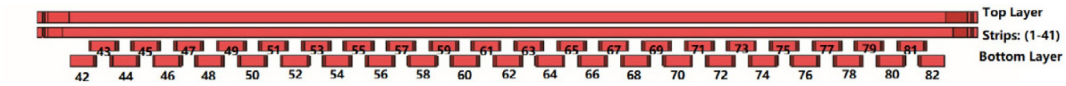


Figure 2.3: The side view of the PSD . Figure from [120]

The STK follows an approach already used in other experiments as Fermi-LAT [122]. It consists of six tracking planes, with each plane made of the two layers of single-sided silicon micro-strip detectors (SSD) measuring the X-Y directions of the incoming particle. To be able to detect the incoming gamma rays, three layers of tungsten of 1 mm thickness are placed before layers 2, 3 and 4 (Figure 2.5) to induce photon conversion. Each silicon layers includes 16 ladders with 4 SSD each, for a total of 192 ladders and 768 SSDs. The thickness of STK is equivalent to 1 radiation length, mostly due to the presence of the tungsten layers. Moreover, a support tray is constructed using an aluminium honeycomb structure sandwiched between two layers of carbon fibre Reinforced Polymer (CFRP), such that it forms a light but steady structure that can sustain the vibrations from launch.

As a charged particle travels through the detector, it ionises the silicon, generating a signal. Using the signal from each ladder, we can reconstruct the trajectory of a CR. In contrast, gamma rays are neutral particles and do not ionise the first layer of STK. However, when passing through one of the tungsten layers, they undergo pair conversion, producing an electron+positron pair that leaves detectable signals in the subsequent layers.

Figure 2.6 shows STK's spatial resolution is lower than 50 μm, for incident angles lower than 40° [117]. Moreover, STK detector can also be used to reconstruct the charge of the incoming particle [125]. Due to saturations effects, STK is only able to measure

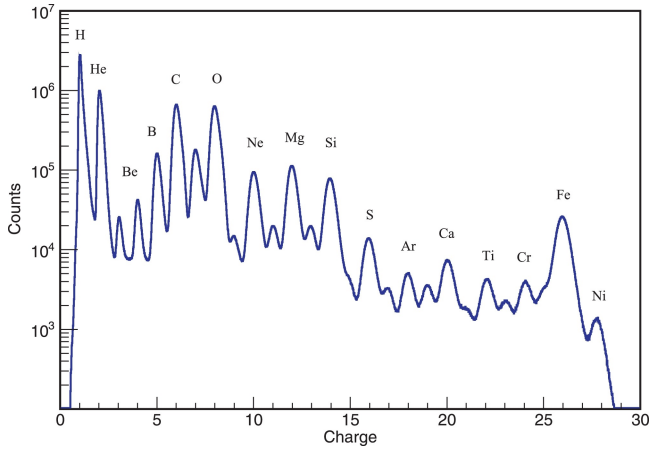


Figure 2.4: PSD's charge spectrum. Figure from [121]

	dy2	dy5	dy8
S0	3.2 GeV-800 GeV	80 MeV-20 GeV	2 MeV-500 MeV
S1	16 GeV-4 TeV	400 MeV-100 GeV	10 MeV-2.5 GeV

Table 2.1: Energy range of the different readout dynode for each PMT [110].

charges from 1 to 6. As shown in [125] ions lighter than Carbon (Z=6) were not affected by this saturation while heavier ions show saturation effects. The charge resolutions were better than 0.3 charge unit from Hydrogen to Neon, as shown in Figure 2.7. Figure 2.8 displays the charge reconstruction for different ions from helium up to Oxygen.

2.1.3 The BGO calorimeter

Bismuth-Germaniun Oxide (BGO) is a scintillator commonly used in calorimetry for electromagnetic particles [126] like electrons and photons due to its excellent stopping power. When a high energy particle interacts with BGO, electrons and positrons will generate an electromagnetic shower and ions will produce an hadronic shower.

DAMPE uses a BGO calorimeter [118]. Its main functions are energy reconstruction, proton/electron separation and serving as a trigger. It is composed of 308 BGO crystal bars set in 14 layers. Each bar has a dimensions of $25 \times 25 \times 600$ mm^3 (Figure 2.9) corresponding to 32 radiation length, which enables a near-3D reconstruction of the shower topology by obtaining the longitudinal and transverse view of the shower development. Each bar is coupled to two PMTs (named S0 and S1) on each end to collect the scintillation light (Figure 2.10). Each PMT has 3 different readout dynodes (dy2, dy5, dy8), which allow to cover a wide dynamic range of energy deposition, as shown in Table2.1.

The detector achieves an energy resolution of 20% at 10 TeV for protons, while for electrons and gamma rays, the resolution is better than 1 % resolution above 10 GeV. Figure 2.11 shows the evolution of the resolution for different energy ranges using beam

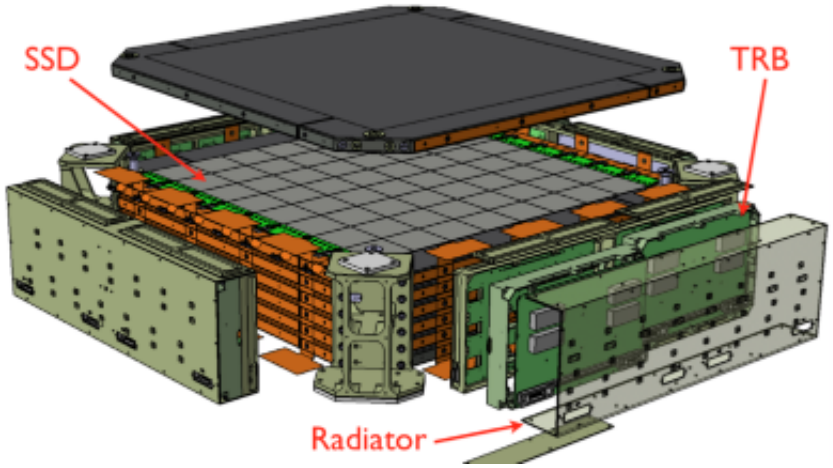


Figure 2.5: View of STK showing the 64 SSDs layers, the readout board (TRB) and the radiator used to dissipated the heat. Figure from [123]

test data.

The electron/hadron separation can be done by looking at the width and depth of the shower in the calorimeter.

2.1.4 Neutron Detector

The electron/hadron separation can be done by looking at the width and depth of the shower in the calorimeter. Nonetheless, hadronic shower are expected to produce a number of thermal neutrons within their showers that is one order of magnitude greater than electromagnetic showers. Those neutrons can be used as an element to improve the discrimination power between electromagnetic and hadronic showers. The NeUtron Detector (NUD), as shown in Figure 2.12, is composed of 4, $30 \times 30 \times 1$ cm^3 blocks of boron-loaded plastic scintillators that can measure the total neutron activity over a few microseconds, thus allowing for a second discriminator between electrons and hadrons. The neutrons produced during in a shower reach NUD from the calorimeter, the neutrons are captured through the photon production reaction:

$$^{10}\mathrm{B} + n \rightarrow {}^{7}\mathrm{Li} + \alpha + \gamma, \tag{2.1}$$

where the ^{10}B captures a n, producing a ^{7}Li, an α and a γ. The gamma emitted in this process is then detected by the 4 PMTs that are attached to the boron-loaded plastic scintillator blocks of the NUD. Figure 2.13 illustrates the NUD signal from a beam test, highlighting the significant difference in signal between electrons and protons. Electrons produce electromagnetic showers, generates a minimal signal, with events registering fewer than 30 channels, and in most cases, fewer than 2 channels. In contrast, hadronic showers produced by protons result in a much larger signal, as these showers are associated with the production of more neutrons, leading to a more significant photon detection

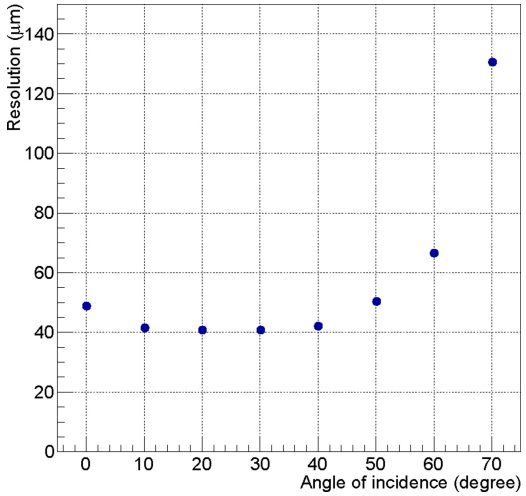


Figure 2.6: Spatial resolution as a function of the impact angle. Below 70 , the resolution is much lower than the 70 μm achieved with a digital position finding algorithm [124]. Figure from [117].

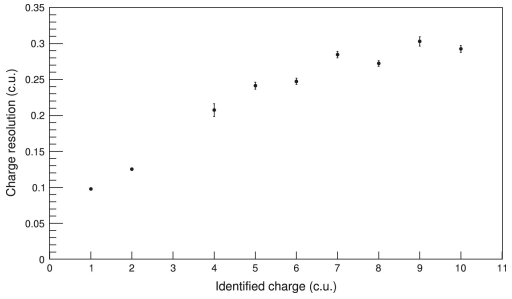


Figure 2.7: STK charge resolution as a function of the identify charge. Figure from [125].

by the PMTs. The difference in signal size between electrons and protons enhances the ability of the NUD to improve the discrimination between electromagnetic and hadronic showers.

2.2 Data Acquisition and trigger

The data acquisition system (DAQ) [110] allows the data collection, implementation of the trigger logic and transfer of the data to the ground. Figure 2.14 shows the structure of the DAQ. It is compose of two main components that include the Payload Data Process Unit (PDPU) and the Payload Management Unit (PMU). The PMU is the control centre of DAMPE. When it receives a trigger signal from the trigger board in the PDPU, it begins to collect science data from the front-end electronics (FEE) on the $-$X/$-$Y sides,

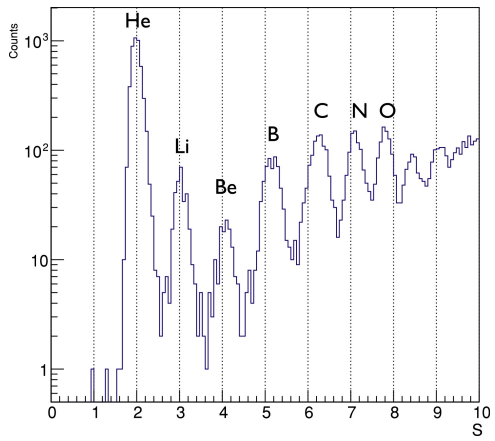


Figure 2.8: STK signal mean distribution for nuclei produced by a lead beam on target, after removing $Z = 1$ particles. The signal mean, with current reconstruction procedures, allows for the identification of ions until Oxygen. Figure from [110].

while data from the FEE on the $+X/+Y$ sides are collected from the PDPU. The trigger board received the signal from BGO calorimeter and makes the decision within $1\,\mu s$

The trigger board receives the signals from eight of the 14 layers of the BGO calorimeter. Then it implements the trigger decision logic with a FPGA chip (Figure 2.15). DAMPE has implemented 4 trigger: unbiased trigger (UBT), MIPS trigger (MIPT), High energy trigger (HET) and Low energy trigger (LET).

Each trigger has a different threshold: the UBT requires that a particle deposited energy in the two top BGO layers above a threshold of 0.4 MIPs at least in one BGO bar; the MIPT is used to select MIP particles (protons) so we ask that an event must have a deposited energy of 0.4 MIPs in layers 3,4,11,12,13 and 14, such that we can select a particle which hasn't shower but crossed all the calorimeter; the HET requires 10 MIP in the 3 first layers while asking for 2 MIP on the 4th layer; the LET requires 0.4 MIPS at the 2 first layers while asking for 2 MIPS for the third and fourth layers. On top of the requirements for each trigger, they are functional at most latitudes with the exception of the MIPT, that is enable at latitude in $[-20, 20]$. Because the DAQ has a dead time of 3.0725 ms [128], all triggers except HET need a pre-scale factor:

- UBT: For latitudes in $[-20, 20]$: 512; Others: 2048
- MIPT: For latitude in $[-20, 20]$: 4
- LET: For Latitude in $[-20, 20]$: 8; Others: 64

2.3 Simulations

Monte Carlo (MC) simulations are essential to evaluate detectors behaviour to incident particles, it is essential both for the design and the optimisation of an experiment.

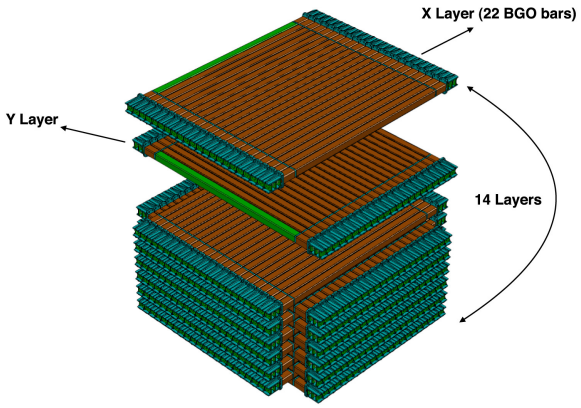


Figure 2.9: BGO bar arrangement in the calorimeter. Figure from [127]

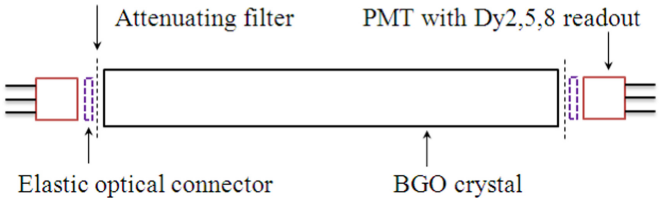


Figure 2.10: Schematic of the BGO readout

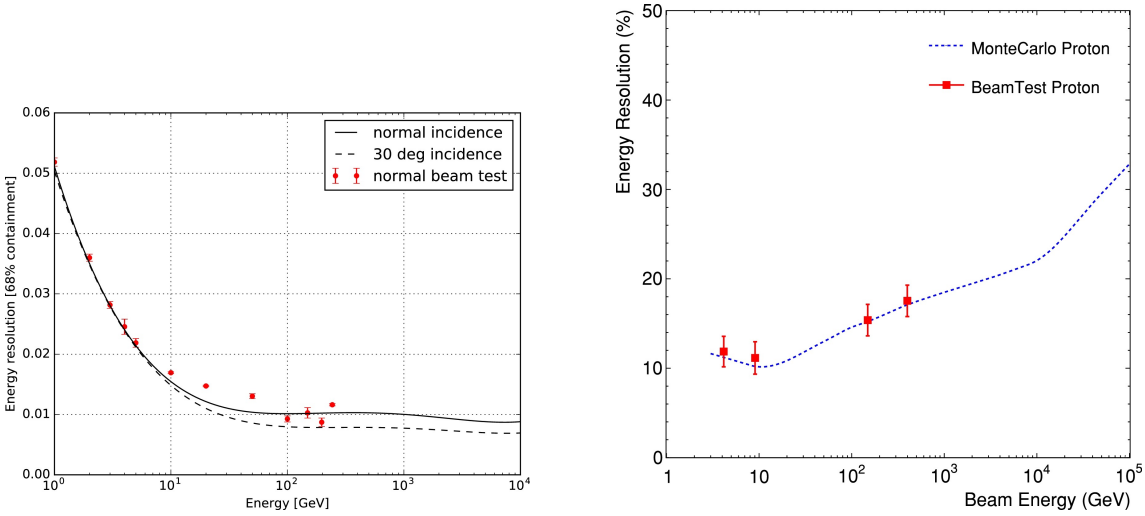


Figure 2.11: Energy resolution of the BGO calorimeter [110]. Left plot: DAMPE beam test results (with electrons) are over-plotted. Right plot: The energy resolution for on-axis protons. The dotted line represents the energy resolution of MC simulated protons after spectral unfolding while the red points represent the beam test data.

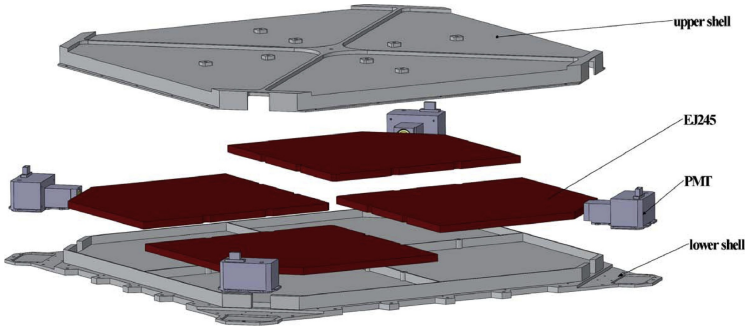


Figure 2.12: Schematic of the NUD detector [110].

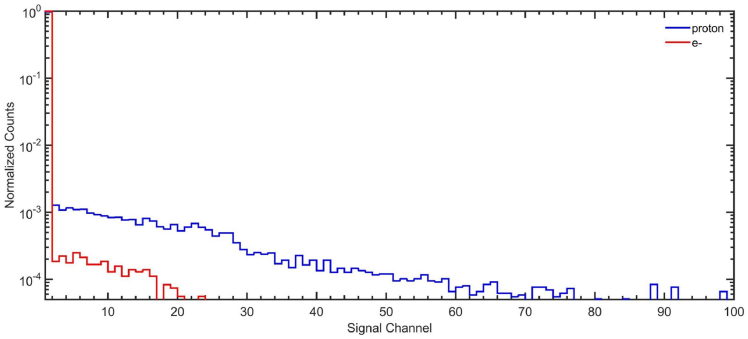


Figure 2.13: NUD signal for 450 GeV protons and 150 GeV electrons (distributions are normalized to unit area) [110].

DAMPE simulations are based on the GEANT4 toolkit [129] widely used by the high energy physics community. Information about geometry, the position and the materials of all detector elements are used by GEANT4 to build a detailed model of the detector. In this simulation, an isotropic particle flux is modeled by a conceptual disk that moves tangentially across the surface of a sphere with radius 1.38 m for electrons, photon and protons, and a radius of 1.0 m for ions heavier than proton. This approach effectively emulates an isotropic distribution of incoming particles.

Signals generated during the simulations go through a digitalisation process to convert hits to ADC signal such that MC samples can be processed by the same reconstruction algorithms as real data.
A "physics list" is taken into account, combined to GEANT4 simulations, to include all known particles and their interactions:

- QGSP_FTFP_BERT: it is use as the default model for simulating electromagnetic processes. For hadronic interactions it includes the Quark-Gluon String Precompound model (QGSP) [130] and the Bertini intranuclear cascade model (BERT) [131].
- FTFP_BERT (The FRITOF Recompound model) [132] is used to described inelastic hadronic interactions for events having energies up to 100 TeV. Mainly used

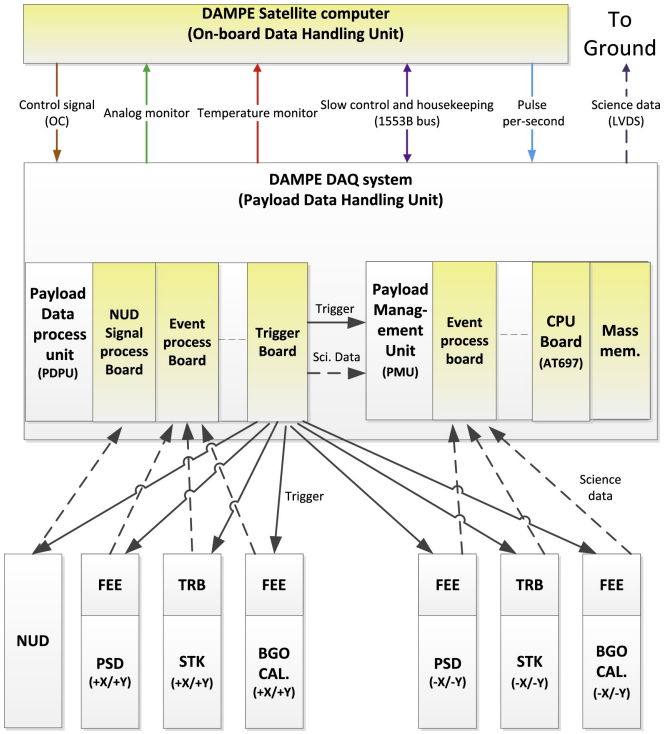


Figure 2.14: Architecture of DAMPE's DAQ [110].

for the cosmic rays measurements of protons and nuclei.

Both models include the option HP (High Precision), used to include neutrons of 20 MeV and lower. This option also allows to include the cross section to describe elastic and inelastic scattering, capture and fission. Mainly used for the NUD detector.

The models are compared to data from experiments at beam facilities like CERN, this comparisons are conducted at energies up to 10 TeV. That being said, there is still some clear mismatches between the predictions of such models and observables. Moreover for energies above 100 TeV, we include a second model EPOS-LHC [133] with Cosmic Rays MonteCarlo (CRMC) [134], [135] implemented an interface between CMRC and Geant4.

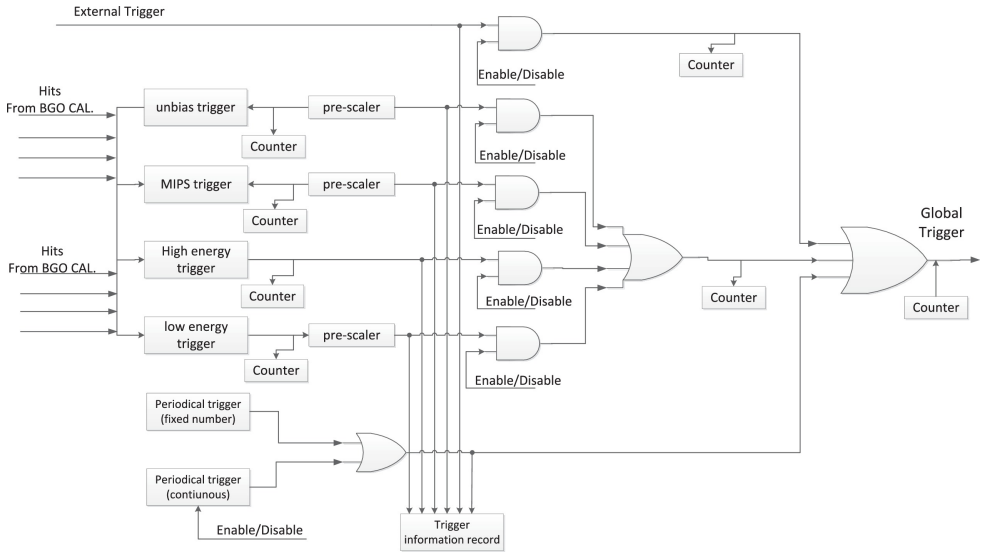


Figure 2.15: Trigger decision logic [110].

Chapter 3

Machine Learning Basics

Machine Learning (ML) is includes a set of methods that use statistical algorithms to learn otherwise hidden features from data. ML is widely used in data-driven applications, ranging from computer vision to speech recognition. Consequently, the physics community has also benefited from ML, applying it to areas such as Monte-Carlo (MC) simulations and data analysis [136], leading to improvements over more traditional methods.

DAMPE has leveraged deep learning techniques for electron-positron and proton discrimination using unsupervised ML methods – such as principal component analysis [137] and supervised one –such as Multilayer Perceptron [138]. Additionally, convolutional neural networks (CNNs) have been employed for track reconstruction [139] and to correct the energy deposition in the calorimeter in the presence of saturation [140]. Other cases of ML in the data will be explore later in this work, see Chapters4 and 5.

This chapter provides a general introduction to machine learning and its working principles, Specifically, with a particular focus on CNNs [141]. Such models can be used to identify electrons (Chapter 5) and reconstruct their energy (Chapter 4).

3.1 Machine learning applications

ML can be use for regression and classification task. Its techniques can be divided into two main categories: *Supervised models* and *Unsupervised models*. Supervised models are trained using labelled data, and are tasked to find a function $f(x, \omega) = y$– with x the inputs, ω the parameters to fit, and y the labels, that maximise the accuracy or precision of the predictions with respect to the true labels. In a problem of binary classification, for instance, the true labels can be either 0 or 1. Unlike supervised learning, unsupervised models are trained on datasets without labels. Instead of learning a direct mapping between inputs and known outputs, these models aim to find patterns, structures, or relationships in the data. Unsupervised learning is particularly useful when dealing with large and complex datasets, where manual labelling is unpractical.

Deep learning (DL) is a specialised branch of ML, that focuses on learning hierarchical representations of data through deep neural networks. Unlike traditional ML methods– which often rely on handcrafted features, deep learning automatically extracts complex patterns and high-level abstractions from raw data. This makes it particularly effective for tasks involving large and high-dimensional datasets, such as image recognition, natural language processing, and scientific data analysis. In particular, next section will discuss the use of CNNs model in deep learning.

DL can be applied to various tasks, which are typically categorised based on their objectives and learning approaches. Possible problems are classification, regression, clustering, dimensionality reduction, and image processing. We focused on the first two, since they are use on in this work. A classifier is used to separate events. A simple example is detecting if a cow is present in a picture or not– for a binary classification problem, or discriminate the same cow based on their fur– for a multiple classification. On the other hand, regressor is used to find complex relations between some variables. An example for this, is finding a relation between the price of chocolate and some economical or social parameters

3.2 Convolutional Neural Networks

In this section, we focus on CNNs. For readers seeking a more detailed overview of the ML domain, additional information can be found here [141]. One of the must common applications of CNNs is image processing, where an picture is represented as a two-dimensional grid of pixels. CNNs excel at capturing spatial hierarchies in data using convolutional operations.

A CNN is composed by the following blocks: *convolutional layers* (which give the name this model), *pooling layers*, *neuron layers*, and *dropout layers*.

Convolutional Layers

A convolutional layer accepts an input I, which can be the initial image, or the output from previous convolutional layer. Then, a sort of "product" is performed between the input and a kernel K– i.e., an $a \times b$ matrix, whose dimensions are set by the user. The resulting output S is a feature map of the input, whose elements:

$$S(i,j) = \sum_{m=1}^{a} \sum_{n=1}^{b} I(i+m, j+n) K(m,n) \tag{3.1}$$

Overall, the convolutional layers are used to highlight different features in an image; see Figures 3.1 and 3.2 for an example. The type of layers can be used multiple times, one after the other, with different kernels. A convolutional layer can be tweaked further, setting the stride s used to move the kernel around the input image. Another enhancement is the symmetrical addition of a pair of rows and columns on either side of the original input. This technique, called padding, avoids the loss of information at the perimeter of the image, where the kernel cannot go further. The output size of a convolutional layer

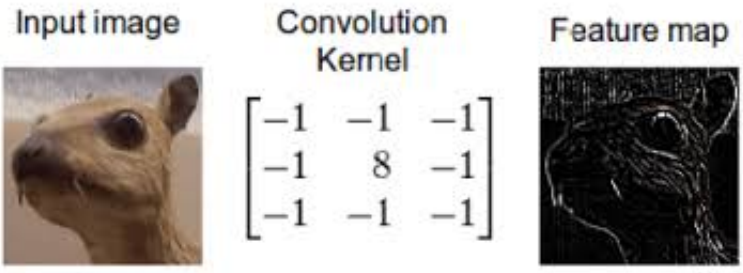


Figure 3.1: Example of how a convolutional layer extracts features from an input image. On the left, the original image given to the convolutional layer; in the centre, the kernel used by the layer, to detect ridges and edges; on the right, the output feature map. Images originally from [142]

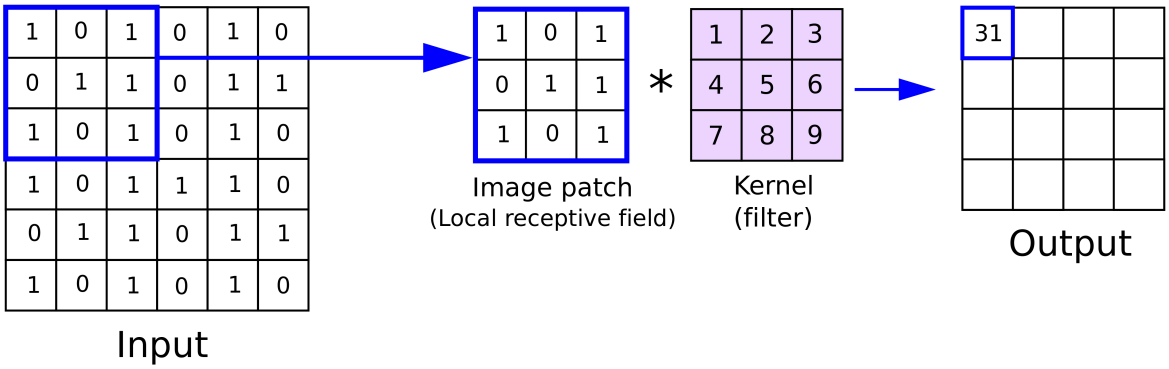


Figure 3.2: A convolutional layer applies a kernel K to a local region (image patch) of the input I. On the left it represented the input's pixel values. The center represents the convolutional function, as described in Eq 3.1. The 3×3 kernel is slid over the input image, performing element-wise multiplication with the overlapping image patch followed by summation. The result for each patch is placed in the corresponding position of the output feature map, as shown in right image. In this example, the dot product of the first image patch and the kernel yields a value of 31 in the top-left cell of the output. Images originally from `https://anhreynolds.com/blogs/cnn.html`

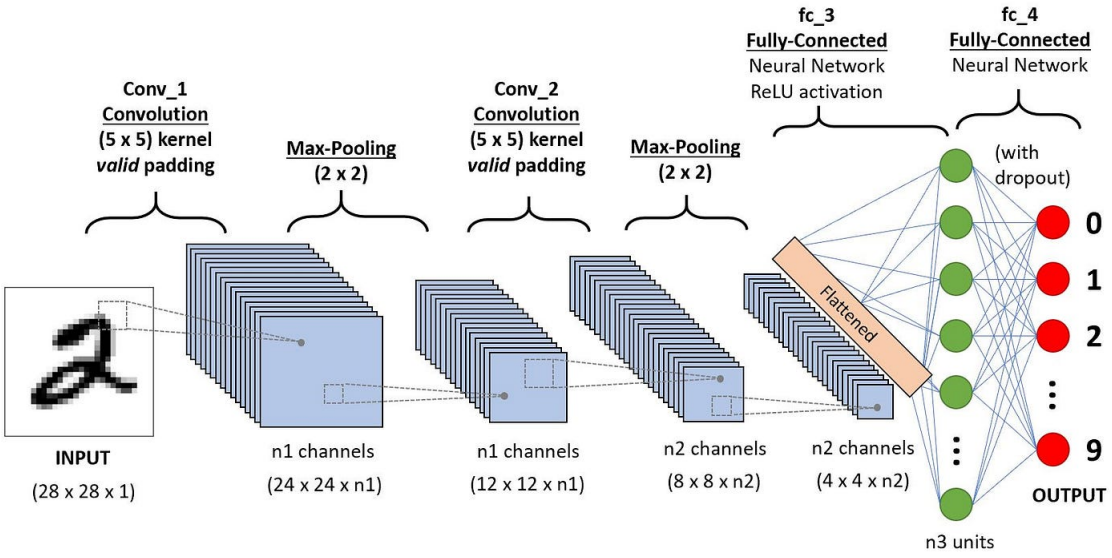


Figure 3.3: Visual representation of a CNN wit. Images originally from `https://www.plainconcepts.com/convolutional-neural-network-guide/`

can be obtained as

$$\mathrm{W_{out}} = \frac{\mathrm{W_{in}} - \mathrm{a} + 2\mathrm{p}}{s} + 1, \mathrm{H_{out}} = \frac{\mathrm{H_{in}} - \mathrm{b} + 2\mathrm{p}}{s} + 1, \tag{3.2}$$

where $\mathrm{W_{in\,(out)}} \times \mathrm{H_{in\,(out)}}$ are the dimensions of the input (output), a $\times$ b the ones of the kernel, s the stride, and p is the number of rows/columns added to either side of the input (e.g. $p = 1$ corresponds to adding one row at the top and at the bottom and on column on the left and on the right side of the original image). Thus allowing to be invariant under local translation, meaning that if the input is shifted the output is also shifted by the same amount.

Pooling Layer

A pooling layer aims to reduce the high dimensionality of the maps resulting from convolutional layers, summarising the feature information. The usage of this type of layers reduce the computational cost as well as the number of hyper-parameters.

In pooling, a shape is set for the pool, which is no more than a lens focusing on specific parts of the map. Then, the pool selects one of the features, according to some filter, e.g., the element with the maximum value in the pool. As for the convolutional layers, also for pooling layers a stride and a padding can be set.

Neuron Layers

After the image as been processed by the convolutional and pooling layers, the final output is reshaped to a one dimensional array, that serves as input to the neuron layers. A neuron layer takes an input and applies some weights to it, that it be assimilated to the parameters to be fitted, as in a linear regression problem, for example. Then, the sum of the weighted input values is passed to a function– called **activation function**, that does not need to be linear or continuos, and that returns the final output value. A simple example of activation functions, is a function that returns 1 if the sum of the weighted input is above a given threshold or otherwise. Other examples of usual functions are the rectified linear unit (ReLu) and sigmoid. The ReLu is defined as :

$$\text{ReLu}(x) = \max(0, x) \tag{3.3}$$

The sigmoid function returns values between 0 and 1, that can be considered as probabilities, and is a differentiable function, which is an important feature in ML models. The function definition is:

$$\text{S}(x) = \frac{1}{1 + e^{-x}} \tag{3.4}$$

Different activation functions can be most suited for different use cases, and not all of them can be used for every task.

Figure 3.3 display an example of CNN architecture with the different components described previously.

Dropout Layer

One of the major issues in ML models is overfitting, i.e. the model learns the specificities of the training data loosing in generalisation. To avoid that, we can randomly drop some neurons from the network during training, thus preventing them from co-adapting too much [143]

After constructing a CNN, it is crucial to evaluate its performance, i.e. how well it learns the task. This requires a loss function to quantify the difference between predictions and true values. The loss function is minimised during training to find the parameters values. In deep learning, a key method for this is the gradient descent, which iteratively adjusts the model parameters to reduce the loss.

Loss function

As any task that requires fitting, we need to estimate how well the predictions of the model compare to the real values. To do so, we choose a function– called loss function, to estimate the goodness of the model. It can be defined by the user, however some well known functions are:

- the mean square error (MSE), i.e., $L(y, \widetilde{y}) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widetilde{y}_i)^2$
- the cross binary entropy, i.e., $L(y, \widetilde{y}) = \sum_{i=1}^{n} y_i \cdot log(\widetilde{y}_i)$

where y is the true value, $\widetilde{y}$ the predicted one and n is the number of entries. The first is useful for regression tasks, while the second for binary classification.
The loss function also allows us to estimate if the model under- or over-fits the data. For example, Figure 3.4 shows a distribution of points with some noise and three different fitting functions: linear, quadratic and 14th-order polynomial. The linear fitting shows underfitting, while the polynomial function shows overfitting, and hence a lack of generalisation. Thus, if the model was presented with a new set of points, it will not be able to generalise and predict their values. In contrast, quadratic fitting results in a general understanding of the data. To avoid such problems, it is important to split the dataset used for training an ML model into three subsets: training, validation, and test. The training sample is used to fix the parameters, such that we minimise the loss. The validation sample is also used to minimise the loss based on the parameters obtained from the training. In this way, we can compare the resulting loss function for both. Finally, the test sample is used, as its name suggests, to evaluate the model on unseen data and check for any bias and evaluating the variance when looking at the output of the model.

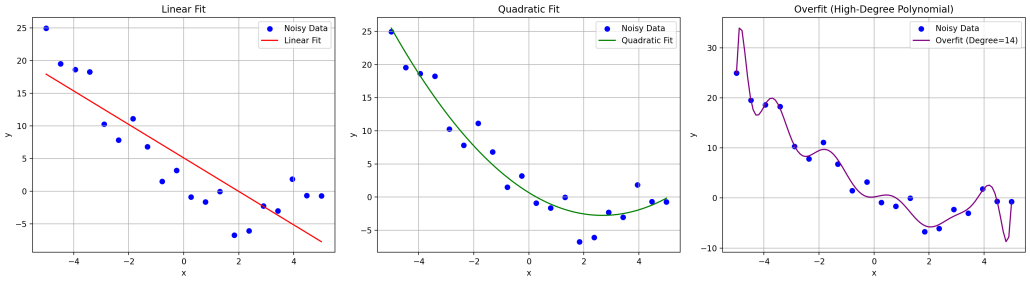


Figure 3.4: Example of under and overfitting. Blue points are a points with some noise. On the left, a set of data points is fitted with a linear function; the same set of points is fitted with a quadratic function, at the centre, and a 14-th order polynomial function. The results show how the complexity of the linear fitting is not enough to describe the behaviour of data, while the polynomial function fits too well to the specific observation, loosing on generality, if provided with a new sample of the same observable. The quadratic function has the better balance to fit the given sample without loosing in generality.

Gradient descend

ML techniques can be simply seen as algorithms that minimise the loss function L for a given task. Despite this simplified view of ML, the high dimensionality of such models does not allow us to analytically find the minimum, and a numerical minimisation is employed. Moreover, multiple local minima may exist. However, we can expect that around a given point, the function is differentiable and use this property to our advantage. By leveraging this, we can find the direction in which the loss decreases the fastest. This technique is called gradient descent (GD), and it is defined in the following way:

$$\vec{x}_{m+1} = \vec{x}_m - l \cdot \vec{\nabla} f(x_m), \tag{3.5}$$

where $\vec{x}_m$ is the previous point, $f(\vec{x}_m)$ the function to be minimised, and l the step size, also called learning rate; since no previous point is available at $m = 0$, $\vec{x}_0$ is chosen randomly. Applied to ML, $\vec{x}_m$ represents the weights and $f(\vec{x}_m)$ is the loss function.

The learning rate is a hyper-parameter chosen by the user, and it can range between 0 and 1. The value of the learning rate should be carefully selected: if it is too high, the optimization process may overshoot the minimum; if it is too low, it might lead to a slow convergence, stalling the learning on a local minimum, or– completely failing to find one. The selection of the learning rate can be done by brute force, and verifying different values of it and comparing the models performances.

We illustrate the notion of gradient descent by means of an example. Let's define a function $f(\vec{x}) = x_{m,i}^2 + x_{m,j}^2$ with $\vec{x}_m$ being a array with two elements $x_{m,i}$ and $x_{m,j}$. Now we define a first value for the initial iteration $m = 0$ the value for $\vec{x}_0 = (1, 3.5)$, and we chose a learning rate $l = 0.1$. At the step $m = 1$, we have moved to the point:

$$\begin{aligned} \vec{x}_1 &= \vec{x}_0 - l \cdot \nabla f(\vec{x}_0), \\ \vec{x}_1 &= (1, 3.5) - 0.1 \cdot (2, 7) = (0.8, 2.8) \end{aligned} \tag{3.6}$$

Then for the second iteration– m=2, we move to the point:

$$\begin{aligned} \vec{x}_2 &= \vec{x}_1 - l \cdot \vec{\nabla} f(\vec{x}_1), \\ \vec{x}_2 &= (0.8, 2.8) - 0.1 \cdot (1.6, 5.6) = (0.64, 2.24) \end{aligned} \tag{3.7}$$

We can continue until we find the minimum, or the obtained value is below a certain threshold. Figure 3.5 shows the the example of the gradient descent after 50 iterations.

While this method can minimise the loss function, some problems may arise. First, GD depends on the data dimensionality; if the size is sufficiently large, it can accurately represent a given function. However, this also makes the gradient descent application more computationally expensive. A workaround is to look at a random subset of the dataset and estimate an approximation of the gradient based on it. This method results in faster iterations, at the expense of lower convergence, and is called stochastic gradient descend (SGD) [144]. A second problem is that the gradient may reach a local minimum. While this is not necessarily a problem if such local minimum is close enough to the global one and we accept a decrease in performance, it may also cause the gradient to oscillate around the global minimum without reaching it. To avoid this, one can take inspiration from Physics and add a certain inertia to the gradient. This is achieved introducing a new hyper-parameter called momentum [141], which depends on the change of $\vec{x}$ (i.e., $\Delta\vec{x}_n = \vec{x}_n - \vec{x}_{n-1}$), that will smooth the direction followed by the gradient descent.

Now that all the steps have been defined, we can train a model. At the beginning of the training, the different weights are randomly set and will be optimised at each iteration. One iteration is completed when all data is used for training the model. Such iterations are called epochs. Multiple epochs are needed to be able to minimise the loss

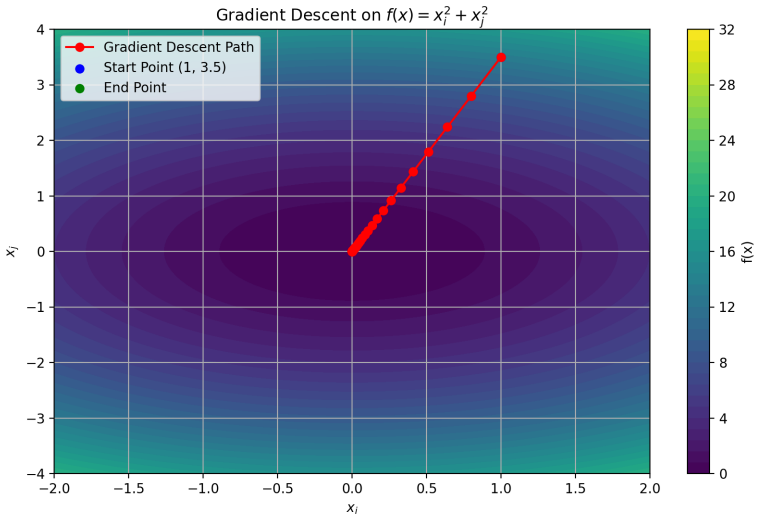


Figure 3.5: Visual example of the gradient descent technique used for a function $f(\vec{x}) = x_i^2 + x_j^2$, with $\vec{x}$ a bi-vector of elements x_i, x_j. The red line shows the gradient descent path for 50 iterations, with a starting point $\vec{x}_0 = (1, 3.5)$.

function. During each epoch, the data is split into smaller samples called batches used for the training.

Chapter 4

CNN method for energy reconstruction for electrons

The work of this chapter has been submitted to JINST [145].

As already mentioned in the Introduction as well as in Chapter 1, the electron+positron flux is a key element for the understanding of possible local sources [2, 146] as well as dark matter [32, 147, 148]. However, the current published results were limited to 5 TeV from DAMPE [46] and 7 TeV from CALET [45]. In order to improve current DAMPE results. To increase the statistics, we expanded the acceptance by selecting events outside the fiducial volume. Such events will be named non-fiducial events. These events have a greater incidence angle, leading to increased shower leakage of the shower outside the BGO calorimeter and additional energy losses due to dead material. To mitigate these effects, [149] developed a technique to correct for them. However, this method was optimised for events with near-normal incidence to the detector (fiducial events). In Section 4.1, we define non-fiducial events. Section 4.2 introduces the classical energy reconstruction method applied to these events. Next, we present the CNN model used for energy reconstruction. Finally, in Section 4.4 we discuss the results based on MC simulations and data.

4.1 Non-fiducial Events

To start, we describe the event selections applied in previous publications [46, 150]. This set of initial selections is based on the shower development in the calorimeter, described as follows:

- Rejecting events that enter the calorimeter from the sides, by imposing a energy fraction in one layer to not exceed 35%.
- Rejecting events where the shower direction cannot be reconstructed.

- Ensuring that the reconstructed shower direction extrapolates to the top and bottom of the BGO sensitive volume, within a distance of 280 mm from the centre, in either the X or Y direction.

These selection aims to select events that are well reconstructed and contained within the calorimeter. Non-fiducial events are defined by inverting the last cut, i.e. by requiring that the bottom extrapolation is not within a distance of 280 mm from the centre of the calorimeter. Figure 4.1 shows a view of the shower for a non-fiducial and fiducial event. On top of that, we excluded events in the SAA and we applied the HET trigger (see Section 2.2).



Figure 4.1: View of a non-fiducial event (left figure) and fiducial event (right figure) crossing the calorimeter. The BGO images combine the X-Z and Y-Z views

4.2 Classical Energy Reconstruction Method

As discussed in section 2.1.3, the energy reconstruction of a particle is based on the energy deposited in the calorimeter [149]. The classical method for energy reconstruction relies on the energy deposited in each BGO bar, which is obtained from the multi-dynode readout. This approach enables energy reconstruction up to 10 TeV with a resolution of 1.2% at energies above 100 GeV for electrons [110]. However, due to leakage and energy losses in the dead material between bars, adjacents bars, and the front of the detector, the reconstructed energy is generally lower than the true energy. The work in [149] took into account this effect and developed a correction method. However, this methods were primarily optimised for fiducial events, which are typically associated with low incidence angles. For these events, the lateral and longitudinal shower development is well understood, allowing for reliable energy reconstruction. When applied to non-fiducial events, the performance degrades because a significant portion of the shower develops outside of the calorimeter. This results in missing energy and poorer reconstruction. As illustrated Figure 4.1, non-fiducial events often exhibit complex shower topologies, necessitating more advanced reconstruction techniques.

To evaluate such methods on non-fiducial electrons, we look at the relative error

between a given reconstruction method and the kinetic energy:

$$r = 1 - \frac{E_{\text{rec}}}{E_{\text{kin}}}, \tag{4.1}$$

where E_{rec} is the reconstructed energy and E_{kin} is the kinetic energy of the initial particle.

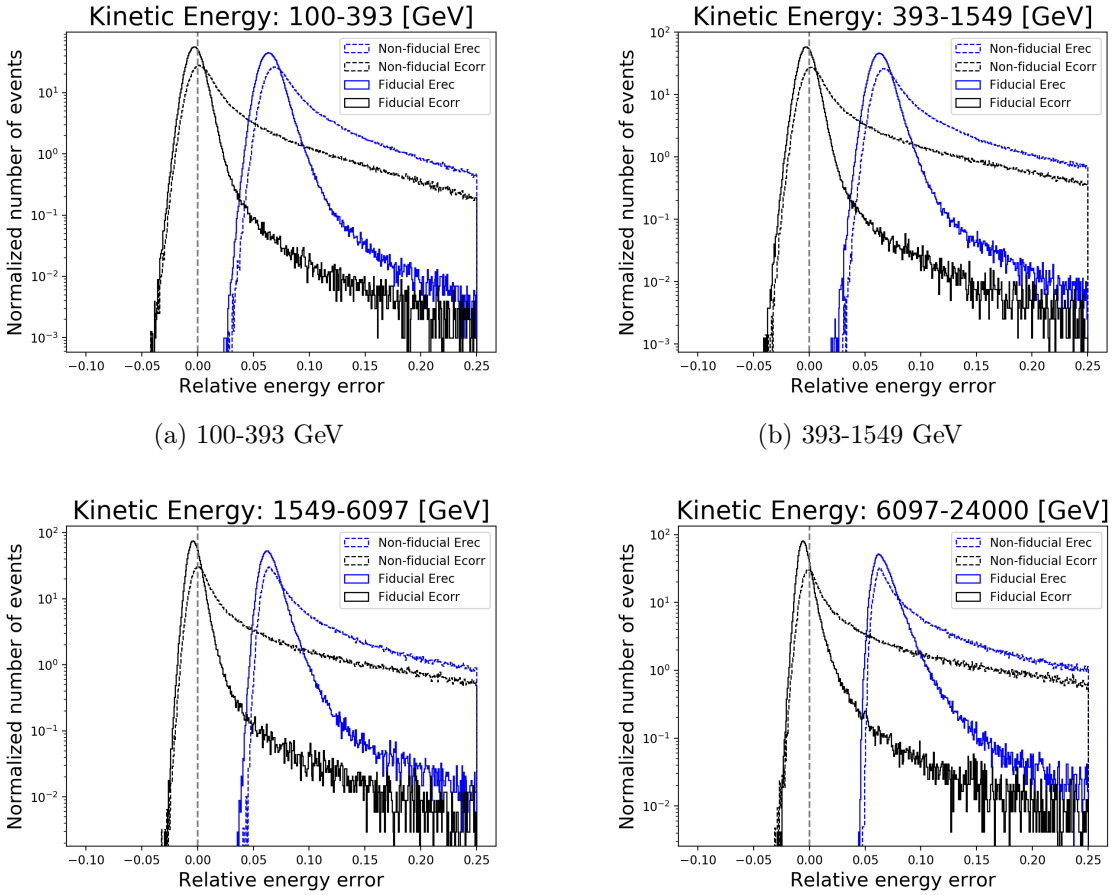
Figure 4.2 showcases the distribution of the relative error for both types of events. First of all, E_{rec} is at least 5% lower than the kinetic energy for both non-fiducial and fiducial events. This is not surprising since part of the shower leaks outside the detector. The distribution of the correction method (E_{corr}) gives a better estimate of the particle's energy, as the bulk of the distribution is centred around zero. Figure 4.3 illustrates the stability of the mean–E_{corr}– at all energies for fiducial events, while for non-fiducial events, we observe that the mean has improved by a factor of 4 but is still not comparable to the fiducial results. An increase in the mean is also observed. Additionally, we observe a long tail (Figure 4.2) for the non-fiducial events resulting in a significant number of events underestimating the particle's reconstructed energy. This is also confirmed by Figure 4.4, which shows the 68/95% energy containment.

This shows that while E_{corr} [149] improves energy measurements by yielding values closer to the kinetic energy for fiducial events—accounting for leakage and dead material effects—it now tends to overestimate the energy relative to the true kinetic energy. For non-fiducial events, a new correction method is required.

4.3 CNN model

In Chapter 3, we introduced the notion of ML techniques and how such models have been widely used in various domains of physics [151]. DAMPE has employed these techniques [137, 139, 140, 150].

We train a CNN to perform a regression of the relative error (Eq 4.1) between E_{rec} and E_{kin}, allowing us to correct the energy of non-fiducial events. The relative error is chosen instead of directly using the kinetic energy, as it allows the model to achieve better performance by keeping values bounded between 0 and 1. As an input, the CNN receives an image representation of the BGO calorimeter (Figure 2.9), combining both the X-Z and Y-Z views. As described in Chapter 2, the calorimeter layers are arranged in a hodoscopic pattern, which creates a correlation between the two views. Previous studies [139, 140] have shown that combining both views yields better model efficiency. To ensure consistency, the input is normalised by the maximum deposited energy in the calorimeter, producing images with pixel values between 0 and 1.
The model was trained based on MC simulations of electrons. The MC simulations are generated using Geant4 (Version 10.05) [129] with FTFP-BERT physics [132] . The MC samples were selected based on the event selection described in Section 4.1. We divided the MC sample into a training sample (containing 70% of the events) and a validation sample (containing 30% of the events). In Chapter 3, we mentioned that splitting data



(a) 100-393 GeV

(b) 393-1549 GeV

(c) 1549-6097 GeV

(d) 6097-24000 GeV

Figure 4.2: Distribution of the relative energy error obtained with electron MC simulations in four different energy bins, ranging from 100 GeV to 24 TeV. Two types of MC samples are used – fiducial and non-fiducial, analysed with two energy reconstruction methods – standard [110] and corrected [149]. Distributions are denoted as follows: fiducial standard (blue solid line), non-fiducial standard (blue dashed line), fiducial corrected (black solid line) and non-fiducial corrected (black dashed line). The grey line indicates ideal case when reconstructed energy equals the true kinetic energy.

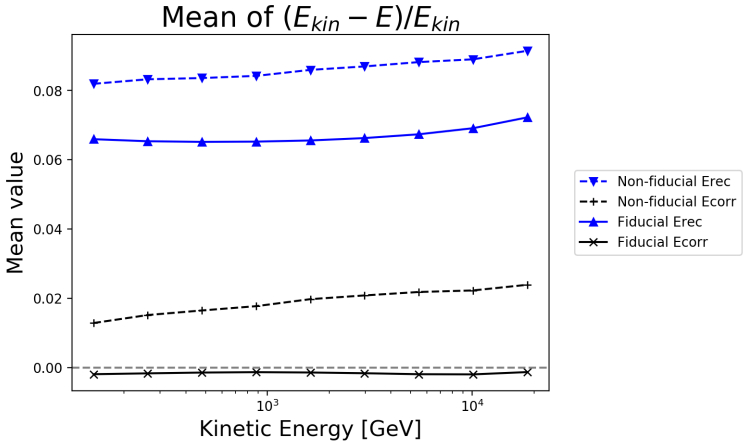


Figure 4.3: Mean of the relative energy error as a function of kinetic energy for the different energy reconstruction methods.

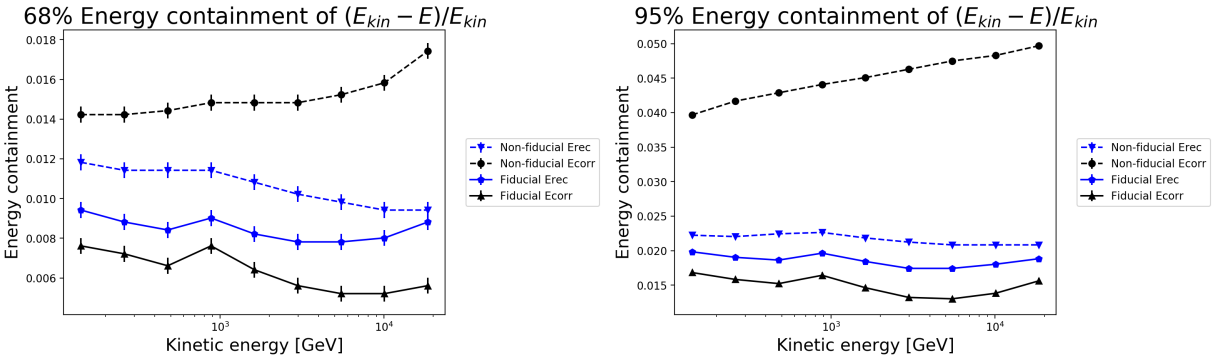


Figure 4.4: Energy resolution as a function of kinetic energy. Left plot: 68% containment, right plot: 95% containment.

into three sets—training, validation, and test—helps assess the model for under-fitting or over-fitting. However, we omitted the test sample because, in principle, it should represent unseen data. In our case, the test sample would still originate from the same MC sample, making it not truly independent. Therefore the full MC sample was used to estimate the model. To be able to have uniform performances over all energies, the training sample was split into 10 bins equally spaced in logarithm of kinetic energy (from 100 GeV to 24 TeV), then we choose the energy bin with the lowest number of events and select the same number of events in each of the other energy bins. The resulting training sample was composed of 1'084'912 events.

The framework used for the training the CNN is Tensorflow 2 [152]. Figure 4.5 showcases the structure of the model. The architecture of the CNN used in this study

was selected based on an exploratory approach. Given the absence of a universally optimal configuration for the task at hand, an initial structure was chosen heuristically, drawing on commonly used architectures in similar contexts. This served as a baseline for experimentation. Although this selection was not guided by a formal optimization process, it allowed for practical implementation. As we will see in Section 4.4, the models yields a sufficiently robust performance for the scope of this research. The convolutional

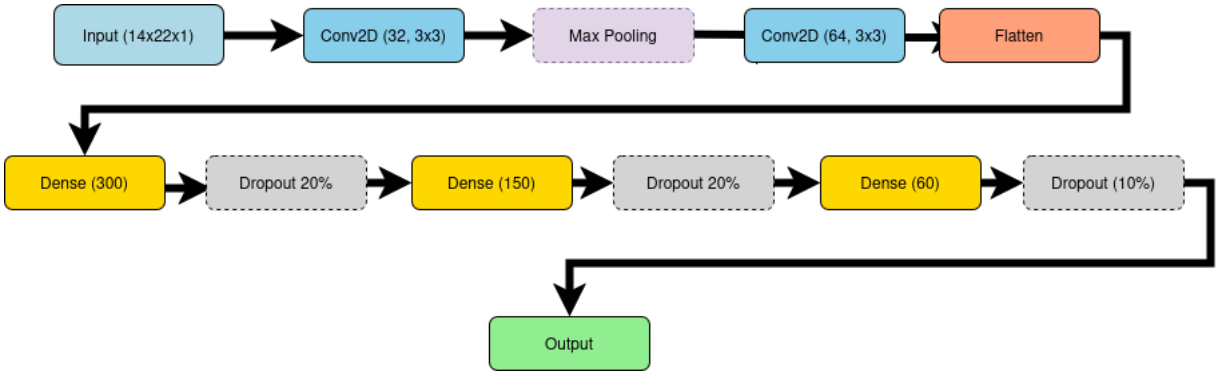


Figure 4.5: CNN model structure [145].

layers in the proposed CNN utilise 3×3 kernels. This kernel size is a common choice in deep learning models, including well-known architectures such as VGGNet [153], due to its favourable trade-off between spatial feature extraction and computational efficiency. A 3×3 kernel effectively captures local patterns (e.g., edges, corners, and textures) while maintaining a manageable number of parameters. The activation function used is a ReLU. It has become the default activation function for many types of neural networks [139, 140, 154] because a model that uses it is easier to train and often achieves better performance. Finally, as for loss function, we choose the mean squared error (MSE) which is a common loss function used in regression tasks. It measures the average squared difference between the predicted value and the true values, ensuring that larger errors are penalised more heavily than smaller ones. Mathematically, it is defined by: We chose an inital value for the learning rate of 0.01, that being said an optimiser adaptive moment estimation (ADAM) [155] was chosen to adjust the learning rate based on the previous value of the gradients.

To regularise the loss we added dropout layers, the key idea is to randomly drop units (along with their connections) from the neuron part of the model. this prevents the model to overfit during training.

$$\frac{1}{\mathrm{n}}\sum_{i=0}^{\mathrm{n}}(\hat{r}_i - r_i)^2, \tag{4.2}$$

where n is the number of training samples, r_i is the relative error (Eq 4.1) and $\hat{r}_i$ is the CNN prediction.

Figure 4.6 shows the evolution of the loss function, the training finished after 40 epochs. We first observe that there is no major deviation between training and validation loss, indicating that there is no under- or over-fitting. Secondly, the loss has converged close to $1.75 \cdot 10^{-4}$ ($1.50 \cdot 10^{-4}$ for the validation loss). Notably, the validation loss is slightly lower than the training loss. This may be attributed to the composition of the validation dataset, which could include simpler events that are easier to predict accurately.

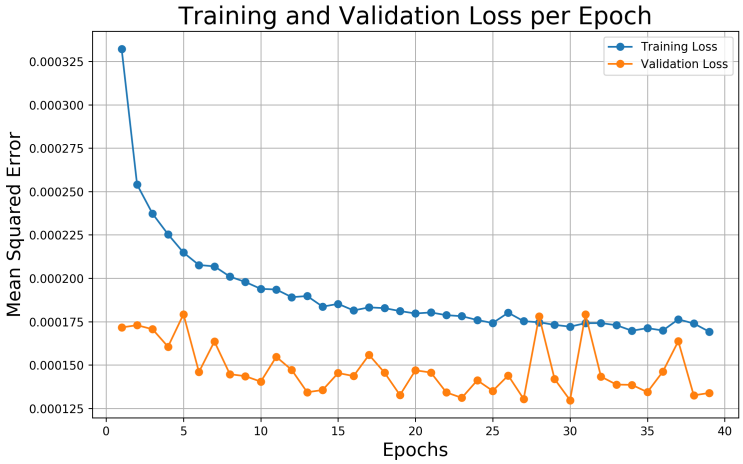


Figure 4.6: Evolution of the loss, defined as the mean squared error for the training sample (blue curve) and validation sample (orange curve).

4.4 Results and Model validation

With the CNN trained, we can now obtain the reconstructed energy E_{cnn}, according to Eq 4.1, as follows:

$$E_{\text{CNN}} = \frac{E_{\text{rec}}}{1 - r}, \tag{4.3}$$

where r is the CNN prediction.

To compare the CNN results to the classical methods, we need to calculate the relative error for E_{CNN}:

$$r_{\text{CNN}} = 1 - \frac{E_{\text{CNN}}}{E_{\text{kin}}} \tag{4.4}$$

4.4.1 MC validation

Figure 4.7 illustrates the distribution of the relative error for CNN and the classical methods for non-fiducial electrons. While the CNN was trained on electrons only, its

performance is expected to be identical for both electrons and positrons since their electromagnetic showers are effectively indistinguishable at energies sufficiently higher than 10 MeV (see Figure 34.11 from [156, page 582]).

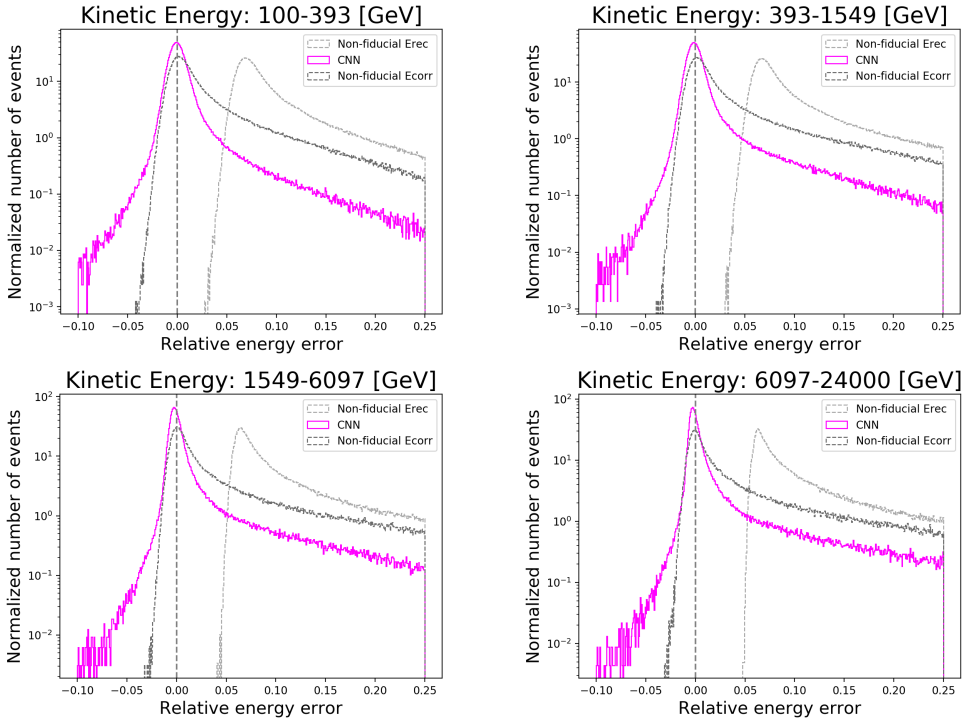


Figure 4.7: The distribution of the relative energy error (Equation 4.1) for non-fiducial electron MC samples ranging from 100 GeV to 24 TeV. Light grey curve: for the standard energy reconstruction method [110], dark grey curve: standard electron energy correction method [149] and magenta curve: CNN method [145]. The grey vertical line indicates the ideal case where the reconstructed energy equals the true kinetic energy.

The distribution of the relative error for the CNN method (magenta curve) has improved the right tail by lowering the number of events with lower estimated energy compare to E_{kin}. Nonetheless, we can see a more pronounced bias toward overestimating the energy. We now look at the mean and the 68/95% energy containment (Figures 4.8 and 4.9. The 68/95% energy containment is used to quantify the spread of a distribution—specifically, to find the smallest radial interval that contains a desired fraction of data points.

Figure 4.8 shows that the mean value for the CNN method stays stable at all energies, demonstrating apparent improvement over the E_{corr} method. The same can be seen for the 68/95% energy containments in Figure 4.9. In particular, the 68% containment for the CNN is consistently lower (better) than for the standard methods and decreases (improves) with energy. The 95% containment shows similar behaviour, although with less improvement compared to the standard methods, especially at high energies. These

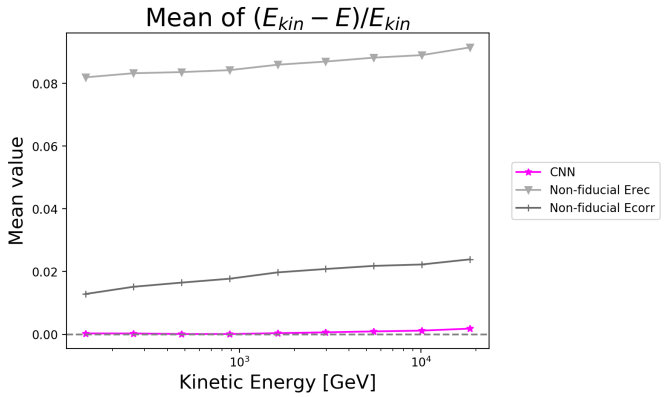


Figure 4.8: Plot of the mean of the relative energy error (Equation 4.1) for non-fiducial electron MC samples ranging from 100 GeV to 24 TeV. Light grey curve: for the standard energy reconstruction method [110], dark grey curve: standard electron energy correction method [149] and magenta curve: CNN method [145]. The horizontal grey line indicates the ideal case when the reconstructed energy equals the true kinetic energy.

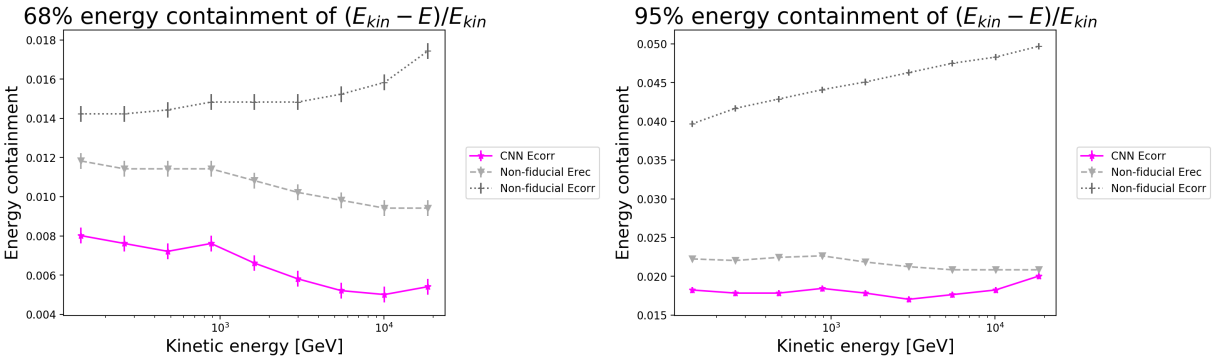


Figure 4.9: The 68% and 95% containment of the relative energy reconstruction errors for non-fiducial electron MC samples as function of the kinetic energy.

results demonstrate a significant improvement compared to the classical methods.

A common technique used in CR data analysis is unfolding [129], which is used to estimate the number of events as a function of the true energy by leveraging knowledge from the detector (typically using MC samples). This process involves constructing a response matrix that characterises the deposited energy as a function of the kinetic energy (Figure 4.10). We will demonstrate that the developed CNN method eliminates the need for unfolding in flux calculations.

To illustrate this, we define a response matrix for each energy reconstruction method and use the Python library pyUnfold [157] to apply the unfolding method. Then, we

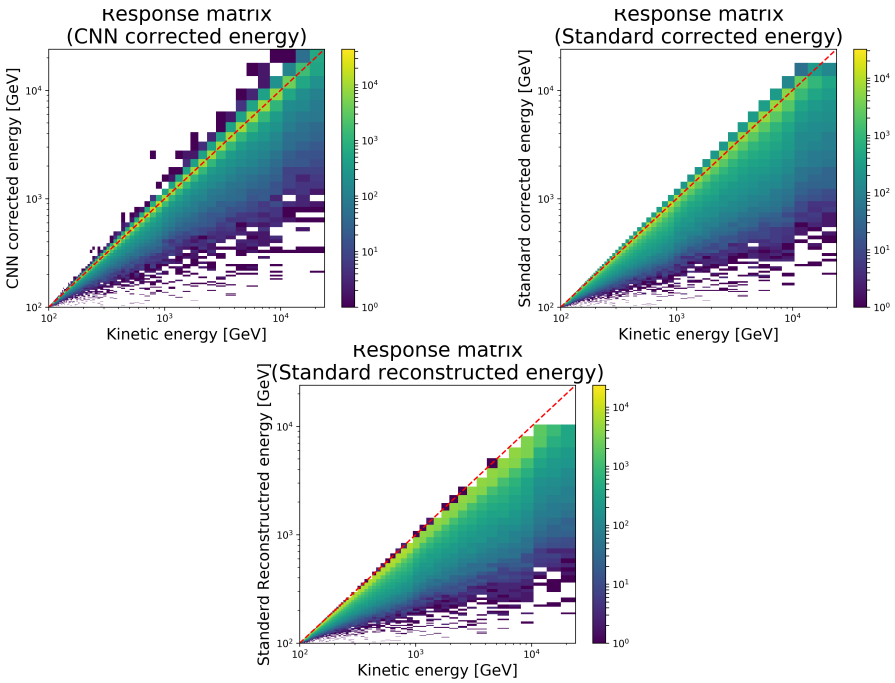


Figure 4.10: Response matrix of the three energy reconstruction methods: CNN-corrected, standard corrected and standard reconstruction (no correction). Red line indicates a diagonal of a matrix.

compare the number of events normalized by the energy (after unfolding) with those derived without such step, showcasing the advantages of the CNN approach.

From Figure 4.11, it can be seen that the normalized CNN energy and the unfolded CNN energy are in good agreement with the kinetic energy.

Figure 4.12 shows the ratio of electron energy spectrum calculated using the CNN-corrected energy method, with and without unfolding, divided by the truth spectrum defined in the MC simulation. We can see that the unfolding procedure does not improve the electron spectrum estimation compared to the case without unfolding. In fact we can see an oscillation. Additionally, the unfolding procedure exhibits edge effects at low and high energies. We conclude that the CNN method can recover the true energy of the event accurately and with minimal bias. As a result, unfolding is not needed to compute the electron flux. This is important because it avoids edge effects.

4.4.2 MC-data validation

In the previous section, we demonstrated the energy resolution of the CNN correction method compared to classical approaches [110, 149]. This study was conducted using Monte Carlo (MC) simulations. However, to apply this technique in future analyses (Chapters 5 and 6), it is essential to validate these findings with real data.

To achieve this, we will examine beam test data. Since beam test data consists of

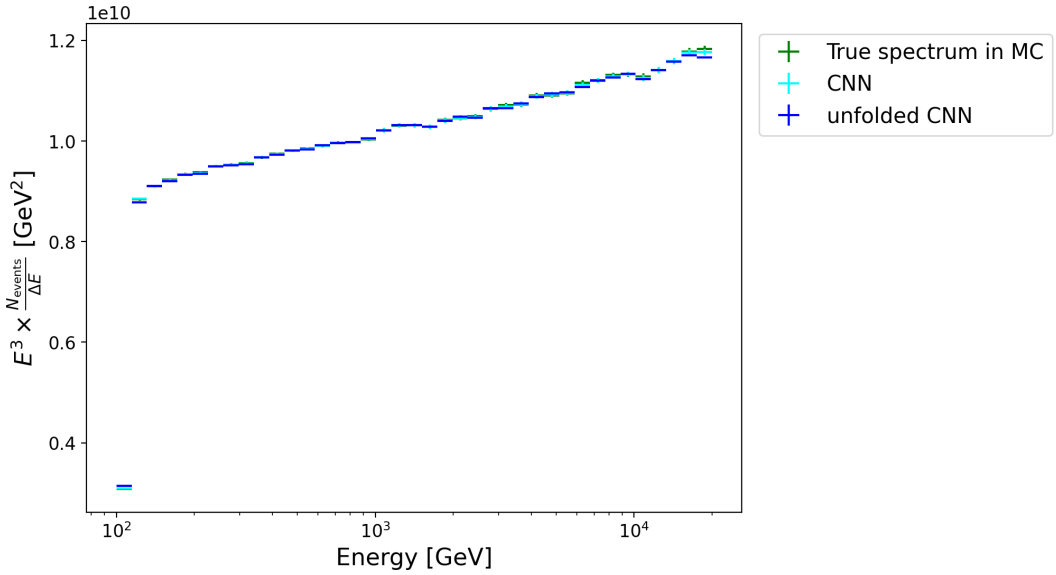


Figure 4.11: Electron kinetic energy spectrum obtained with the CNN method with and without the unfolding procedure (cyan and blue lines respectively) compared with the truth spectrum defined in MC.

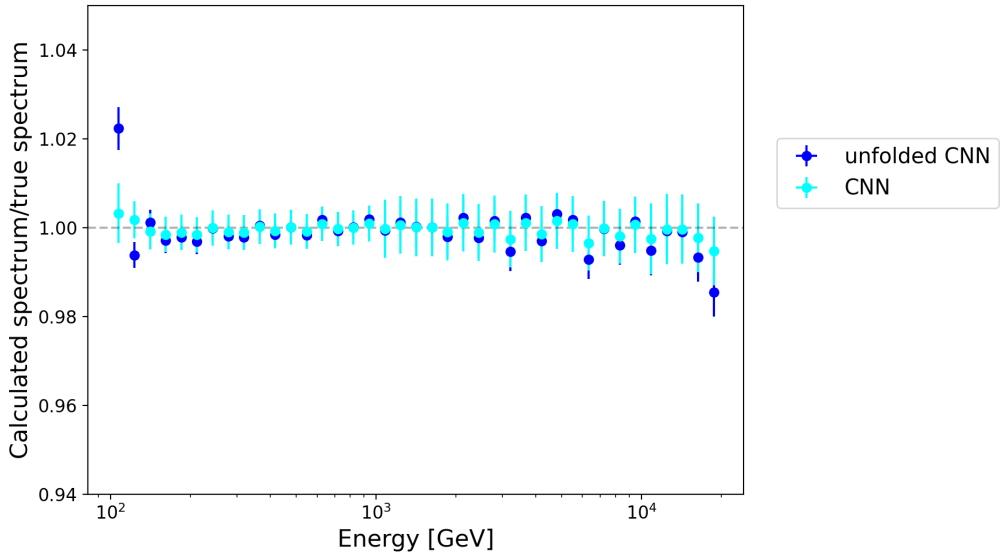


Figure 4.12: Distribution of the reconstructed energy for beam test data. Full line are

fiducial events, it cannot be used exclusively to validate the CNN method; therefore, we also need to incorporate real data from DAMPE. Which will be discus afterwards.

Beam test validation

The use of beam data can validate the CNN reconstruction method proposed here, as it provides access to the initial energy of the particles. This allows for a comparison between classical methods and the CNN-based approach. However, the beam test consists of fiducial events. Consequently, these model estimations will not be a one-to-one comparison but can provide a preliminary evaluation of the model's performance.

The beam test took place at CERN [118, 158], where electron beams with an initial energy of 100 GeV were used. We use MC samples from beam test simulations with a kinetic energy of 100 GeV.

Figure 4.13 shows the energy distribution for each method applied to beam test electrons.

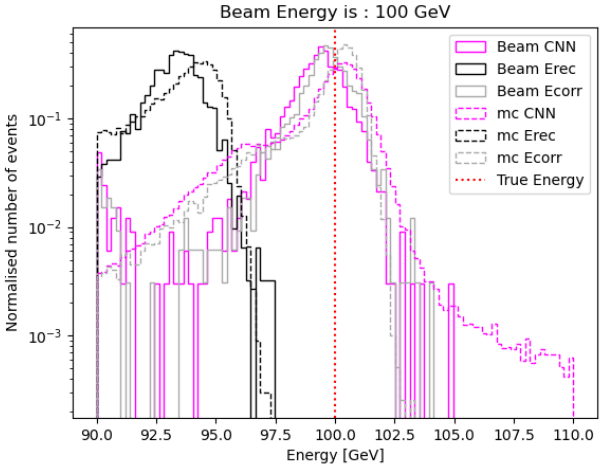


Figure 4.13: Ratio of the calculate spectrum, with and without unfolding, divided by the truth kinetic energy spectrum defined in MC.

Sources	Mean [GeV] (data/MC)	Standard deviation (data/MC)
E_{cnn}	98.98/99.49	1.95/3.69
E_{corr}	99.31/99.36	2./1.78

Table 4.1: Mean and standard deviation for the three energy method for beam test samples of 100 GeV. For beam test data and MC samples.

From Figure 4.13 and Table 4.1, we observe an improvement in the CNN method compared to the reconstructed energy (E_{rec}). However, there is a discrepancy between data and MC for E_{cnn}, with a difference of 0.51 GeV, whereas for E_{corr}, the discrepancy is an order of magnitude smaller.

The overall conclusion is that the CNN method provides a better estimation of the

initial energy compared to the classical method. That being said, in the case of fiducial events, the classical correction proves to be more effective. As mentioned at the beginning of this section, since the model was trained only on non-fiducial events, it is expected to be less efficient in this scenario.

Validation with Flight data

As mentioned earlier, validation with beam test data is not ideal since the CNN model was not trained on events with an almost normal incidence to the detector. To address this, we select a sample of events (both data and MC) within a narrow energy range, using reconstructed energy as a proxy since the kinetic energy is not accessible for real data. The main challenge is that real data consists of a mixture of electrons and protons[1]. To obtain a pure sample of electrons, we must effectively remove protons. However, we cannot rely on the classical discriminator ζ developed for electron-positron analysis [46]. This issue will be examined in more detail in Chapter 5. Instead, we select a clean sample of electrons using a new CNN, specifically trained to distinguish electrons from protons.

The CNN model was trained in the same way as the CNN described in Section 4.3, the exception been that we also added MC simulated protons. As before, the training sample was split into 10 bins equally spaced in logarithm of kinetic energy (from 100 GeV to 24 TeV), then we choose the energy bin with the lowest number of events (this was made between the training electron and proton sample) and select the same number of events in each of the other energy bins. This approach ensures that both species are equally represented, preventing the CNN from favouring one species over another. The final size is 1'025'768 for the training sample and 293'356 for the validation sample. Another difference from the previous CNN is that this time the BGO images were normalised using the reconstructed energy (E_{rec})

We continue using Tensorflow 2 [152] for training the CNN. Figure 4.14 showcases the structure of the model. This time, the model was trained during 100 epochs.

Figures 4.15 and 4.16 show the output of the classifier for 5 energy samples divided in bins of E_{rec}: 200-205 GeV, 300-305 GeV, 400-410 GeV, 460-500 GeV and 960-1000– the last two energy bins are larger due to low statistics. As described in [138], we also removed the sigmoid function from the output layer. The output of the CNN corresponds to a value in the $[-\infty, +\infty]$, where negative values correspond to proton like events and positive values correspond to electron like events. On top of that proton MC were shifted manually to the left by 0.5 to improve the agreement with data. The MC were scaled to data based on the CNN output distribution in the following way:

$$\begin{aligned} s_{\mathrm{p}} &= I_{\mathrm{data}}[-12,-4]/I_{\mathrm{p}}[-12,-4], \\ s_{\mathrm{e}} &= \frac{I_{\mathrm{data}}[2,20] - I_{\mathrm{p}}[2,20]}{I_{\mathrm{e}}[2,20]}, \end{aligned} \tag{4.5}$$

[1] As discussed in Section 4.1, preselection cuts were applied to remove ions, obvious protons and poorly reconstructed events

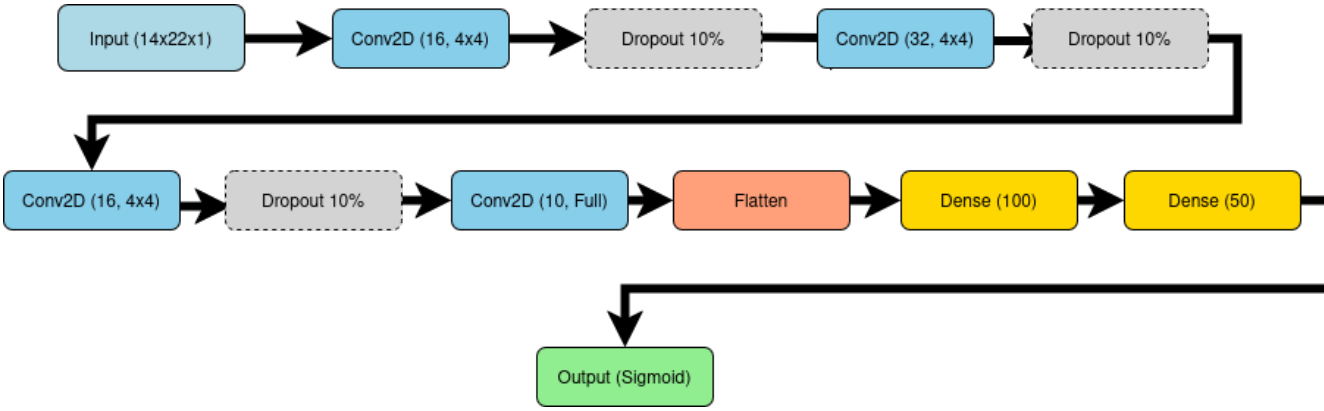


Figure 4.14: CNN model structure.

where s_{p} and s_{e} are the scaling factor for protons and electrons MC, $I_{\mathrm{data}}[-12,-4]$ is the integral of the real data distribution between [-12,-4], $I_{\mathrm{p}}[-12,-4]$ is the integral of the proton distribution, $I_{\mathrm{e}}[2,20]$ is the electron integral in the region [2,20].

We see an overall agreement between MC and data. Some mismatch are observed on the left tail, in the proton region. This is of no problem since the region of interest is locate in the signal region (positive values) where we need to estimate the background fraction.

CNN score of 4 is used to select a clean sample. Table 4.2 reports the number of events that pass this cut. Finally, in Figure 4.17 we plot for each energy range the distribution of energy reconstructed with the CNN method for both real data and MC as shown. Making the cut looser, to 3, results in a worsening of the agreement between data and MC. While a tighter cut, to 5, improves the agreement. Thus, showing that the background contribution plays a role. Additionally, distinct features can be observed in the tails of the distribution. These features arise because, for a given initial energy of the particle, it can deposit only a fraction of its energy, as shown by the response matrix in Figure 4.10. The CNN's ability to more accurately estimate kinetic energy leads to the emergence of additional structures within the tail, corresponding to various kinetic energy values. To have a more quantitative way of showing the comparison between data and MC as shown in Figures 4.18 and 4.19, we looked at the mean value and the 68/95% energy containment of the distributions. The comparison between data and MC is performed using two metrics: the mean value and the 68/95% energy containment. The results show a strong agreement in the mean values between data and MC, indicating consistency in the central tendencies. However, the energy containment demonstrates tighter bounds in the MC compared to the real data. Despite this difference, both data and MC exhibit similar overall behaviour, as illustrated in Figures 4.18 and 4.19, which validate the alignment of the models in capturing key trends.

We can conclude the CNN method improves the accuracy of the estimation of the

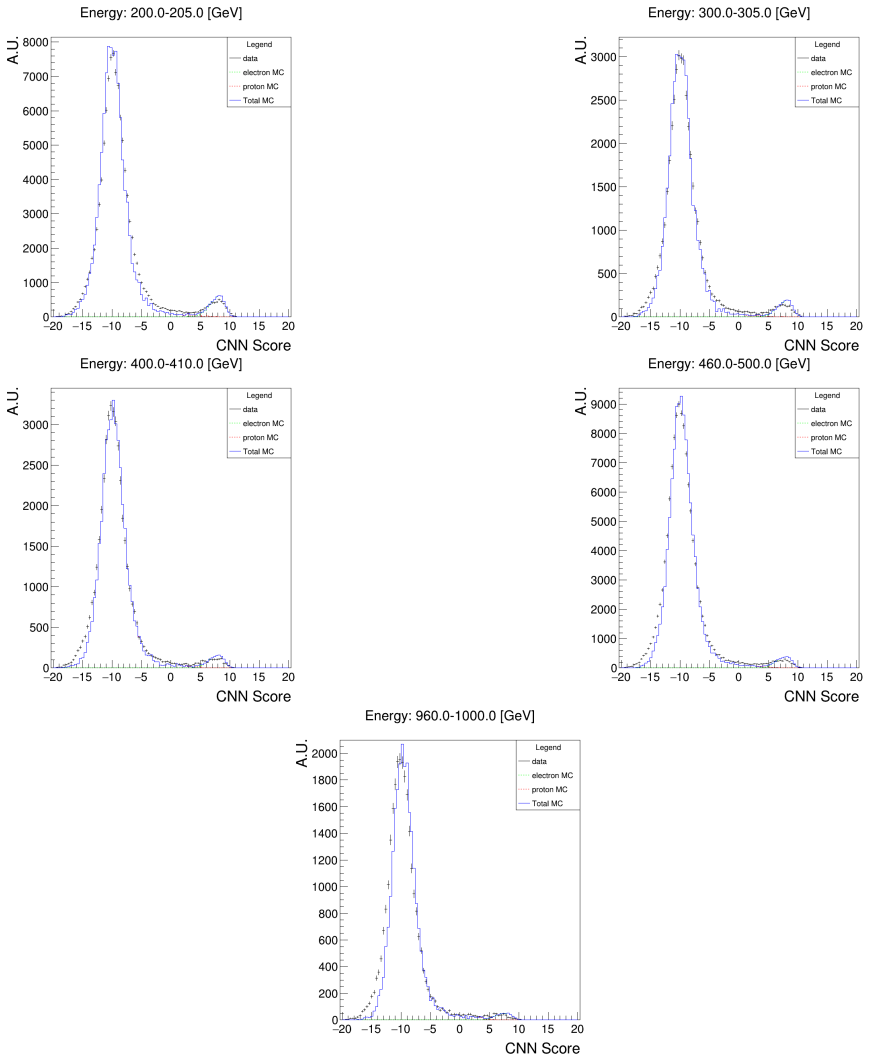


Figure 4.15: Distribution of CNN classifier for data and MC samples for different energy bins in linear scale.

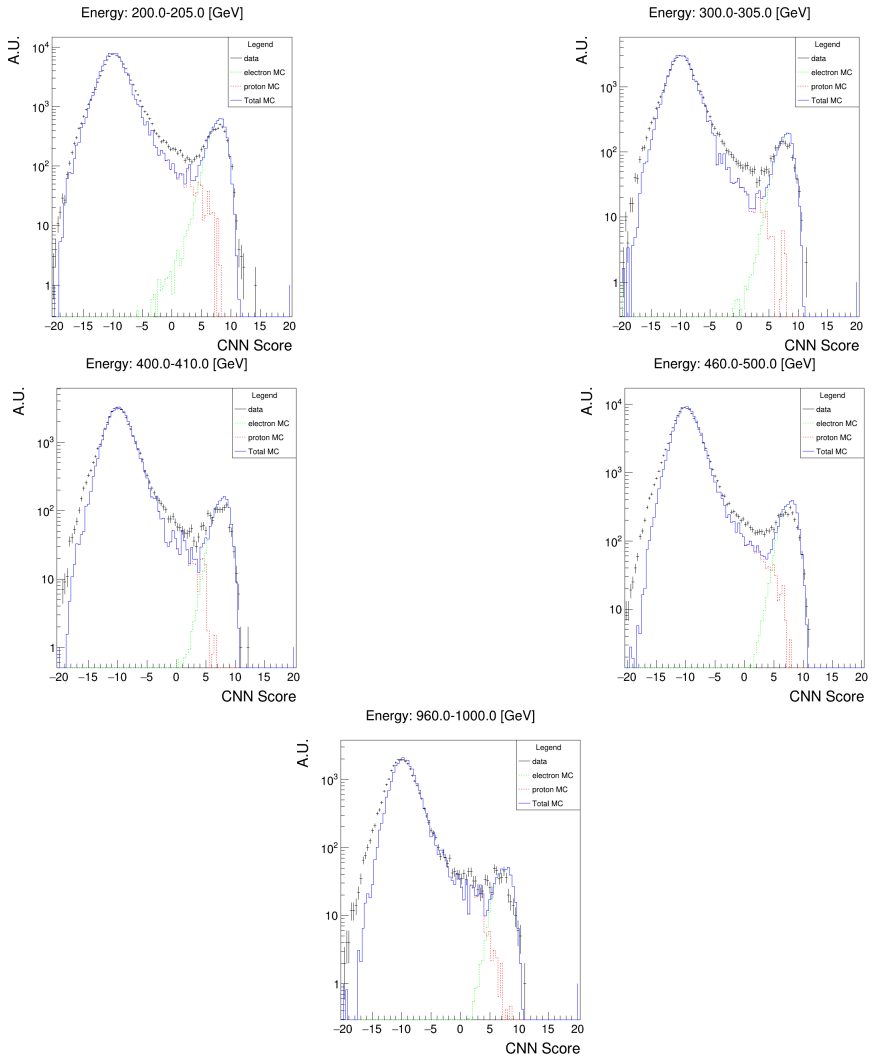


Figure 4.16: Distribution of CNN classifier for data and MC samples for different energy bins in logarithmic scale.

Table 4.2: Number of MC and data events passing the CNN classifier, for a cut value of 4. Last column represents the fraction of background that passes the cut.

Energy range [GeV]	proton MC	electron MC	data	background fraction
200-205	279	4763	4793	5.8%
300-305	63	1508	1429	4.4%
400-410	63	1246	1215	4.5 %
460-500	254	3045	2988	8.5 %
960-1000	49	401	434	11.3 %

initial energy of the electrons compared to the classical methods [110, 149]. We have shown that while there is some differences between MC and data, we can observe a overall agreement.

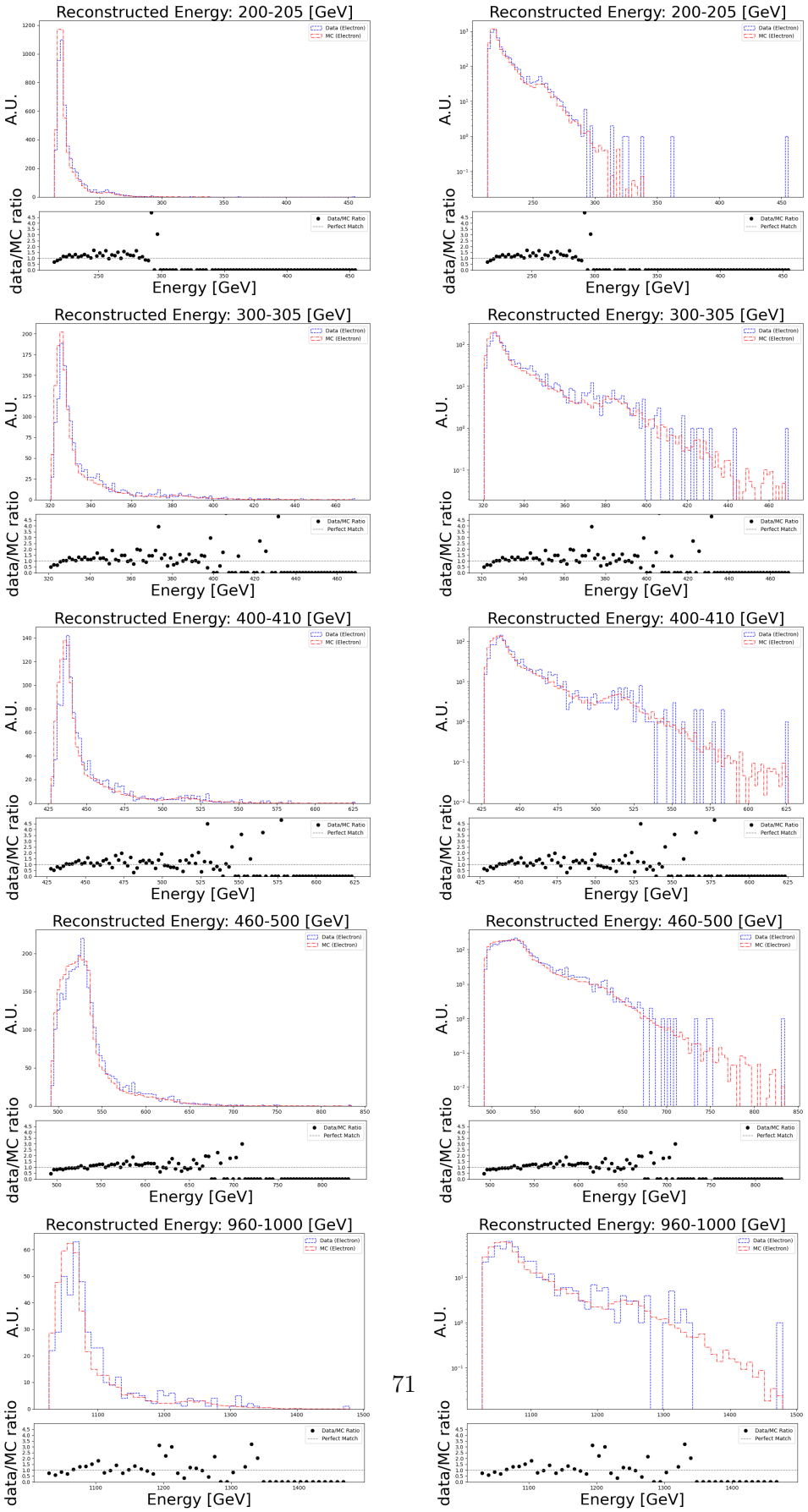


Figure 4.17: Distribution of CNN energy correction method for data and MC samples. Lower plot are the ratio between real and MC data. Left plots are in linear scale, right plots are in log scale.

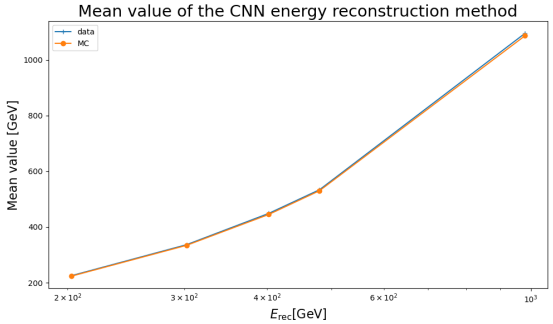


Figure 4.18: Mean value of the CNN reconstructed energy for data and MC as a function of the standard energy reconstruction method [110].

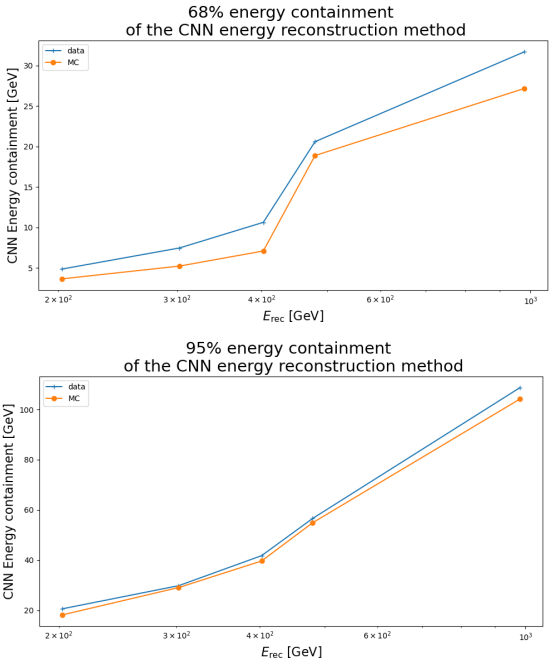


Figure 4.19: The 68% and 95% containment of the CNN energy reconstruction method as a function of standard energy reconstruction method [110].

Chapter 5

Convolutional Method for Identification of Non-fiducial Electrons with the DAMPE Experiment

The work of this chapter is to be submitted to JINST [159]. Some sections have been revised for coherence within this thesis.

CRE are a key component of CRs and could enable to direct detection of CR sources or DM, as discussed in Chapter 1. However, measuring the CRE flux above 5 TeV is particularly challenging due to the steeply falling spectrum. Additionally, electrons constitute less than 1% of the proton population in cosmic rays, making them extremely rare. As a result, distinguishing these events from the proton background is crucial. The BGO calorimeter plays a key role in this process since the PSD and STK detectors can only reconstruct absolute charge values. The ζ variable [46] was used as a discriminator between protons and electrons, based on the shower information from the BGO calorimeter, in order to later obtain a measurement of the CRE flux. Nevertheless, previous studies [137, 138] have employed ML techniques to enhance separation power. However, this methods were applied only to fiducial events, where the shower is almost fully contained within the calorimeter. In contrast, our objective is to increase the available statistics by enlarging the acceptance– specifically, by including events with a larger incidence angles, which result in greater shower leakage outside the BGO calorimeter. Section 5.1 presents the ζ variable [46] and motivation behind the replacement of such method with ML application. Section 5.2 describes the CNN model and training process. In Section 5.3, we present the results and compare the performance of this method with the classical approach.

5.1 A Classical Discriminator for Protons and Electrons

The distinction between electromagnetic and hadronic shower is made at the level of the calorimeter. The work in [46] introduces a variable call ζ, which combines the information of the depth and length of the shower. Electromagnetic showers are dominated by losses through ionization and bremsstrahlung for leptons, hence their shower will be shallow and less wide. On the other hand, hadronic showers will have a more complex interaction with the detector leading to deeper and wider showers that are not fully contained.

The depth and width of a shower are defined as discussed below.

Depth

The depth of the shower, $\mathcal{F}_{\text{last}}$ is the fraction of energy deposited in the last layer with non-zero energy deposition.

$$\mathcal{F}_{\text{last}} = \frac{E_{\text{last}}}{E_{\text{total}}}, \tag{5.1}$$

where E_{last} is the energy deposited in the last layer that saw activity, E_{total} is the total energy deposited in the calorimeter.

Width

To determine the width of the shower we have first to find the track. To do so, we calculate the centre of gravity (COG) for each layer l.

$$COG_l = \frac{E_{imax-1}x_{imax-1} + E_{imax}x_{imax} + E_{imax+1}x_{imax+1}}{E_{imax-1} + E_{imax} + E_{imax+1}}, \tag{5.2}$$

where E_{imax} is the maximum energy deposited in a given layer with the bar been at the position x_{imax}, $E_{imax\pm1}$ is the maximum deposited energy in the previous and next layer which are at the position $x_{imax\pm1}$.

From there we define the root mean square (RMS) for each layer l:

$$\text{RMS}_l = \sqrt{\frac{\sum_{i=1}^{22} E_i(x_i - COG_l)^2}{E_l}}, \tag{5.3}$$

where i correspond to the ith bar of layer l, E_i is the energy deposited in bar i, x_i is the position of the ith bar and E_l is the total energy deposited in the layer l.
By combining the RMS of each layer, we can obtain the information from the width of the shower:

$$\text{sumRMS} = \sum_{l=1}^{14} RMS_l \tag{5.4}$$

As described in [46], we can combine both informations into one variable, ζ:

$$\zeta = \frac{\text{sumRMS}^4 \cdot \mathcal{F}_{last}}{8 \times 10^6} \tag{5.5}$$

Figure 5.1 shows the distribution of ζ for electrons MC, protons MC and real data. We observe that below 1 TeV ζ can discriminate between electrons and protons. As expected, electrons have a narrower shower which results is a small ζ (not greater than 20 mm^4) while proton exhibit a value ranging from a few mm^4 to 100 mm^4. A systematic cut on ζ was applied, removing events with ζ greater than 100. At those values only protons are found. Starting at 1 TeV, the separation becomes nearly impossible. This can be attributed to the increased leakage of the shower compared to fiducial events, which results in a smaller estimation of ζ and causes the fusion of both distributions.

5.2 Convolutional Neural Network

Figure 5.1 shows that, due to shower leakage, the non-fiducial events appears shorter and less extended from the perspective of the *zeta* variable. As a consequence, a more advanced method is needed to distinguish electrons and protons. In this section, we will discuss the development of a CNN discriminator to separate non-fiducial (see section 4.1) electrons from protons.

As discussed in section 3.2, CNNs can be used for image processing for event classification. In this work the CNN input are BGO images (14×22 pixels), where each pixel value correspond to the energy deposited. Contrary to the model for energy reconstruction task, we normalise the images by the reconstructed energy. Moreover, this time we not only use electrons MC but also protons MC, since our objective is to train CNN classifier to separate non-fiducial electrons and non-fiducial protons. The MC were processed following the preselection described in Section 4.1 such that only non-fiducial events are selected. For this work, two models are trained: one for events with energy below 1 TeV and another for events with energy from 1 TeV to 24 TeV. To do so, we separates MC samples based on their CNN reconstructed method energy developed in Chapter 4. Then we split the resulting sample into a training sample with 70% and a validation sample corresponding to 30% of the total MC sample. Contrary to the previous model trained for the energy reconstruction, here we do not balance the samples such that we could have enough large statistics for training the model. While we attempted to balance the number of electrons and protons in the samples, this approach reduces the overall statistics available for training and validation. For the low energy (LE) model, the training sample has a total of 2'339'937 events (electrons: 1'440'939 and protons: 898'998) and the validation sample has 1'002'832 events (electrons: 617'546 and protons: 385'286). At high energy (HE), the training sample consists of a total of 1'615'015 events (electrons: 663'948 and protons: 951'067) and the validation sample contains 692'151 events (electrons: 284'550 and protons: 407'601).

The model structure is trained using Tensorflow 2.0 [152], it's structure is the following: The CNN architecture was adapted from internal work carried out as part of

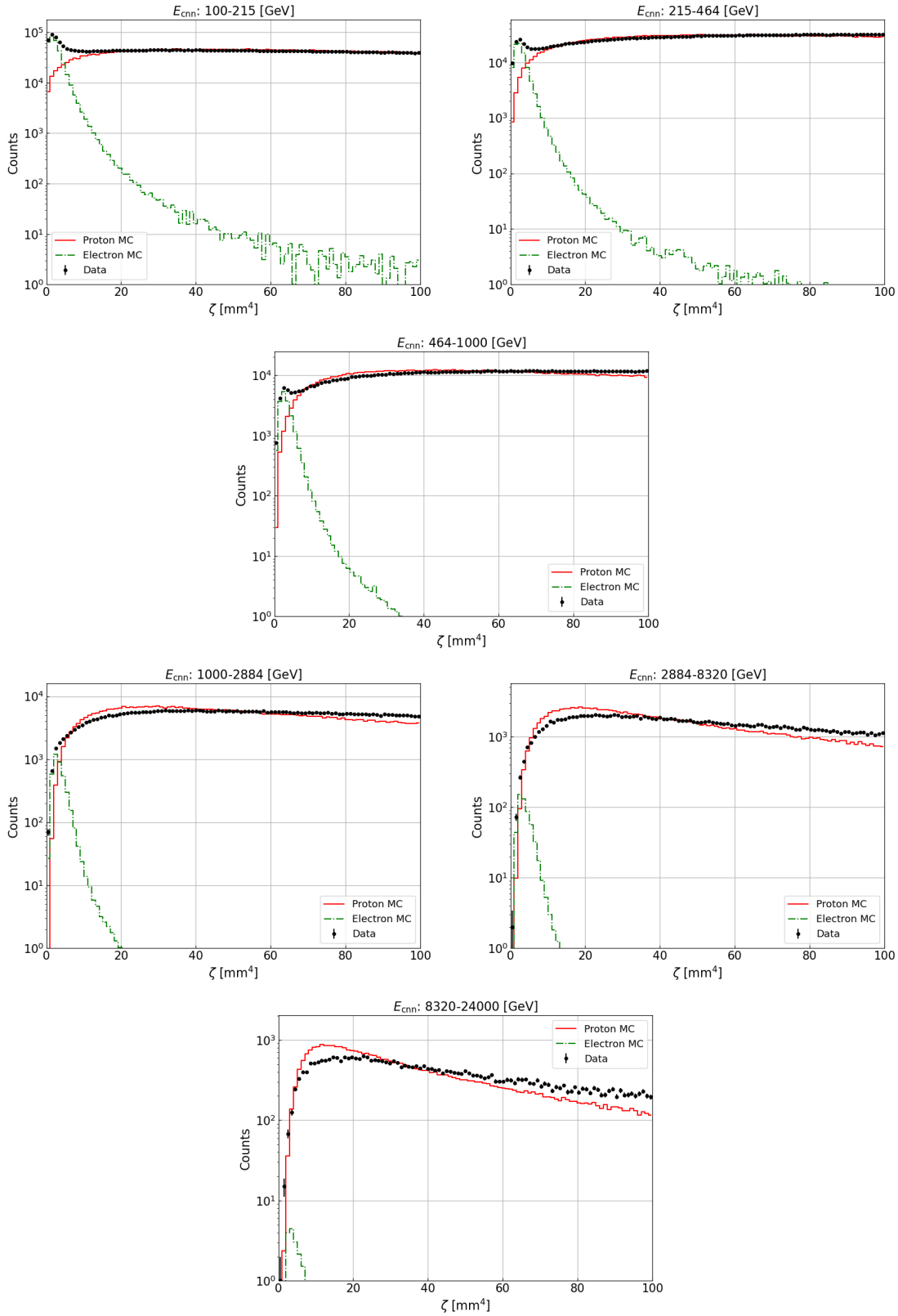


Figure 5.1: ζ distribution for non-fiducial events. Real data (black dots), electrons MC (green line) and protons MC (red line) for 5 energy bins from 100 GeV to 24 TeV.

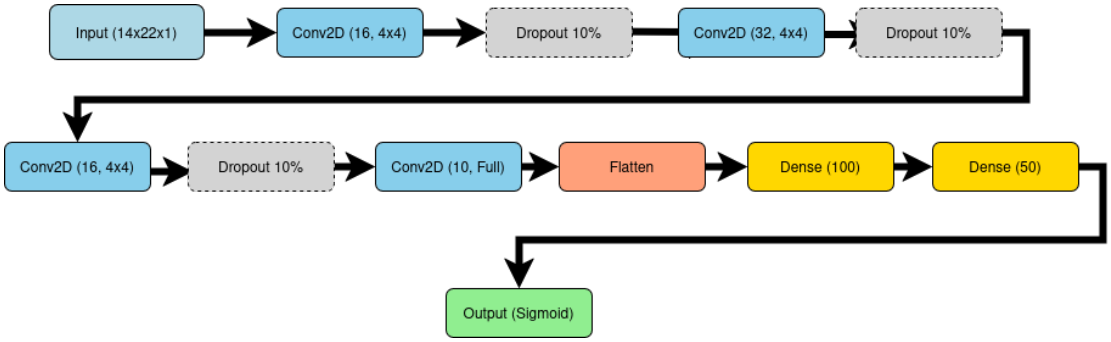


Figure 5.2: CNN structure use for the low energy and high energy model.

a previous study for the discrimination of fiducial electrons and protons. Each convolutional layer (Conv2D) employs a kernel size of 4×44×4, except for the final layer, which matches the output size of the preceding layer. Dropout is applied after every convolutional layer except the last one to prevent overfitting. The output of the final convolutional layer is flattened and passed through two fully connected layers with 100 and 50 neurons, respectively, followed by the output layer. To monitor the training process and detect potential under or over-fitting, the binary cross-entropy loss is tracked across epochs for the training and validation samples. Following the approach outlined in [138], the sigmoid function was omitted from the output neuron.

We trained both models for 200 epochs, Figure 5.3 shows the evolution of the binary cross-entropy for the train and validation samples. From which, we can see a stable loss for both and no evidence of under- and over-fitting.

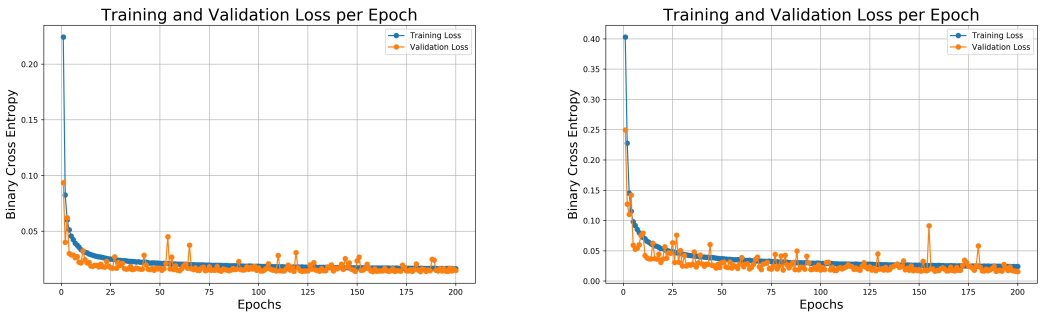


Figure 5.3: Evolution of the loss, defined as the binary-cross entropy for the training sample (blue curve) and validation sample (orange curve). Left plot: Model trained below 1 TeV. Right plot: Model trained above 1 TeV.

5.3 Results

Our objective was to develop a new discriminator to separate electrons from protons. To achieve this, we aim to surpass the current performances of the classical method with the variable ζ [46]. Since each discriminator operates on a different scale and requires a different cut value, a direct one-to-one comparison is not possible. To objectively evaluate both techniques, we use the Receiver Operating Characteristic (ROC) curve, which compares the background rejection as a function of signal efficiency.

The signal efficiency is defined as follows:

$$e_{signal} = \frac{N_{\text{signal,pass}}}{N_{signal}}, \tag{5.6}$$

and the background fraction is defined as:

$$e_{background} = \frac{N_{\text{background,pass}}}{N_{\text{background}}}, \tag{5.7}$$

where $N_{\text{signal,pass}}$ is the number of signal events that pass a discriminator cut, N_{signal} is the total number of MC electron events, $N_{\text{background,pass}}$ is the number of background events that pass the discriminator cut and $N_{\text{background}}$ is the total number of background events. The statistical error for signal efficiency and background are defined as follows:

$$\Delta e_{signal} = e_{signal} \times \sqrt{(\frac{\Delta N_{\text{signal,pass}}}{N_{\text{signal,pass}}})^2 + (\frac{\Delta N_{signal}}{N_{signal}})^2}, \tag{5.8}$$

$$\Delta e_{background} = e_{background} \times \sqrt{(\frac{\Delta N_{\text{background,pass}}}{N_{\text{background,pass}}})^2 + (\frac{\Delta N_{background}}{N_{background}})^2}, \tag{5.9}$$

Where $\Delta N_{\text{signal,pass}}$ is the statistical error of non-fiducial electrons that pass the discriminator cut, $\Delta N_{\text{signal,pass}}$ is the statistical error of total non-fiducial electrons, $\Delta N_{\text{background,pass}}$ is the statistical error of non-fiducial protons that pass the discriminator cut, $\Delta N_{\text{background,pass}}$ is the statistical error of total non-fiducial protons.

Figures 5.4 and 5.5 display the ROC curves, we can see that the CNNs perform significantly better, achieving a lower background fraction at the same signal efficiency. This represents an improvement of approximately one order of magnitude compared to the classical discriminator. To quantify this improvement as a function of the energy, we analyse the background fraction at fixed signal efficiencies of 85% and 90%. Figure 5.6 shows the evolution of the background fraction for the LE and HE models, as well as for the ζ variable, at signal efficiency of 85%. For both CNN models, we confirm the observations from the ROC curves, that the CNNs have an improved separation power. LE model demonstrates improved performances at all energies, with more than one order of magnitude better discrimination compared to the classical method. In the HE model, performance begins to degrade above 3 TeV, a trend also observed for ζ,

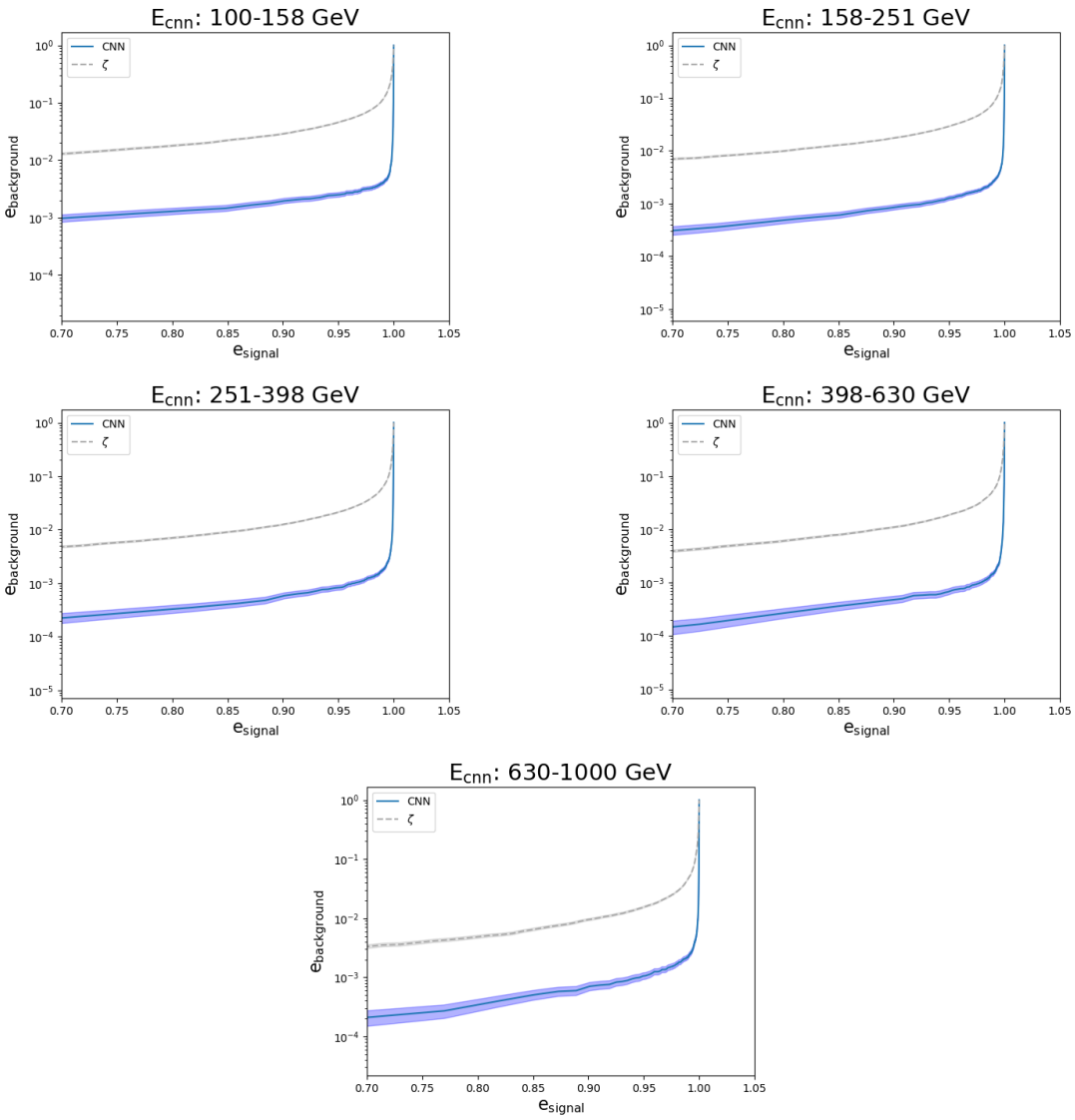


Figure 5.4: Comparison of ROC curves on MC for the ζ variable (grey curve) and the LE CNN classifier (blue curve). A lower curve indicates better performance. Shade bands represent statistical uncertainties.

though the degradation appears less rapid for the CNN. Nevertheless, the CNN maintains a background fraction lower than that of the classical discriminator, remaining below 10^{-3}. Figure 5.7 presents the same results for a fixed signal efficiency of 90%. The overall trend is similar to the 85% case, albeit with a higher background fraction.

MC studies of the CNN models have demonstrated improved discrimination power compared to the classical method [46]. While efficiency studies solely on MC help assess the quality of the model, it is important to see how the CNNs perform on real data.

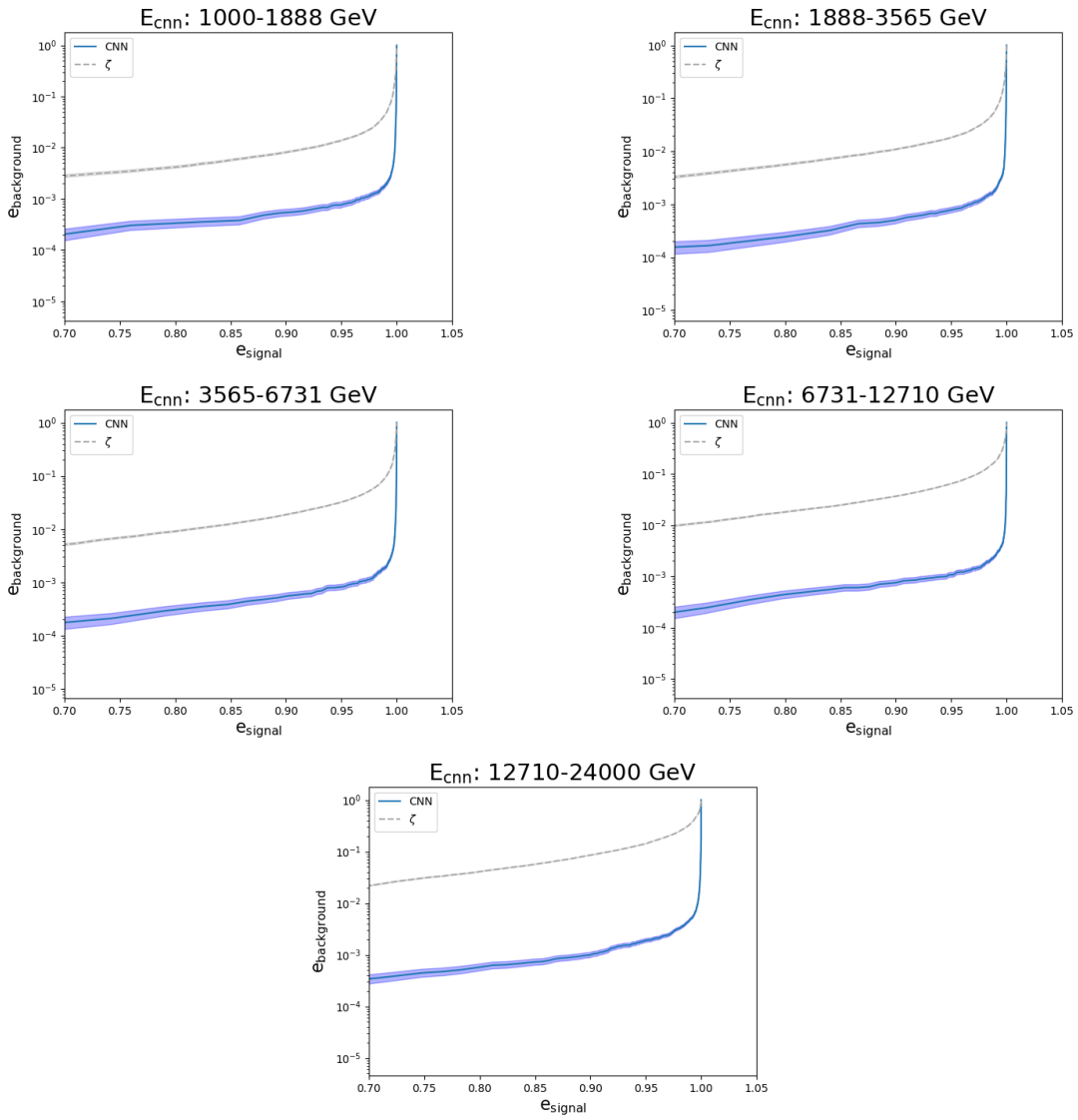


Figure 5.5: Comparison of ROC curves on MC for the ζ variable (grey curve) and the HE CNN classifier (blue curve). A lower curve indicates better performance. Shade bands represent statistical uncertainties.

To this end, we compare the shapes of the output distributions for real data and MC. We first scale the electron and proton MC by fitting the electron fraction f using the following expression:

$$N_{\text{data}} = (1 - f)N_{\text{proton}} + f \cdot N_{\text{electrons}}, \tag{5.10}$$

where N_{data} is the number of data events, N_{proton} is the normalized number of proton MC events, and $N_{\text{electrons}}$ is the normalized number of electron MC events.

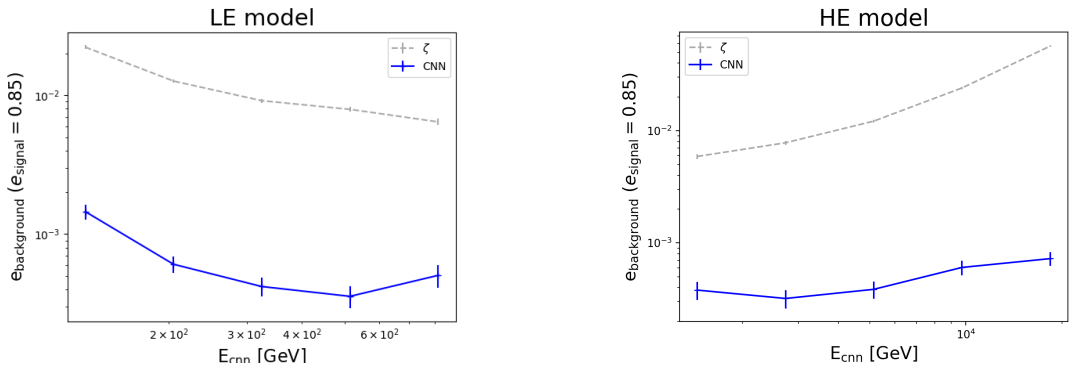


Figure 5.6: Comparison of ROC curves on MC for the ζ variable (grey curve) and the HE CNN classifier (blue curve). A lower curve indicates better performance.Shade bands represent statistical uncertainties.

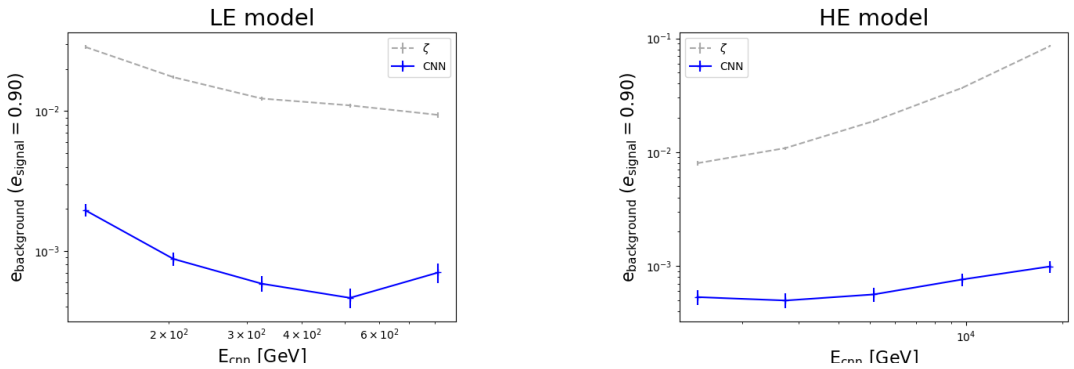


Figure 5.7: Background fraction as a function of the energyζ variable (grey curve) and the HE CNN classifier (blue curve). A lower curve indicates better performance. Shade bands represent statistical uncertainties.

Figures 5.8 to 5.11 present the CNN output distributions (for both low and high energy models) for data and MC, after applying the scaling. We a discrepancy appears between the proton MC and data, particularly in the proton bulk and left tail. Despite this mismatch, there is good agreement in the proton's right tail and the electron region. Importantly, this discrepancy does not impact future CRE analyses, as only the region near the signal–CNN value of 4– is relevant for background estimation. That being said, such discrepancy should be taken into account as an uncertainty.

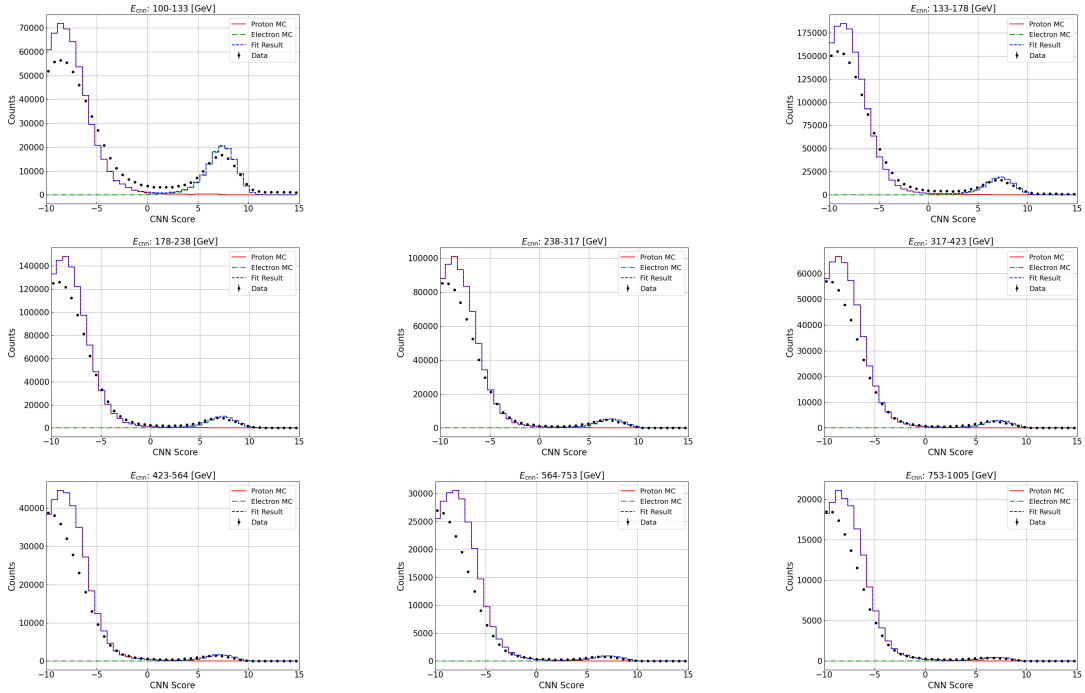


Figure 5.8: CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are real data. Energies below 1 TeV. Linear scale.

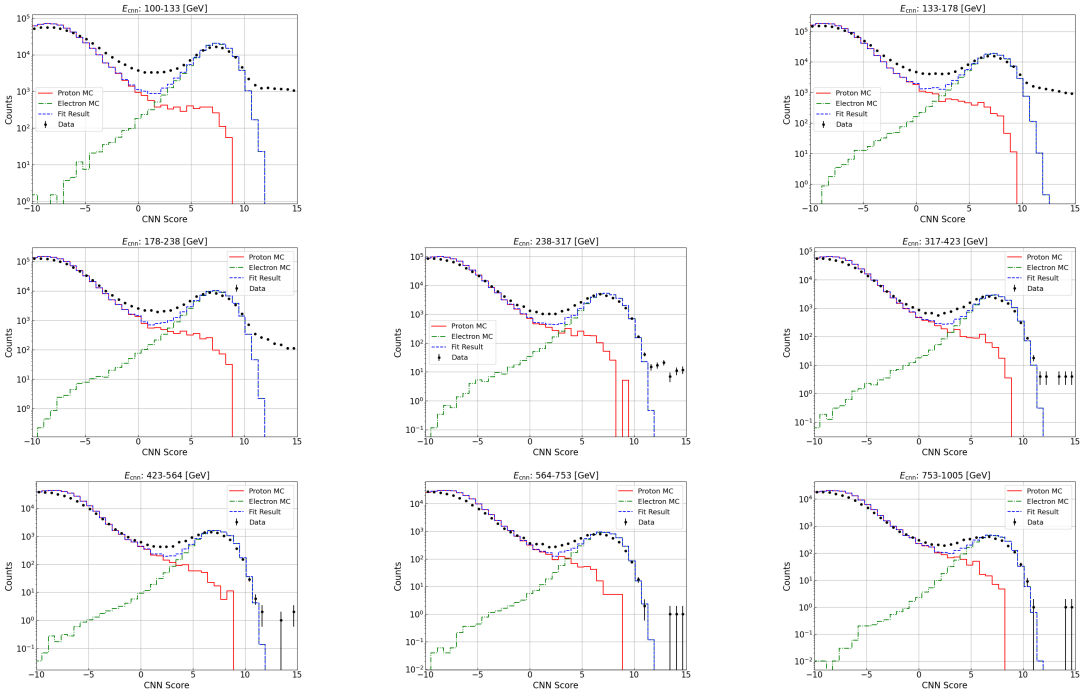


Figure 5.9: CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are real data. Energies below 1 TeV. Logarithmic scale.

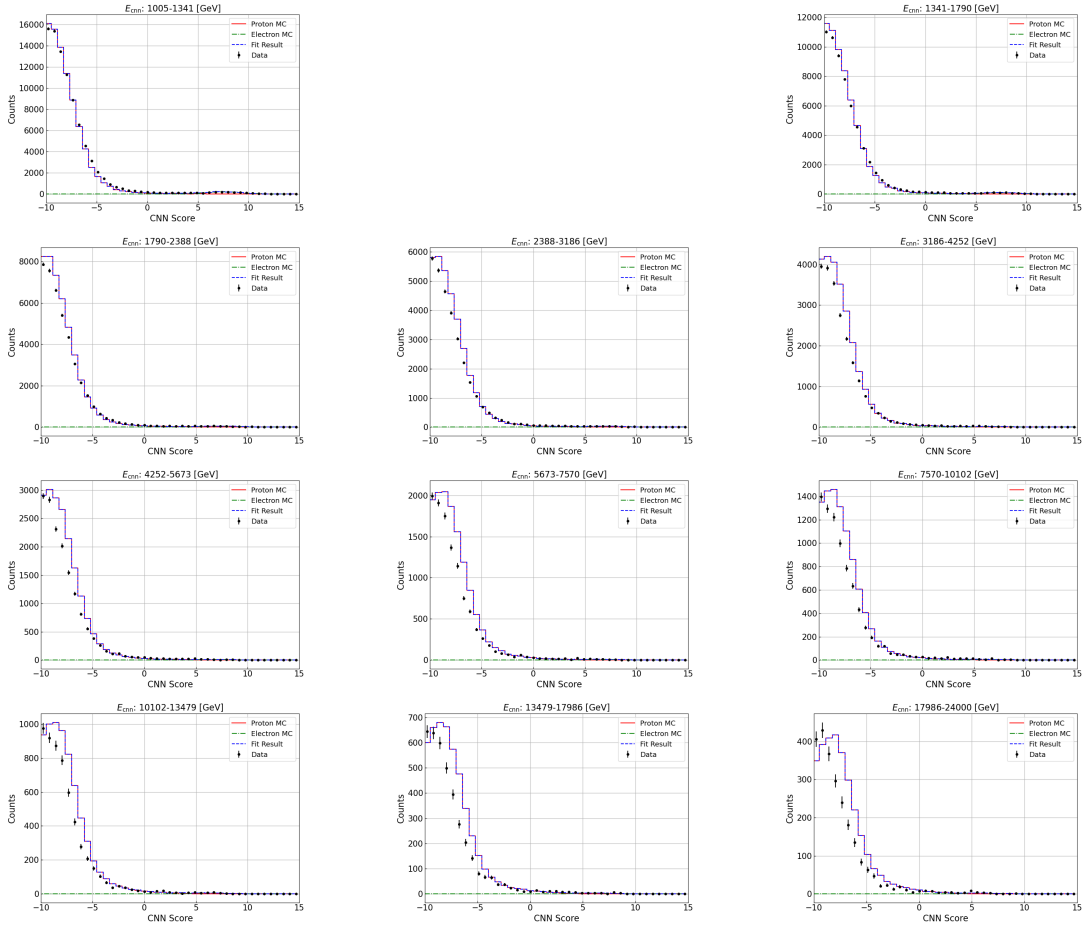


Figure 5.10: CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are real data. Energies above 1 TeV. Linear scale.

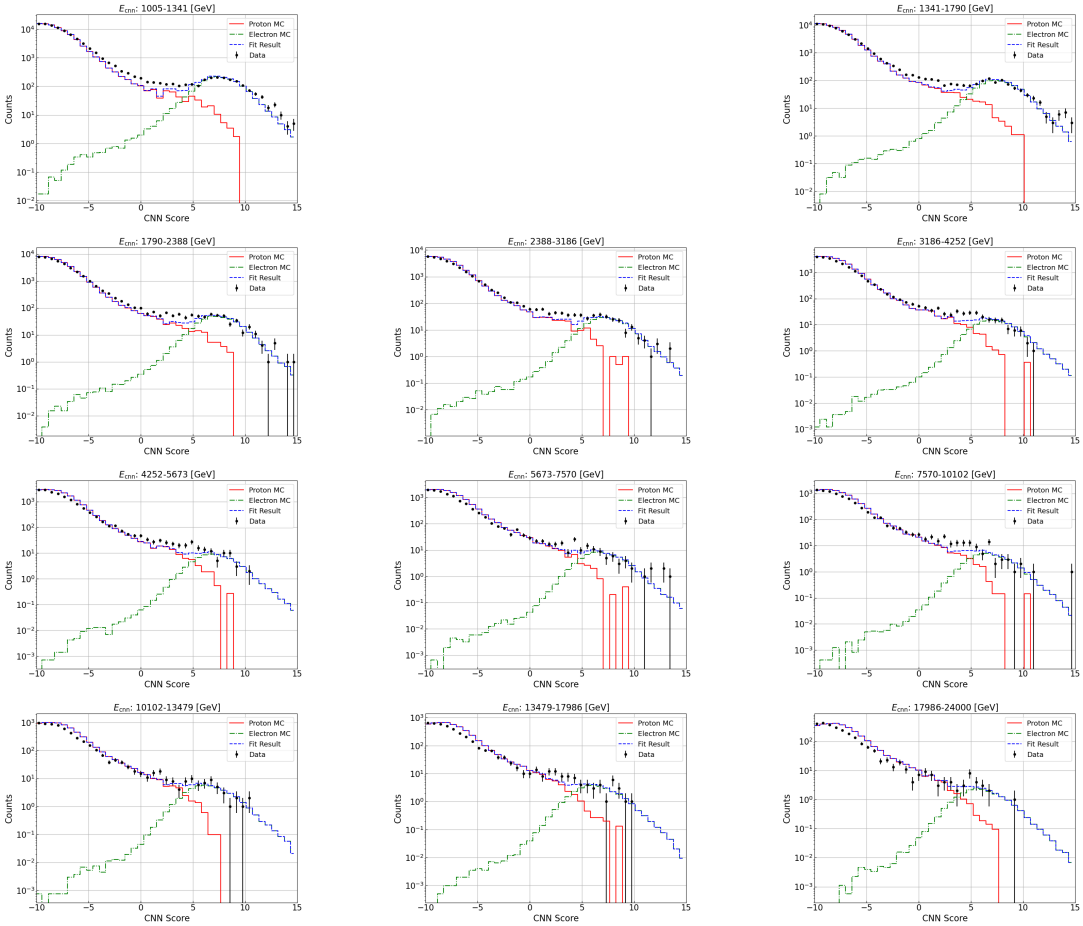


Figure 5.11: CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are real data. High energy model. Log scale.

Chapter 6

Measurement of the Non-Fiducial Electron+Positron Flux

In Chapter 1 we discussed the possible contribution of local sources to the CRE flux at the TeV energy scale. However, to be able to confirm or contradict such models, it is important to measure the electron flux at these energies. DAMPE has been tackling such a task since its launch in 2015, with the aim to measure the CRE flux up to 10 TeV. At high energies, CRE undergo energy losses during propagation due to their low mass, primarily from processes such as synchrotron radiation and inverse Compton scattering [75]. Consequently, it is anticipated that electrons at energies higher than a few TeV originate from nearby sources, i.e. within a distance of about one kiloparsec, not older than about 10^5 years [2, 146]. In addition, the aforementioned energy losses lead to a steeper energy spectrum, which results in a lower CRE flux compared to other species, making them relatively rare at high energies. As mentioned in Chapters 4 and 5, in order to increase the statistics, we opened the acceptance by selecting events outside the fiducial volume. Moreover, in Chapter 4 we developed a CNN to improve the estimation of the initial energy of the particle. On top of that, in Chapter 5 we constructed a CNN to separate electrons from protons. Such tools can be used to measure the CRE flux up to 3 TeV.

The differential flux of a CR particle in a given energy range $[E, E + dE]$ is defined as:

$$\phi(E, E + dE) = \frac{N_{\text{signal}}(E, E + dE)}{\Delta T \cdot A \cdot dE}, \tag{6.1}$$

where:

- $N_{\text{signal}}(E, E + dE)$ is the number of particles observed by the experiment in the energy range $[E, E + dE]$,
- ΔT is the life time of the experiment, corresponding to the time during which DAMPE has collected data,

- A is the geometrical acceptance corrected by the efficiency,
- dE is the width of the considered energy range.

This formula can be extended for the electron flux measurement by adding a term to take into account the efficiency ϵ. Another element to take into account is the background contribution $N_{background}(E, E+dE)$, since no cut is 100% efficient, that will pass the selection. This background can be estimated and subtracted from the total number of candidates $N(E, E+dE)$. Therefore the formula becomes:

$$\phi(E, E+dE) = \frac{N_{total}(E, E+dE) - N_{background}(E, E+dE)}{\Delta T \cdot A_0 \cdot \epsilon \cdot dE}, \tag{6.2}$$

where N_{total} is the number of particles that pass all cuts–including background, A_0 is the geometrical acceptance about 0.31 $m^2 \cdot sr$ at all energies. In the following section we will estimate each of this terms and finally discuss the resulting flux.

6.1 Data and Livetime of DAMPE

6.1.1 Flight data

DAMPE has collected data since the 30th of December 2015. In this analysis only data up to the 31st of December 2023 will be included, resulting in a total of more than 16 billions events with an average of 5 million events per day. Figure 6.1 shows the total number of events collected per day by DAMPE. A drop in the daily counts can be observed starting in mid 2022, this effect is due to the increase in solar activity which results in a stronger magnetic field and therefore a shielding against particles coming from outside the solar system.

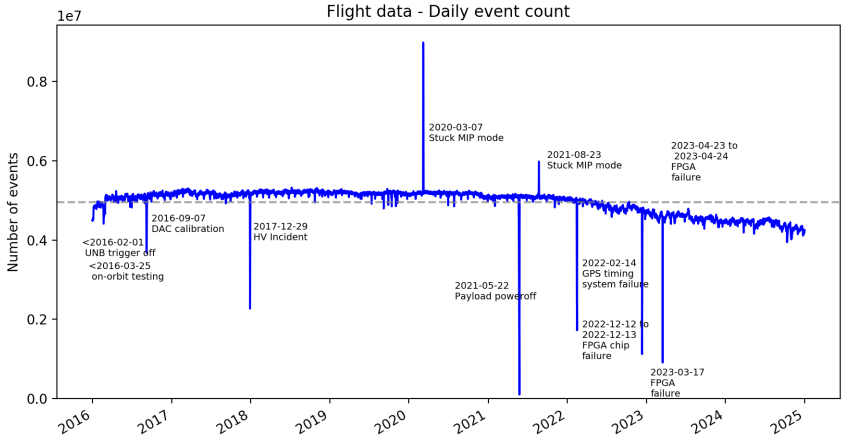


Figure 6.1: DAMPE daily event counts. The horizontal line shows the average number of events taken per day.

Here we list the excluded days from the analysis:

- 2016-02-22 to 2016-02-24: DAMPE operated on in-orbit testing stage
- 2016-09-07 to 2016-09-08: Series of fine calibrations occupying a lot of observation time.
- 2017-12-29 and 2017-12-30: High voltage accident resulting in loss of data.
- 2020-03-07 to 2020-03-08: Issues with the operation mode: DAMPE was stuck in MIP mode, resulting in a very high number of events. MIP mode correspond to the activation of the MIP trigger (see Section 2.2), such mode is used for calibration of the detectors.
- 2021-05-22 to 2021-05-25: Satellite battery fault.
- 2021-08-23 : DAMPE was stuck in MIP mode.
- 2022-02-14 and 2022-02-15: GPS failure.
- 2022-12-12 and 2022-12-13: FPGA failure resulting in loss of data.
- 2023-03-17 : FPGA failure.
- 2023-04-23 and 2023-04-24: loss of housekeeping data due to FPGA failure. This results in the impossibility of generating the files used in all analysis.

Moreover, for the first two months the unbiased trigger was not enabled, resulting in a lower number of events.

As described in Section 4.1, we apply the same set of cuts. Only $\sim 5\%$ pass the selection. Figure 6.2 shows the number of skim events for different energy ranges.

On top of that we remove events when DAMPE passed through the SAA [22].

Livetime

DAMPE has been in operation since the end of 2015. However, data acquisition is not continuous. DAMPE records data for an average of 18.5 hours per day. The rest corresponds to the detector dead time, $\approx$ 3ms after the acquisition of an event (it results in 3 hours per day). The livetime for DAMPE is calculated using an SQL database at the Purple Mountain Observatory in China. In the work presented here, the total livetime is $217'244'668$ seconds, corresponding to 2514 days.

6.1.2 MonteCarlo Simulations

To understand the interaction of incoming particles with DAMPE, it is necessary to use simulations as described in Section 2.3. In this work we use proton and electron simulations, we use the DAMPE software and Geant4, with the FTFP_BERT physics list. The event count amounts to: 5058361 protons MC and

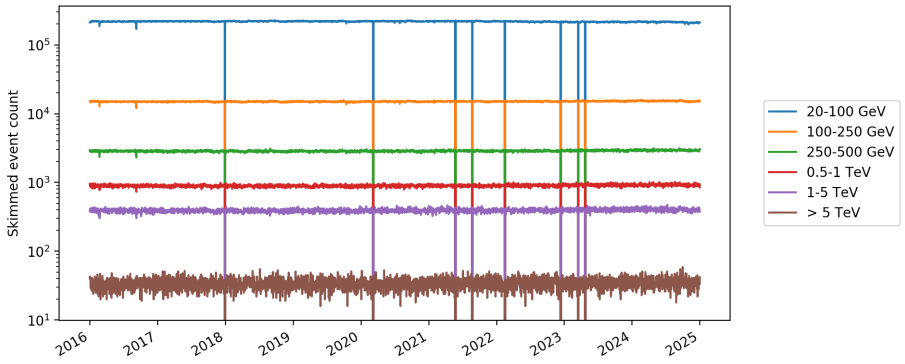


Figure 6.2: Daily event count that passes the selection described in Section 4.1 from DAMPE. Blue line shows events with energies in the range [20,100] GeV, orange line events with energy within [100,250] GeV, green line for energies in the range [250,500] GeV, red line for energies in the range [0.5,1] TeV, purple line for [1,5] TeV and brown line for events with energies above 5 TeV.

Each of these energy bins was simulated following a power law with an E^{-1} spectrum. Then the events were given a weight following a E^{-3} spectrum for electrons and $E^{-2.7}$ for protons. The main reason been that electron flux follows a E^{-3} power law spectrum and proton flux follows a $E^{-2.7}$ power law.

6.2 Event Selection

As described in section 4.1 for which we select events based on the fiducial selection, we also apply the HET. On top of that we apply the following cuts:

- Using ML track reconstruction technique [139] to confirm that the shower direction is within a distance of 280 mm from the centre of the calorimeter,
- inBGOlength: Energy dependent cut on the track length in BGO,
- Energy dependent cut on the Lvalue (fitted shower max position with respect to inBGOlength),
- PSD ions: cut on the PSD charge to remove ions,
- STK ion: cut on the STK charge to remove heavy ions,
- STK helium: cut on the STK charge to remove helium,
- Energy fractions: an energy-dependent cut is applied based on the ratios of energy deposited in the BGO calorimeter layers (R01, R02)—where R01 is the ratio between the first and second layers, and R02 is the ratio between the first and third

layers. The cut also uses the energy fractions deposited in specific BGO layers: F0 and F1 correspond to the first and second layers, respectively; combined fractions include $F01 = F0 + F1$ (sum of the first and second layers), $F23 = F2 + F3$ (sum of the third and fourth layers), and $\mathcal{F}_{\text{last}}$,

- $sumRMS$ and cosTheta: energy dependent cut on sum of the RMS of the BGO layers and the cos of the angle between the BGO track reconstructed using [139], and the correlation between $sumRMS$ and cosTheta,

The event selection efficiency, shown in Figure 6.3, is above 94% below 1 TeV and decreases towards the high energy to 89%. These different cuts have the main objective of removing ions

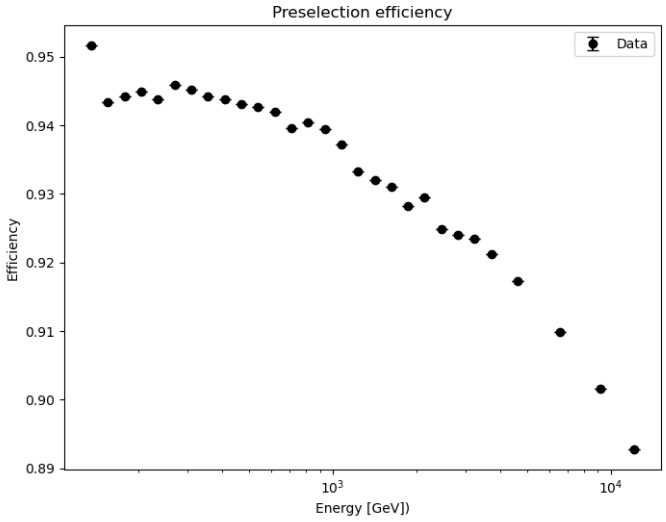


Figure 6.3: Preselection efficiency as a function of the energy.

6.3 Electron Identification and Background Estimation

The final selection of the electrons from the data is based on the CNN developed in Chapter 5. Two models were trained on a mixed sample of electrons and protons, as described in section 5.1. The first model is used to separate events with energies below 1 TeV, while the second model is used for events with energies of 1 TeV and above. The reliability of the model was first assessed by calculating the ROC curves, as shown in Figure 5.5. We then evaluated the evolution of the background at fixed signal efficiency (85% and 90%), which resulted in an overall improvement compared to ζ. Finally, we examined the agreement between data and MC simulations, as illustrated in Figures 5.8 to 5.11.

As described in Chapter 5 the MC samples were scaled to data in the following Eq 5.10.

After scaling MC to data, we select a CNN score of 4 as the final cut, such that below such value it is considered background. This distributions are shown in Figures 6.4 to 6.7 displaying the distribution of the CNN score for MC and data. MC events were scaled to data by fitting the electron fraction defined in Eq 5.10. As commented in section 5.3, at low energy (below 1 TeV) a discrepancy is observed between the proton MC and data, primarily in the proton bulk and left tail. Despite this mismatch, there is good agreement in the proton's right tail and the electron region, but it results in an uncertainty of the estimation of the proton background, which will be taken into account in the systematic uncertainties.

For the high energy distributions, we can see that agreement between MC and data is better. However, we do observe that the separation for the last 5 energy bins becomes worse, where the statistics are also limited.
Moreover, from the distribution plots (Figures 6.4 to 6.7) it is clear that the proton background is present in the signal region and thus needs to be estimated. A simple way of estimating such background is to count the relative ratio of background in the signal region:

$$F_{\text{background}} = \frac{N(\text{CNN} > cut)_{\text{proton}}}{N^{tot}_{\text{proton}}}, \tag{6.3}$$

Figure 6.8 shows the fraction of background events. From it we can see that below 1 TeV the background fraction reaches up to $\sim 2\%$ and above 1 TeV it goes from $\sim 2\%$ up to $\sim 10\%$.

Figure 6.9 shows the efficiency of the CNN cut. Efficiency of the CNN selection after applying all analysis cuts except the CNN one ($n-1$ selection). As we can see, the preselection efficiency is above 90%. The CNN efficiency is different below and above 1 TeV, since we use two CNN models. Its efficiency is above 94% below 1 TeV, and above 1 TeV the efficiency stays above 88%. This behaviour shows the good separation power of the model.

6.4 Flux Calculation

Now that we have all the ingredients, we can use Eq. 6.2 to calculate the flux from 100 GeV up to 3 TeV. For each bin, we consider the event count and background estimation, both of which are based on the CNN, as well as the effective acceptance and the livetime.

Figure 6.10 displays the statistical uncertainties relate to the data, background, CNN cut and acceptance.

First, we verify that the new results are in agreement with the previous analysis, as shown in Figure 6.11. We observe that while the new flux is below previous published work [46] below 1 TeV, both are fluxes are within uncertainties. For energies above 1 TeV, the agreement worsens. This is likely due to an underestimation of the background, indicating that a cut-and-count method for background estimation is not well suited for this

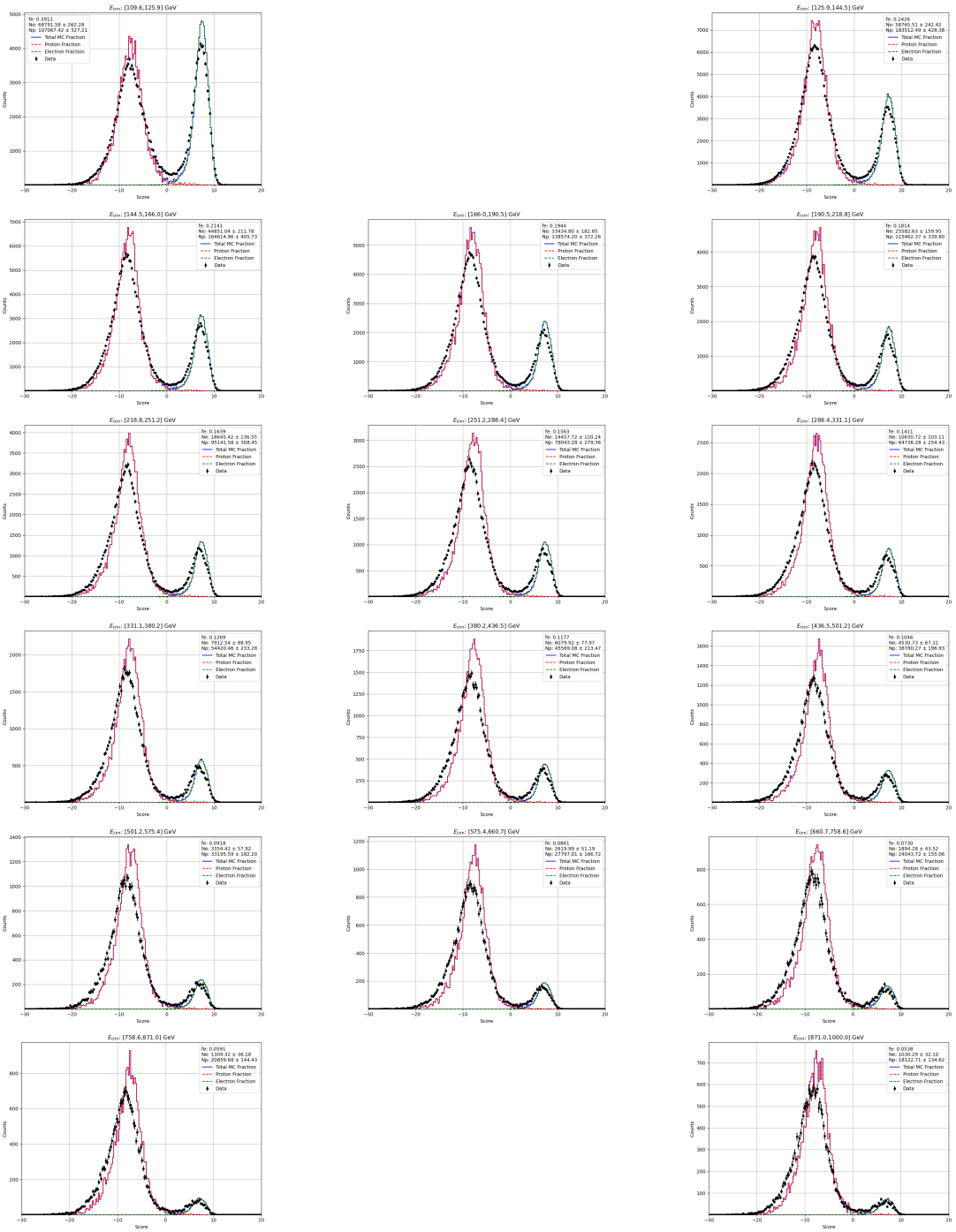


Figure 6.4: CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are data. Energies below 1 TeV. Linear scale.

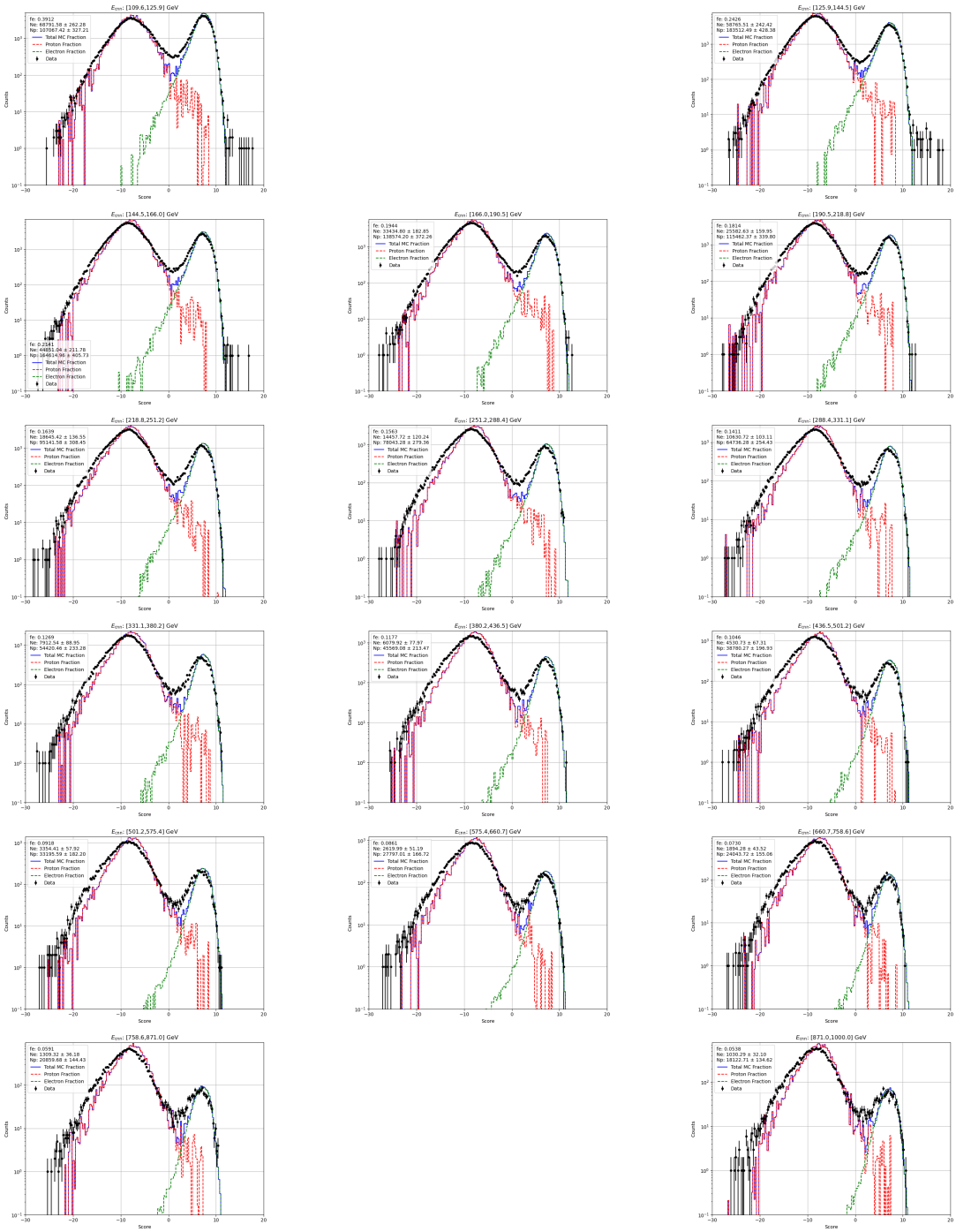


Figure 6.5: CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are data. Energies below 1 TeV. Log scale.

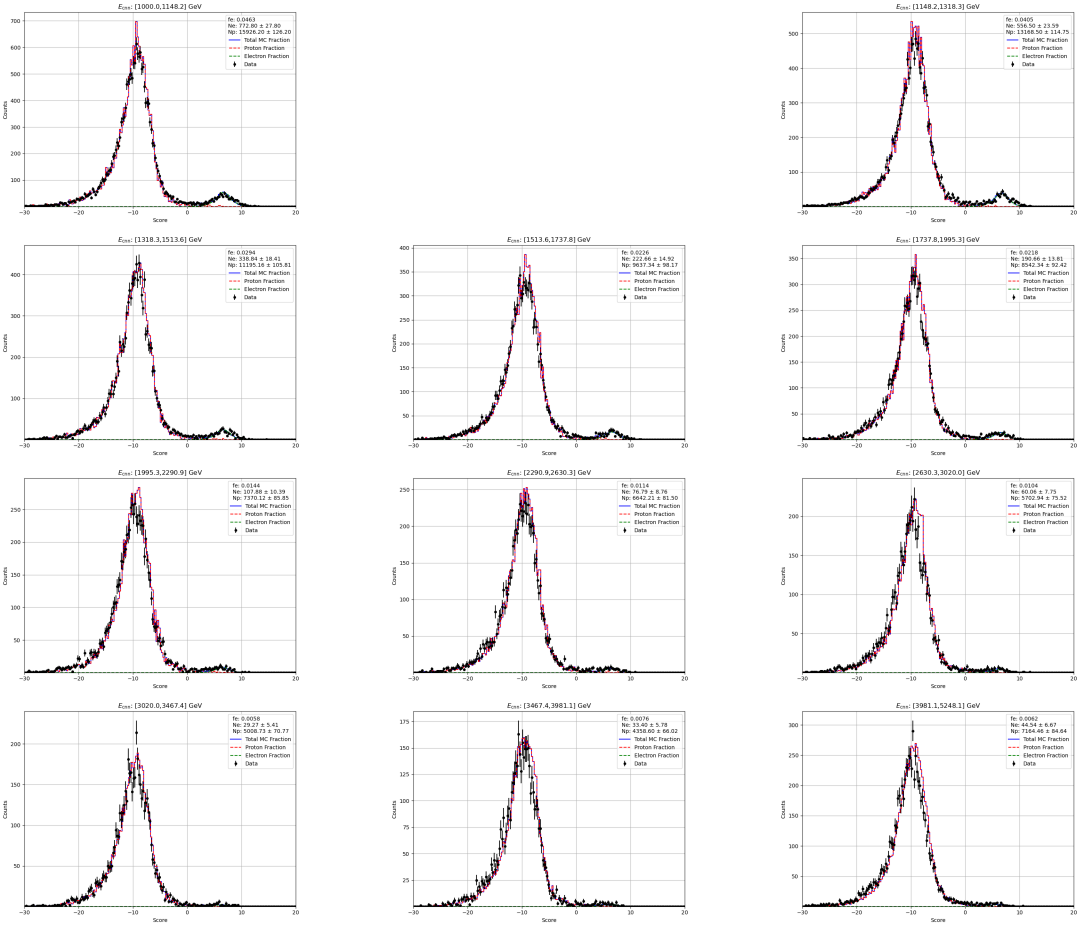


Figure 6.6: CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are data. Energies above 1 TeV. Linear scale.

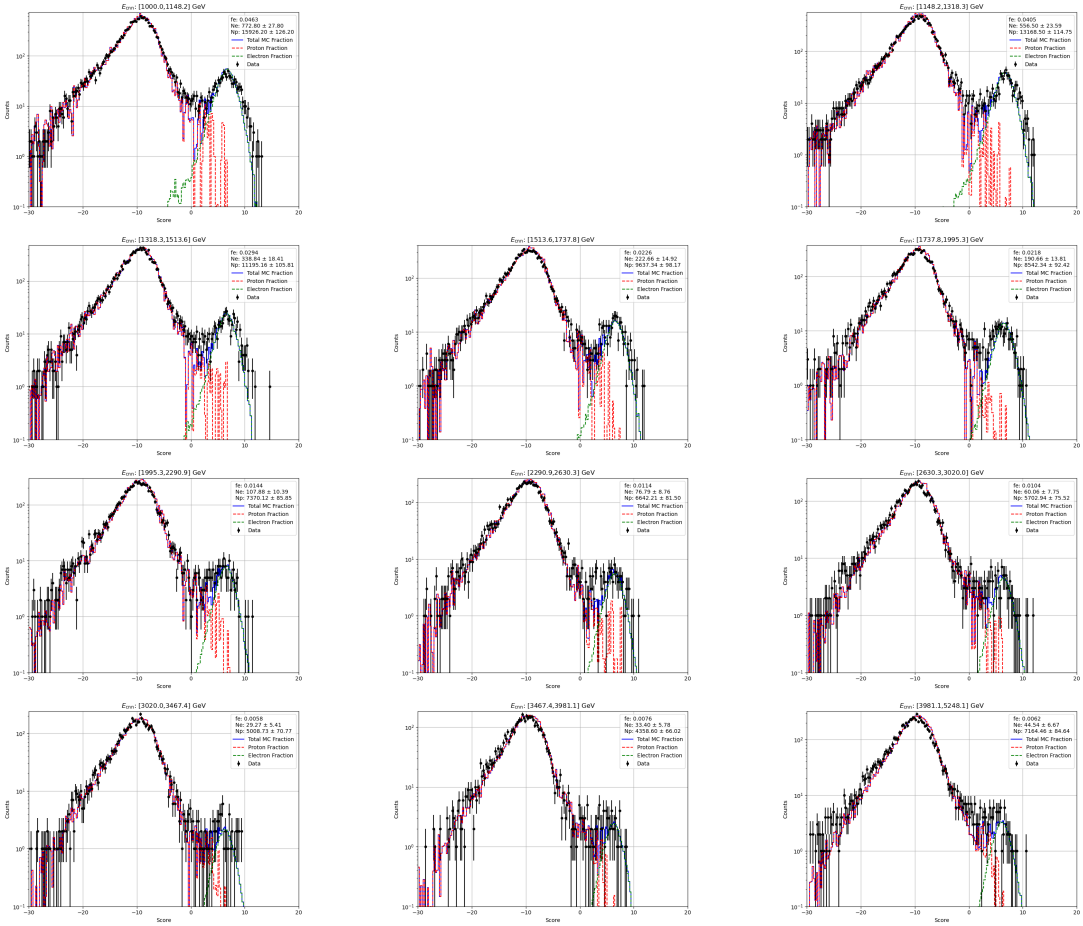


Figure 6.7: CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are data. Energies above 1 TeV. Log scale.

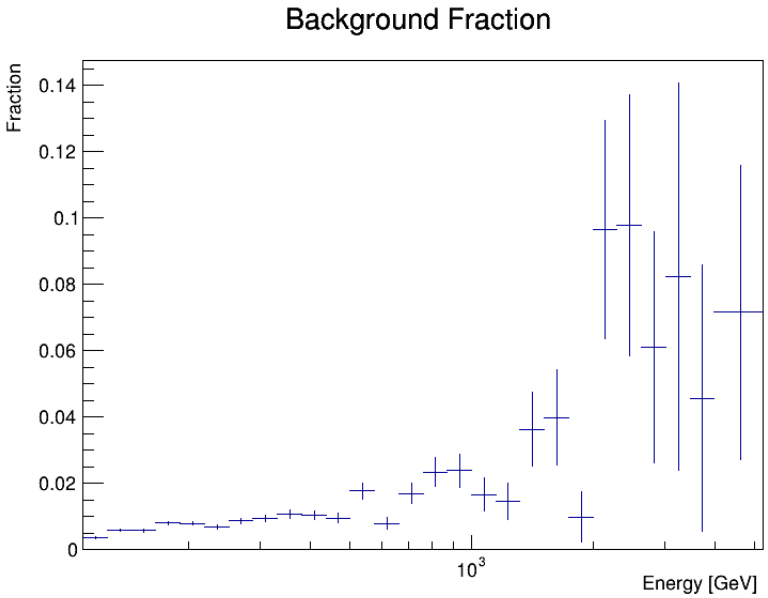


Figure 6.8: Number of background events that pass the CNN cut of 4.

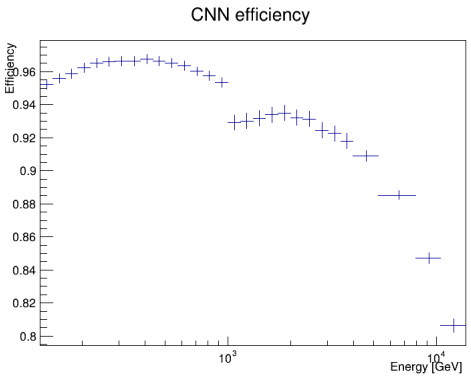


Figure 6.9: CNN Efficiency plot for electron MC.

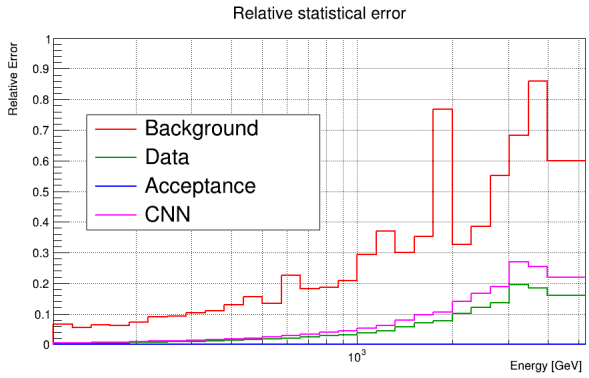


Figure 6.10: Statistical uncertainties as a function of energy.

analysis. Future work should aim to improve such estimations, for example, by fitting the background using interpolation, which could help mitigate the effects of statistical fluctuations.

We then compare the current results with other experiments –ground- and space-based experiments– see Figure 6.12. Our results are in agreement with Fermi-LAT and within the uncertainties band results up to 2 TeV. As the previous flux [46], our results are in tension with AMS-02 [81] and CALET [41]. From 200 GeV up to 1 TeV, VERITAS results also agree with our results. That being said, VERITAS uncertainties are significant, making it difficult to confirm a specific behaviour.

To verify the stability of the measured flux, we calculated it using a loose CNN cut (CNN $>$ 3) and a tight cut (CNN $>$ 5). Figure 6.13 compares the measured fluxes obtained with the nominal, loose, and tight cuts. We observe good agreement between the three measurements below 1 TeV. Above this energy, larger differences appear due to reduced separation power; however, these differences remain within the uncertainties. This confirms the overall stability of the CNN-based separation. Figure reffig 6.13 shows the relative ratios between the nominal cuts and the fluxes with tight and loose cut. Comparing the relative ratio with the uncertainties, we can see that with the exception of the first point, the relative ratio is below uncertainties, thus showing the stability and good agreement between fluxes.

6.5 Statistical Analysis

Having established the measured flux, we now aim to investigate potential spectral features. As discussed in Section 1.7, the contribution of local sources is expected to manifest in the multi-TeV range. Since our objective is to demonstrate the feasibility of measuring the CRE flux using non-fiducial events, we follow the approach of [46] and fit a single broken power law. The fitting procedure incorporates the values listed in

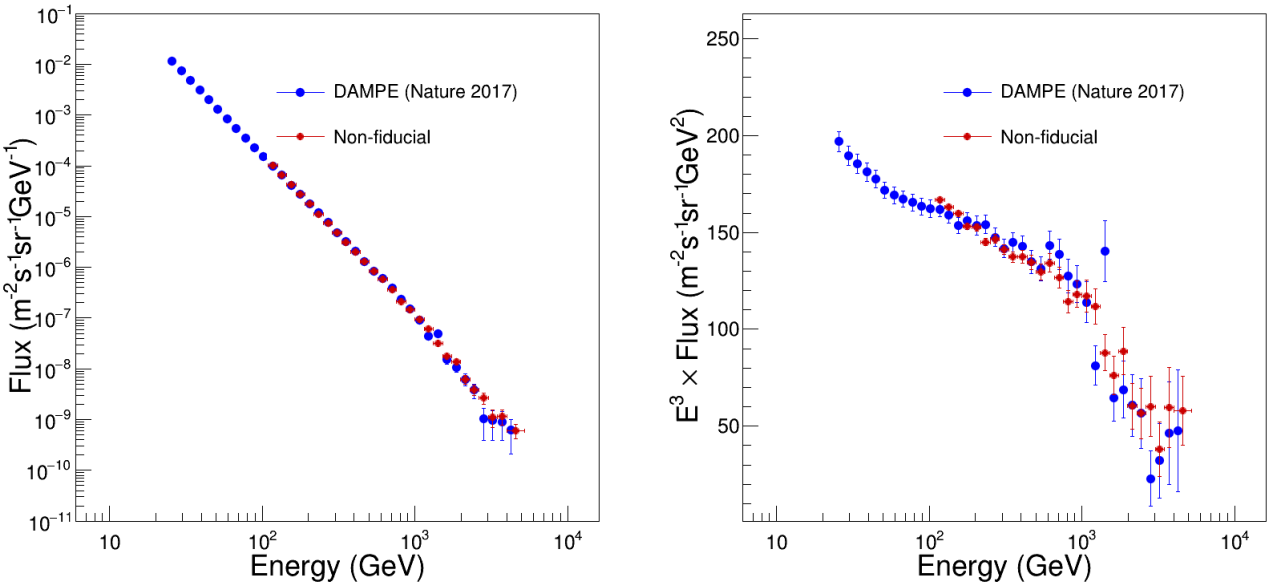


Figure 6.11: Comparison of the CRE flux (left plot) and E^3 scale flux (right flux). Blue point is the flux from 2017 [46]. Red point are the non-fiducial flux with data until the end of 2023. Error bars only take into account statistical uncertainties.

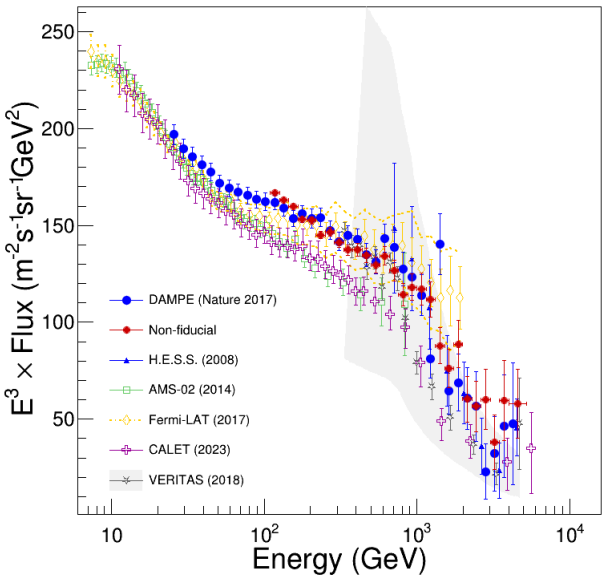


Figure 6.12: CRE flux for different experiments and results from this work. Red points (this work). Blue triangles from HESS [160]. Green squares from AMS-02 [81]. Yellow diamonds from Fermi-LAT [82]. Purple crosses from CALET [41]. Grey starts from VERITAS [161].

Table 6.1: The new DAMPE CRE spectrum (in units of $m^{-2}s^{-1}sr^{-1}GeV^{-1}$) with statistical and systematic uncertainties. E is the energy bin edges, $< E >$ is the bin centre, A is the effective acceptance, N is the number of signal events, Bckg. fraction is the background fraction

, $\Phi \pm \sigma_1 \pm \sigma_2$ is the flux measurement with σ_1 the statistical uncertainties and σ_2 is the uncertainties related to the background and efficiency.

E [GeV]	E_{centre} [GeV]	A [$m^2 \times sr$]	N	Bckg. fraction	$\Phi \pm \sigma_1 \pm \sigma_2$
109.6-125.9	117.4	0.200±0.001	64652	$(0.36 \pm 0.02)\%$	$(1.03 \pm 0.004 \pm 0.006) \cdot 10^{-4}$
125.9-144.5	134.8	0.228±0.001	54817	$(0.58 \pm 0.03)\%$	$(6.67 \pm 0.03 \pm 0.05) \cdot 10^{-5}$
144.5-166.0	154.8	0.236±0.002	42049	$(0.58 \pm 0.04)\%$	$(4.31 \pm 0.02 \pm 0.03) \cdot 10^{-5}$
166.0-190.5	177.7	0.242±0.002	31446	$(0.8 \pm 0.1)\%$	$(2.73 \pm 0.02 \pm 0.02) \cdot 10^{-5}$
190.5-218.8	204.0	0.246±0.002	24165	$(0.8 \pm 0.1)\%$	$(1.80 \pm 0.01 \pm 0.02) \cdot 10^{-5}$
218.8-251.2	234.2	0.249±0.003	17653	$(0.7 \pm 0.1)\%$	$(1.13 \pm 0.01 \pm 0.01) \cdot 10^{-5}$
251.2-288.4	268.9	0.252±0.003	13716	$(0.9 \pm 0.1)\%$	$(7.54 \pm 0.06 \pm 0.10) \cdot 10^{-6}$
288.4-331.1	308.8	0.254±0.004	10082	$(0.9 \pm 0.1)\%$	$(4.79 \pm 0.05 \pm 0.07) \cdot 10^{-6}$
331.1-380.2	354.5	0.255±0.004	7503	$(1.1 \pm 0.1)\%$	$(3.08 \pm 0.04 \pm 0.05) \cdot 10^{-6}$
380.2-436.5	407.1	0.259±0.005	5768	$(1.0 \pm 0.1)\%$	$(2.04 \pm 0.03 \pm 0.04) \cdot 10^{-6}$
436.5-501.2	467.4	0.259±0.006	4291	$(1.0 \pm 0.1)\%$	$(1.32 \pm 0.02 \pm 0.03) \cdot 10^{-6}$
501.2-575.4	536.6	0.261±0.006	3171	$(1.8 \pm 0.2)\%$	$(8.38 \pm 0.15 \pm 0.22) \cdot 10^{-7}$
575.4-660.7	616.1	0.260±0.007	2469	$(0.8 \pm 0.2)\%$	$(5.74 \pm 0.12 \pm 0.17) \cdot 10^{-7}$
660.7-758.6	707.4	0.260±0.009	1778	$(1.7 \pm 0.3)\%$	$(3.58 \pm 0.09 \pm 0.13) \cdot 10^{-7}$
758.6-871.0	812.2	0.260±0.010	1225	$(2.3 \pm 0.4)\%$	$(2.13 \pm 0.06 \pm 0.09) \cdot 10^{-7}$
871.0-1000.0	932.5	0.259±0.012	959	$(2.4 \pm 0.5)\%$	$(1.46 \pm 0.05 \pm 0.07) \cdot 10^{-7}$
1000.0-1148.2	1070.7	0.252±0.013	695	$(1.7 \pm 0.5)\%$	$(9.55 \pm 0.37 \pm 0.55) \cdot 10^{-8}$
1148.2-1318.3	1229.3	0.251±0.016	502	$(1.5 \pm 0.5)\%$	$(6.03 \pm 0.27 \pm 0.41) \cdot 10^{-8}$
1318.3-1513.6	1411.4	0.251±0.020	306	$(3.6 \pm 1.1)\%$	$(3.13 \pm 0.19 \pm 0.27) \cdot 10^{-8}$
1513.6-1737.8	1620.5	0.250±0.024	201	$(4.0 \pm 1.4)\%$	$(1.79 \pm 0.13 \pm 0.19) \cdot 10^{-8}$
1737.8-1995.3	1860.6	0.251±0.026	172	$(1.0 \pm 0.8)\%$	$(1.38 \pm 0.11 \pm 0.16) \cdot 10^{-8}$
1995.3-2290.9	2136.3	0.251±0.035	97	$(9.7 \pm 3.1)\%$	$(6.18 \pm 0.69 \pm 0.97) \cdot 10^{-9}$
2290.9-2630.3	2452.8	0.249±0.041	69	$(9.8 \pm 3.8)\%$	$(3.84 \pm 0.51 \pm 0.72) \cdot 10^{-9}$
2630.3-3020.0	2816.1	0.250±0.047	53	$(6.1 \pm 3.4)\%$	$(2.69 \pm 0.39 \pm 0.57) \cdot 10^{-9}$
3020.0-3467.4	3233.4	0.248±0.067	26	$(8.2 \pm 5.6)\%$	$(1.13 \pm 0.24 \pm 0.34) \cdot 10^{-9}$
3467.4-3981.1	3712.4	0.245±0.062	29	$(4.6 \pm 3.9)\%$	$(1.17 \pm 0.23 \pm 0.33) \cdot 10^{-9}$
3981.1-5248.1	4556.4	0.242±0.053	38	$(7.2 \pm 4.3)\%$	$(6.13 \pm 1.06 \pm 1.55) \cdot 10^{-10}$

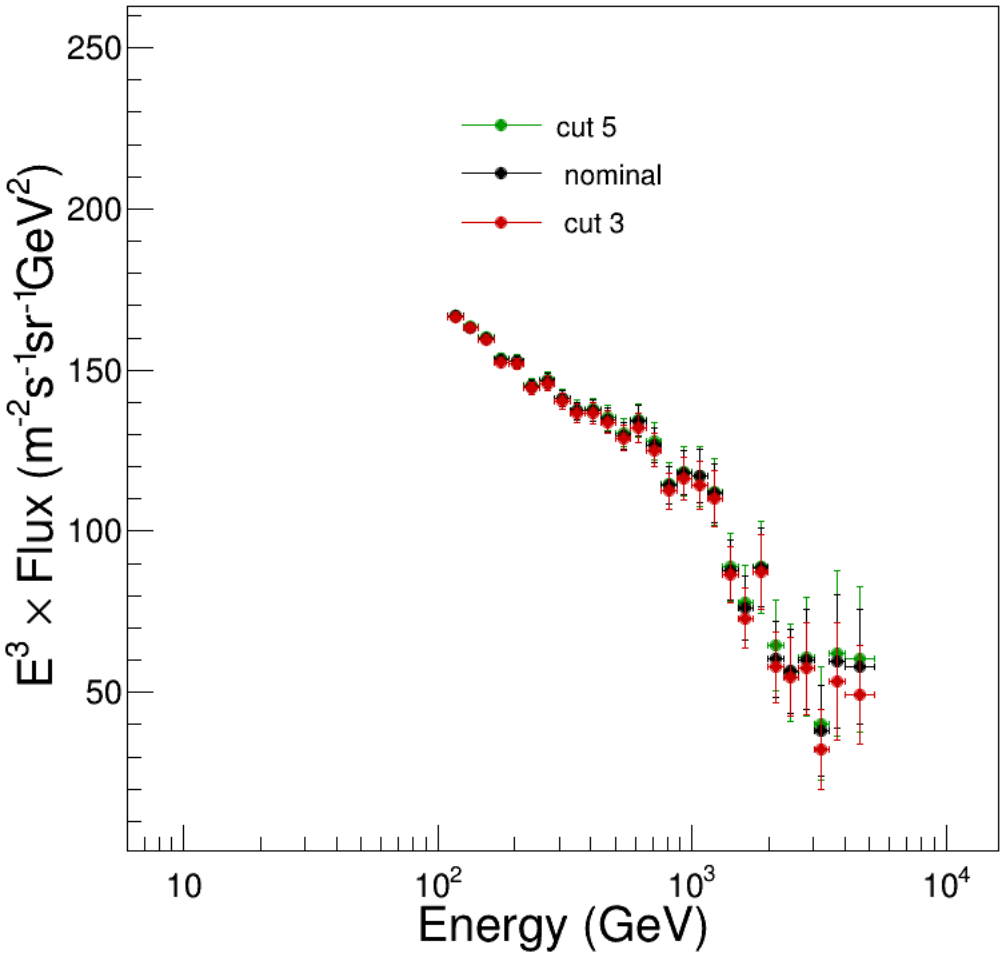


Figure 6.13: CRE flux measurement of this work for different cuts. Green points: flux measurement with a CNN cut of 5. Black point: nominal flux measurement (CNN cut of 4). Red points: flux measurement with a CNN cut of 3.

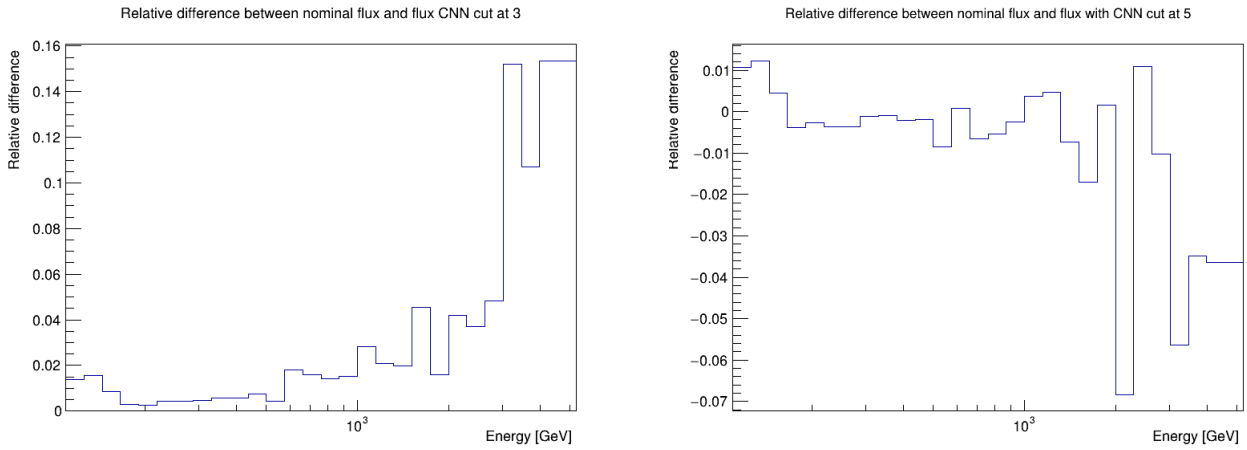


Figure 6.14: Distribution of the relative ratio between nominal flux and the measured flux with a loose cut (CNN cut at 3) and tight cut (CNN cut at 5). Left plot shows the relative ration for the loose cut. Right plot shows the relative ratio for the tight cut.

Table 6.1, the centre of the energy bin is defined in the following way:

$$E_{\text{centre}} = \left[\frac{E_{\text{HE}}^2 - E_{\text{LE}}^2}{2 \cdot (E_{\text{HE}} - E_{\text{LE}})}\right]^{\frac{-1}{3}}, \tag{6.4}$$

where E_{LE} is the lower edged of the energy bin, E_{HE} is the higher edge of the energy bin. To account for systematic effects during the fitting, such effects can be modeled as a sequence of unknown correction factor. They correspond to nuisance parameters that are derived in the fit of the CRE flux when we test different parametrisation.

We fit our data from 100 GeV to 2.63 TeV using a single broken power law. We include 7 nuisance parameters to account for the uncertainties. Following [162] the smoothing factor Δ is set to 0.1. The fitting function is:

$$\Phi = \Phi_0 \left(\frac{E}{100\text{GeV}}\right)^{-\gamma_1} \left[1 + \left(\frac{E}{E_b}\right)^{\frac{\gamma_2 - \gamma_1}{\Delta}}\right]^{-\Delta}, \tag{6.5}$$

where Φ_0 is the normalisation factor, E_b is the break energy, γ_1 is the spectral index before the break and γ_2 after the break.

Single Broken Power Law

The results of the fit are shown in Figure 6.15 and below we show the values of the different fitted parameters:

- Normalisation factor: $\Phi_0 = 1.17 \pm 0.02 \times 10^{-7}$

- Power law indices: $\gamma_1 = 3.16 \pm 0.01$ and $\gamma_2 = 3.854 \pm 0.162$
- Break energy: $E_b = 993 \pm 108\,\text{GeV}$
- $\chi^2 = 33.7$

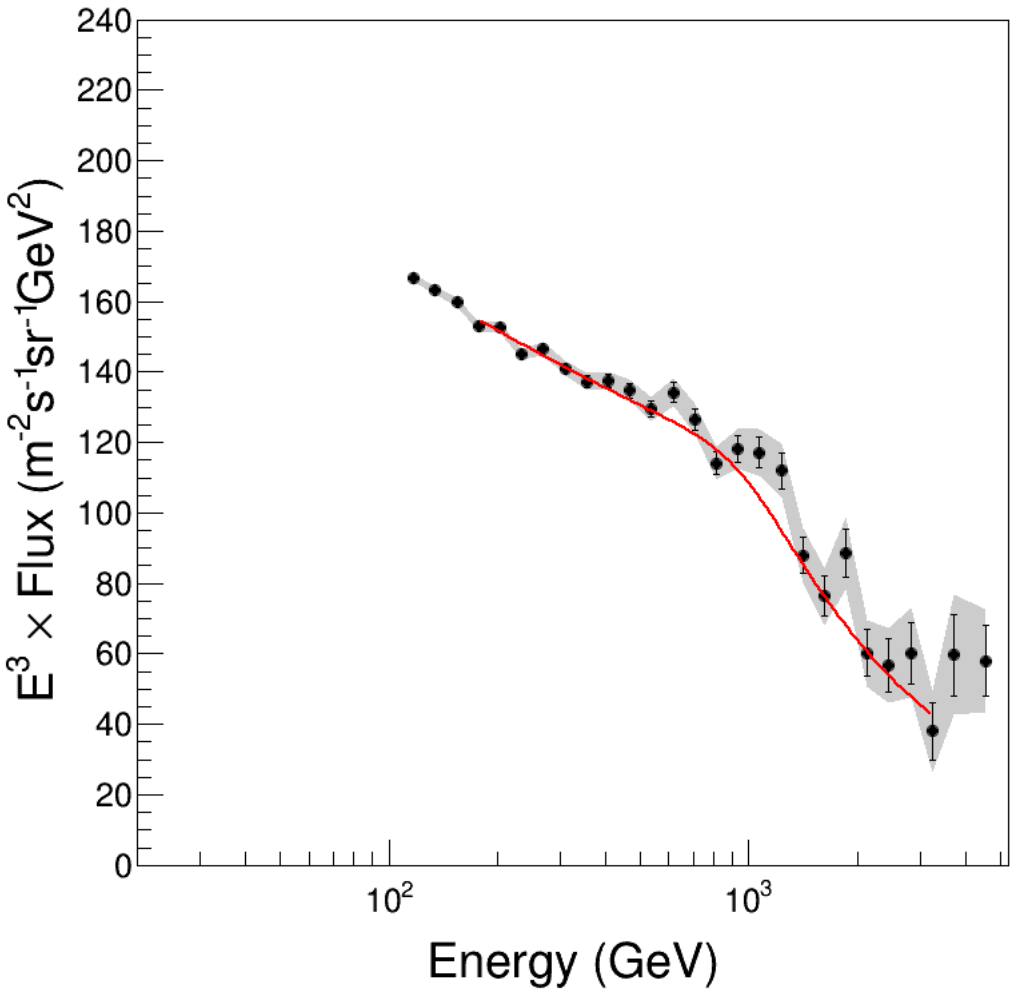


Figure 6.15: Single broken power law fitting of the CRE flux. Grey band includes the uncertainties from the background estimation.

In conclusion, the analysis presented in this chapter demonstrates that non-fiducial electrons can be successfully reconstructed and used to measure the CRE flux from 100 GeV up to 3 TeV. Our results show good agreement with the previous DAMPE measurement [46], validating the approach. This also indicates that non-fiducial events can be considered in future extensions to higher energies, particularly for studies in the multi-TeV range, where contributions from local sources of CRE are expected [2–4].

Conclusion

The discovery of cosmic rays by Victor Hess in 1915, coupled with the development of advanced detector technologies, marked the beginning of a new era in cosmic ray physics. This breakthrough not only confirmed the existence of radiation originating from space but also enabled the identification of particle types and their electric charges. While cosmic rays have been shown to consist primarily of protons, helium nuclei, and some heavier ions, electrons and positrons—due to their relatively low mass—are particularly valuable at multi-TeV energies for tracing potential local sources and investigating possible dark matter signatures. At these high energies, however, detecting electrons and positrons becomes increasingly challenging and demands sophisticated analysis techniques. In this context, machine learning offers a powerful and promising approach for enhancing event classification and interpretation.
The DAMPE experiment has been operating continuously since 2015. It has already confirmed previous observations and has significantly expanded our knowledge of cosmic rays. Notably, in 2017, DAMPE reported the observation of a spectral break in the cosmic-ray electron and positron (CRE) spectrum. Thanks to its uninterrupted data collection, DAMPE has extended CRE observations into the multi-TeV range. However, at these energies, the main limiting factor becomes statistics.

To mitigate this limitation, the acceptance of DAMPE can be increased by including events with larger incidence angles, namely non-fiducial events. A major caveat of selecting such events is the increased energy leakage in the BGO calorimeter due to shallower penetration depths. This leakage leads to poorer energy resolution and degrades the accuracy of electron–proton discrimination, since the calorimeter plays a crucial role in particle identification. Consequently, the current variables used for classification become less effective, necessitating more advanced methods for reliable separation.

Machine learning is well-suited for this type of analysis. In this work, we apply CNN techniques for two primary purposes. First, we develop a CNN model for energy reconstruction. As previously demonstrated, classical techniques underestimate the true energy of particles since energy looses and leakage are not taken into account. While existing correction methods improve energy reconstruction for fiducial events, they cannot be directly applied to non-fiducial events.

To address this limitation, we trained a CNN using an image-based representation of the shower development in the BGO calorimeter. We evaluated the performance of each technique by analysing the relative error using Monte Carlo (MC) simulations. The results showed that the CNN method yielded a mean relative error distribution centred near zero and maintained stability across all energy ranges. We further examined the 68% and 95% energy containment radii. For the 68% containment, we observed notable improvement at higher energies. The 95% containment remained stable overall, with the exception of the highest energy bin, where performance was comparable to that of the classical correction method.

However, to validate the CNN approach for use with real data, it is essential to study its behaviour on experimental datasets. For this purpose, we first used beam test data, which provides known initial particle energies. While the CNN achieved performance comparable to classical methods, these beam test events featured only small incidence angles and were not represented in the training set. Therefore, the results, although promising, are insufficient to fully validate the model.

To further assess the method, we turned to actual flight data. Since this dataset contains both electrons and protons, we trained a dedicated CNN classifier to separate the two species. After event selection, we applied the CNN energy reconstruction model across different reconstructed energy bins ($E_{\rm rec}$). A general agreement between data and MC was observed. Although some differences remained in the 68% and 95% energy containment comparisons between data and simulation, the overall consistency supports the conclusion that the CNN-based energy reconstruction method is robust and effective.

The second application of machine learning involved the development of a discriminator between electrons and protons. To achieve this, we trained two separate CNN models: one for events with $E_{\rm cnn} < 1$ TeV, and another for events with $E_{\rm cnn} \geq 1$ TeV. As in the previous case, we used an image representation of the particle showers in the calorimeter. This time, the training datasets consisted of both electrons and protons.

After training the CNNs, we evaluated their performance against the traditional discriminator ζ by comparing their ROC curves. The CNN-based method demonstrated a significantly better separation, outperforming ζ by approximately one order of magnitude.

We then validated the models using real data and Monte Carlo (MC) simulations. To do this, we normalized the electron and proton MC distributions to the data by fitting the electron fraction. A discrepancy was observed between the proton MC and the data, particularly in the bulk and left tail of the proton distribution. However, overall agreement was found in the right tail of the proton spectrum and in the electron region. Crucially, this mismatch does not affect future cosmic-ray electron (CRE) analyses, as only the region near the signal is relevant for background estimation.

Thanks to the implementation of both CNN models, we have updated the DAMPE CRE flux measurement using eight years of data. The current analysis applies a series of selection cuts to remove poorly reconstructed events as well as to suppress contamination from protons, helium, and heavier nuclei. For energies below 1 TeV, the preselection efficiency exceeds 93%, while for energies above 1 TeV and up to approximately 10 TeV, it decreases slightly to about 89%. After applying the CNN classifier, the efficiency

remains above 95% below 1 TeV, and gradually decreases from 92.5% to 80% as energy increases. This demonstrates that the CNN is effective in distinguishing electrons from protons across a broad energy range. Despite this, a residual background remains in the signal region (defined as $\mathrm{CNN}_{\mathrm{score}} > 4$). This remaining background contribution was estimated using a cut-and-count method. By accounting for these effects, we have not only refined and updated previous results but have also successfully extended the CRE flux measurement up to 3 TeV. To explore potential spectral features in the updated CRE flux, the data from 100 GeV to 3 TeV was fitted using both single broken power law models. The analysis includes 7 nuisance parameters to account for uncertainties related to the background estimation and CNN efficiency.

We conclude by showing that this work allowed to show that by opening the acceptance it is possible to include non-fiducial events and the possibility to recover similar results than previous work. This motivates futures work to extended the CRE flux to the multi-TeV region by including non-fiducial events.

Future Prospects

As mentioned in Chapter 6, the CRE flux measurement presented in this work does not include systematic uncertainties, which represent a critical component of any precision measurement. Therefore, a key improvement to this study would be the inclusion of such uncertainties. Another important aspect is the estimation of the background, which, as previously noted, is underestimated above 1 TeV. Additionally, the current cut-and-count method is particularly sensitive to statistical fluctuations. These effects could be mitigated through interpolation techniques, which provide a smoother and more stable background estimation.

Another promising direction for improvement lies in the expanded use of machine learning. The development of attention mechanisms [163], which allow models to dynamically focus on the most relevant parts of the input, offers a significant advantage over traditional approaches such as multilayer perceptrons (MLPs) or convolutional neural networks (CNNs), whose weights are fixed during each iteration. These newer techniques may offer enhanced performance and adaptability in complex tasks such as energy reconstruction and particle identification.

While this work focused solely on non-fiducial events, it would be valuable to combine them with fiducial events, potentially increasing the available statistics by a factor of 1.5 compared to fiducial-only analyses. One possible approach would be to merge both datasets after event selection, using the Ω classifier [138] for fiducial events and the CNN model for non-fiducial ones. However, a key challenge in this strategy lies in combining the uncertainties from both analyses in a consistent manner.

To circumvent this issue, an alternative approach would be to train a new deep neural network from scratch—based on MLP, CNN, or even more advanced architectures such as transformers—using the combined dataset. This would allow for a unified treatment of the data and associated uncertainties, potentially leading to improved performance and robustness.

List of Figures

1.1 The induction charging of a gold-leaf electroscope is an early instrument used to detect the presence of electric charge on a body. The one shown in the figure is composed of two metallic foils at the bottom, connected to a metallic rod with at the top. The foils are encapsulated in a glass bottle to be protected from air friction. By Sylvanus P. Thompson [public domain], via Wikimedia Commons. 3

1.2 Electroscope built by Theodor Wulf. By Alessandro De Angelis [public domain], via Wikimedia Commons. 5

1.3 Plots representing the number of Ion pair as a function of the altitude. Left plot: Results from Hess ascent in 1912. Right plot: results from Kolhôrster ascent. Image by A. De Angelis [5] 6

1.4 Schematic view of ATIC structure [15]. 8

1.5 Schematic of the CREAM instrument configuration [19]. 9

1.6 Representation of the AMS-02 detector layout [23]. 10

1.7 Representation of a shower development and different detection techniques. Image Author: Konrad Bernlöhr, from `https://www.mpi-hd.mpg.de/hfm/CosmicRay/ShowerDetection.html`. 13

1.8 Side view of a shower propagating through the atmosphere. Left figure: shower from a 50 GeV photon. Right plot: 1 TeV proton shower. From `https://www.iap.kit.edu/corsika/`. 14

1.9 Plot of the differential energy flux, scaled by E^2 for different particle species. The different markers present the results of different experiments. The grey arrow at the bottom shows the collision energy of the experiment LHC in the fixed target frame [57]. 16

1.10 Schematic view of an AGN, with the different components. Depending on the angle between the observer and the plane of the AGN, we observe different behaviours. Image from `https://fermi.gsfc.nasa.gov/science/eteu/agn/`. 19

1.11 Interaction of a particle with a strong shock wave. (a) Shock waves propagating at speed U, separating the medium into two regions with different pressure and temperature conditions. (b) Flow of interstellar gas surrounding the shock front. (c) Flow of gas in the frame in which the upstream is stationary and the speed distribution is isotropic. (d) Flow of gas in the frame in which the downstream is stationary and the speed distribution is isotropic. Image from [69]. 23
1.12 The AMS electron (blue data points) and positron (red data points, multiplied by 10) spectra . For display purposes the electron data point at $\sim$ 830 GeV is slightly shifted horizontally to avoid overlap with the positron point [36]. 28
1.13 The spectral indices of the electron flux γ_{e^-} and of the positron flux γ_{e^+} as a function of energy. The green band represents the 68 % confidence level interval of a fit to the positron flux. The blue band represents the 68 % confidence level interval of a fit to the electron flux. [36]. 28
1.14 electron+positron flux measured by DAMPE and comparison with other experiments [46]. H.E.S.S [80], AMS-02 [81], Fermi-LAT [82] 29
1.15 Positron flux measured by AMS-02 [32] together with earlier measurements from PAMELA [83], Fermi-LAT [40], MASS [84], CAPRICE [85], AMS-01 [86], and HEAT [87]. 30
1.16 Example of expected spectra of electrons with an energy cut off of $E_c =$ 20 TeV, escape time after SNR formation $\tau = 5 \times 10^3$ yr [2]. 31
1.17 Rotation curve data for M31 [95]. 31
1.18 Overview of three types of searches for dark matter particles (represented by χ), all based on interactions with standard model particles (represented by f). The three Feynman diagrams on the right are specifications of the diagram on the left, where the arrows on the left diagram indicate the direction in which the diagram should be read. On the right, the dashed lines represent the mediator, and g_{ff} and g_{DM} are the coupling strengths at the standard model and dark matter vertex, respectively [100]. 32
1.19 Upper limits on the self interacting DM-nucleon cross section as a function of DM mass. Plot from Chapter 27 of [97]. 33
1.20 Comparison of the AMS positron flux (red data points) with predictions of a dark matter model [109] (red line) based on a DM particle with a mass of 1.2 TeV [23] and classical cosmic-ray contribution (black line) . . 34

2.1 Schematic representation of DAMPE and its layout. Figure from [110]. . . 35
2.2 Full view of the PSD. Figure from [120] 37
2.3 The side view of the PSD . Figure from [120] 37
2.4 PSD's charge spectrum. Figure from [121] 38
2.5 View of STK showing the 64 SSDs layers, the readout board (TRB) and the radiator used to dissipated the heat. Figure from [123] 39

2.6 Spatial resolution as a function of the impact angle. Below 70 , the resolution is much lower than the 70 μm achieved with a digital position finding algorithm [124]. Figure from [117]. 40

2.7 STK charge resolution as a function of the identify charge. Figure from [125]. 40

2.8 STK signal mean distribution for nuclei produced by a lead beam on target, after removing $Z = 1$ particles. The signal mean, with current reconstruction procedures, allows for the identification of ions until Oxygen. Figure from [110]. 41

2.9 BGO bar arrangement in the calorimeter. Figure from [127] 42

2.10 Schematic of the BGO readout 42

2.11 Energy resolution of the BGO calorimeter [110]. Left plot: DAMPE beam test results (with electrons) are over-plotted. Right plot: The energy resolution for on-axis protons. The dotted line represents the energy resolution of MC simulated protons after spectral unfolding while the red points represent the beam test data. 42

2.12 Schematic of the NUD detector [110]. 43

2.13 NUD signal for 450 GeV protons and 150 GeV electrons (distributions are normalized to unit area) [110]. 43

2.14 Architecture of DAMPE's DAQ [110]. 44

2.15 Trigger decision logic [110]. 45

3.1 Example of how a convolutional layer extracts features from an input image. On the left, the original image given to the convolutional layer; in the centre, the kernel used by the layer, to detect ridges and edges; on the right, the output feature map. Images originally from [142] 48

3.2 A convolutional layer applies a kernel K to a local region (image patch) of the input I. On the left it represented the input's pixel values. The center represents the convolutional function, as described in Eq 3.1. The 3×3 kernel is slid over the input image, performing element-wise multiplication with the overlapping image patch followed by summation. The result for each patch is placed in the corresponding position of the output feature map, as shown in right image. In this example, the dot product of the first image patch and the kernel yields a value of 31 in the top-left cell of the output. Images originally from `https://anhreynolds.com/blogs/cnn.html` 48

3.3 Visual representation of a CNN wit. Images originally from `https://www.plainconcepts.com/convolutional-neural-network-guide/` 49

3.4 Example of under and overfitting. Blue points are a points with some noise. On the left, a set of data points is fitted with a linear function; the same set of points is fitted with a quadratic function, at the centre, and a 14-th order polynomial function. The results show how the complexity of the linear fitting is not enough to describe the behaviour of data, while the polynomial function fits too well to the specific observation, loosing on generality, if provided with a new sample of the same observable. The quadratic function has the better balance to fit the given sample without loosing in generality. 51
3.5 Visual example of the gradient descent technique used for a function $f(\vec{x}) = x_i^2 + x_j^2$, with $\vec{x}$ a bi-vector of elements x_i, x_j. The red line shows the gradient descent path for 50 iterations, with a starting point $\vec{x}_0 = (1, 3.5)$. 53

4.1 View of a non-fiducial event (left figure) and fiducial event (right figure) crossing the calorimeter. The BGO images combine the X-Z and Y-Z views 55
4.2 Distribution of the relative energy error obtained with electron MC simulations in four different energy bins, ranging from 100 GeV to 24 TeV. Two types of MC samples are used – fiducial and non-fiducial, analysed with two energy reconstruction methods – standard [110] and corrected [149]. Distributions are denoted as follows: fiducial standard (blue solid line), non-fiducial standard (blue dashed line), fiducial corrected (black solid line) and non-fiducial corrected (black dashed line). The grey line indicates ideal case when reconstructed energy equals the true kinetic energy. 57
4.3 Mean of the relative energy error as a function of kinetic energy for the different energy reconstruction methods. 58
4.4 Energy resolution as a function of kinetic energy. Left plot: 68% containment, right plot: 95% containment. 58
4.5 CNN model structure [145]. 59
4.6 Evolution of the loss, defined as the mean squared error for the training sample (blue curve) and validation sample (orange curve). 60
4.7 The distribution of the relative energy error (Equation 4.1) for non-fiducial electron MC samples ranging from 100 GeV to 24 TeV. Light grey curve: for the standard energy reconstruction method [110], dark grey curve: standard electron energy correction method [149] and magenta curve: CNN method [145]. The grey vertical line indicates the ideal case where the reconstructed energy equals the true kinetic energy. 61
4.8 Plot of the mean of the relative energy error (Equation 4.1) for non-fiducial electron MC samples ranging from 100 GeV to 24 TeV. Light grey curve: for the standard energy reconstruction method [110], dark grey curve: standard electron energy correction method [149] and magenta curve: CNN method [145]. The horizontal grey line indicates the ideal case when the reconstructed energy equals the true kinetic energy. . 62

4.9 The 68% and 95% containment of the relative energy reconstruction errors for non-fiducial electron MC samples as function of the kinetic energy. . . 62
4.10 Response matrix of the three energy reconstruction methods: CNN-corrected, standard corrected and standard reconstruction (no correction). Red line indicates a diagonal of a matrix. 63
4.11 Electron kinetic energy spectrum obtained with the CNN method with and without the unfolding procedure (cyan and blue lines respectively) compared with the truth spectrum defined in MC. 64
4.12 Distribution of the reconstructed energy for beam test data. Full line are . 64
4.13 Ratio of the calculate spectrum, with and without unfolding, divided by the truth kinetic energy spectrum defined in MC. 65
4.14 CNN model structure. 67
4.15 Distribution of CNN classifier for data and MC samples for different energy bins in linear scale. 68
4.16 Distribution of CNN classifier for data and MC samples for different energy bins in logarithmic scale. 69
4.17 Distribution of CNN energy correction method for data and MC samples. Lower plot are the ratio between real and MC data. Left plots are in linear scale, right plots are in log scale. 71
4.18 Mean value of the CNN reconstructed energy for data and MC as a function of the standard energy reconstruction method [110]. 72
4.19 The 68% and 95% containment of the CNN energy reconstruction method as a function of standard energy reconstruction method [110]. 72

5.1 ζ distribution for non-fiducial events. Real data (black dots), electrons MC (green line) and protons MC (red line) for 5 energy bins from 100 GeV to 24 TeV. 76
5.2 CNN structure use for the low energy and high energy model. 77
5.3 Evolution of the loss, defined as the binary-cross entropy for the training sample (blue curve) and validation sample (orange curve). Left plot: Model trained below 1 TeV. Right plot: Model trained above 1 TeV. . . . 77
5.4 Comparison of ROC curves on MC for the ζ variable (grey curve) and the LE CNN classifier (blue curve). A lower curve indicates better performance. Shade bands represent statistical uncertainties. 79
5.5 Comparison of ROC curves on MC for the ζ variable (grey curve) and the HE CNN classifier (blue curve). A lower curve indicates better performance. Shade bands represent statistical uncertainties. 80
5.6 Comparison of ROC curves on MC for the ζ variable (grey curve) and the HE CNN classifier (blue curve). A lower curve indicates better performance.Shade bands represent statistical uncertainties. 81
5.7 Background fraction as a function of the energyζ variable (grey curve) and the HE CNN classifier (blue curve). A lower curve indicates better performance. Shade bands represent statistical uncertainties. 81

5.8 CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are real data. Energies below 1 TeV. Linear scale. 82
5.9 CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are real data. Energies below 1 TeV. Logarithmic scale. 83
5.10 CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are real data. Energies above 1 TeV. Linear scale. 84
5.11 CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are real data. High energy model. Log scale. 85

6.1 DAMPE daily event counts. The horizontal line shows the average number of events taken per day. 87
6.2 Daily event count that passes the selection described in Section 4.1 from DAMPE. Blue line shows events with energies in the range [20,100] GeV, orange line events with energy within [100,250] GeV, green line for energies in the range [250,500] GeV, red line for energies in the range [0.5,1] TeV, purple line for [1,5] TeV and brown line for events with energies above 5 TeV. 89
6.3 Preselection efficiency as a function of the energy. 90
6.4 CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are data. Energies below 1 TeV. Linear scale. 92
6.5 CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are data. Energies below 1 TeV. Log scale. 93
6.6 CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are data. Energies above 1 TeV. Linear scale. 94
6.7 CNN prediction distribution for different energy bins. Proton MC in red. Electron MC in green. Total MC in blue. Black points are data. Energies above 1 TeV. Log scale. 95
6.8 Number of background events that pass the CNN cut of 4. 96
6.9 CNN Efficiency plot for electron MC. 96
6.10 Statistical uncertainties as a function of energy. 97
6.11 Comparison of the CRE flux (left plot) and E^3 scale flux (right flux). Blue point is the flux from 2017 [46]. Red point are the non-fiducial flux with data until the end of 2023. Error bars only take into account statistical uncertainties. 98

6.12 CRE flux for different experiments and results from this work. Red points (this work). Blue triangles from HESS [160]. Green squares from AMS-02 [81]. Yellow diamonds from Fermi-LAT [82]. Purple crosses from CALET [41]. Grey starts from VERITAS [161]. 98
6.13 CRE flux measurement of this work for different cuts. Green points: flux measurement with a CNN cut of 5. Black point: nominal flux measurement (CNN cut of 4). Red points: flux measurement with a CNN cut of 3. 100
6.14 Distribution of the relative ratio between nominal flux and the measured flux with a loose cut (CNN cut at 3) and tight cut (CNN cut at 5). Left plot shows the relative ration for the loose cut. Right plot shows the relative ratio for the tight cut. 101
6.15 Single broken power law fitting of the CRE flux. Grey band includes the uncertainties from the background estimation. 102

List of Tables

1.1 Order of magnitudes of radius and magnetic fields in accelerations sites, and the maximum attainable energy [5]. 16

2.1 Energy range of the different readout dynode for each PMT [110]. 38

4.1 Mean and standard deviation for the three energy method for beam test samples of 100 GeV. For beam test data and MC samples. 65
4.2 Number of MC and data events passing the CNN classifier, for a cut value of 4. Last column represents the fraction of background that passes the cut. 70

6.1 The new DAMPE CRE spectrum (in units of $m^{-2}s^{-1}sr^{-1}GeV^{-1}$) with statistical and systematic uncertainties. E is the energy bin edges, $< E >$ is the bin centre, A is the effective acceptance, N is the number of signal events, Bckg. fraction is the background fraction 99

Bibliography

[1] Anderson CD. The Apparent Existence of Easily Deflectable Positives. Science. 1932;76(1967):238-9. Available from: https://www.science.org/doi/abs/10.1126/science.76.1967.238.

[2] Kobayashi T, Komori Y, Yoshida K, Nishimura J. The Most Likely Sources of High-Energy Cosmic-Ray Electrons in Supernova Remnants. apj. 2004;601(1):340-51. Available from: https://dx.doi.org/10.1086/380431.

[3] Luo Q, Qiao BQ, Liu W, Cui SW, Guo YQ. Statistical study of the optimal local sources for cosmic ray nuclei and electrons. Astrophys J. 2022 May;930(1):82. Available from: https://iopscience.iop.org/article/10.3847/1538-4357/ac6267.

[4] Evoli C, Amato E, Blasi P, Aloisio R. Stochastic nature of Galactic cosmic-ray sources. Phys Rev D. 2021 Dec;104:123029. Available from: https://link.aps.org/doi/10.1103/PhysRevD.104.123029.

[5] De Angelis A, Pimenta MJM. Introduction to Particle and Astroparticle Physics : Questions to the Universe. 1st ed. Undergraduate Lecture Notes in Physics. Milano: Springer Milan; 2015.

[6] Clay J. The origin of penetrating radiation. Proceedings of the Royal Academy of Amsterdam. 1927;30. Available from: https://www.dwc.knaw.nl/DL/publications/PU00011919.pdf.

[7] Rossi B. Directional Measurements on the Cosmic Rays Near the Geomagnetic Equator. Phys Rev. 1934 Feb;45:212-4. Available from: https://link.aps.org/doi/10.1103/PhysRev.45.212.

[8] Grupen C, Shwartz B. Particle Detectors. 2nd ed. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press; 2008.

[9] Abraham J, Aglietta M, Aguirre IC, Albrow M, Allard D, Allekotte I, et al. Properties and performance of the prototype instrument for the Pierre Auger Observatory. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2004;523(1):50-95. Available from: https://www.sciencedirect.com/science/article/pii/S0168900203033497.

[10] Hinton JA. The status of the HESS project. New Astronomy Reviews. 2004;48(5):331-7. 2nd VERITAS Symposium on the Astrophysics of Extragalactic Sources. Available from: https://www.sciencedirect.com/science/article/pii/S1387647303003555.

[11] DeYoung T. The HAWC observatory. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2012;692:72-6. 3rd Roma International Conference on Astroparticle Physics. Available from: https://www.sciencedirect.com/science/article/pii/S0168900212000666.

[12] Science with the Cherenkov Telescope Array. WORLD SCIENTIFIC; 2019. Available from: https://www.worldscientific.com/doi/abs/10.1142/10986.

[13] Kawai H, Yoshida S, Yoshii H, Tanaka K, Cohen F, Fukushima M, et al. Telescope Array Experiment. Nuclear Physics B - Proceedings Supplements. 2008;175-176:221-6. Proceedings of the XIV International Symposium on Very High Energy Cosmic Ray Interactions. Available from: https://www.sciencedirect.com/science/article/pii/S0920563207007992.

[14] Guzik TG, Adams JH, Ahn HS, Bashindzhagyan G, Chang J, Christl M, et al. The ATIC long duration balloon project. Advances in Space Research. 2004;33(10):1763-70. The Next Generation in Scientific Ballooning. Available from: https://www.sciencedirect.com/science/article/pii/S0273117703011566.

[15] Ganel O, Adams JH, Ahn HS, Ampe J, Bashindzhagyan G, Case G, et al. Beam tests of the balloon-borne ATIC experiment. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2005;552(3):409-19. Available from: https://www.sciencedirect.com/science/article/pii/S0168900205014014.

[16] Adams J. Silicon matrix detector for ATIC. In: 26th International Cosmic Ray Conference (ICRC26), Volume 5. vol. 5; 1999. p. 76.

[17] Chang J, Adams JH, Ahn HS, Bashindzhagyan GL, Christl M, Ganel O, et al. An excess of cosmic ray electrons at energies of 300–800 GeV. Nature. 2008 Nov;456(7220):362-5. Available from: https://doi.org/10.1038/nature07477.

[18] Panov AD, Adams JH, Ahn HS, Bashindzhagyan GL, Batkov KE, Chang J, et al.. The results of ATIC-2 experiment for elemental spectra of cosmic rays; 2006. Available from: https://arxiv.org/abs/astro-ph/0612377.

[19] Seo ES, Ahn HS, Beatty JJ, Coutu S, Choi MJ, DuVernois MA, et al. Cosmic-ray energetics and mass (CREAM) balloon project. Advances in Space Research. 2004;33(10):1777-85. The Next Generation in Scientific Ballooning. Available from: https://www.sciencedirect.com/science/article/pii/S027311770301158X.

[20] Ahn HS, Allison P, Bagliesi MG, Barbier L, Beatty JJ, Bigongiari G, et al. ENERGY SPECTRA OF COSMIC-RAY NUCLEI AT HIGH ENERGIES. The Astrophysical Journal. 2009 nov;707(1):593. Available from: https://dx.doi.org/10.1088/0004-637X/707/1/593.

[21] Yoon YS, Ahn HS, Allison PS, Bagliesi MG, Beatty JJ, Bigongiari G, et al. Cosmic-ray Proton and Helium Spectra from the First CREAM Flight. apj. 2011 feb;728(2):122.

[22] Kurnosova LV, Kolobyanina TN, Logachev VI, Razorenov LA, Sirotkin IA, Fradkin MI. Discovery of radiation anomalies above the South Atlantic at heights of 310–340 km. Planetary and Space Science. 1962;9(8):513-6. Available from: https://www.sciencedirect.com/science/article/pii/0032063362900570.

[23] Aguilar M, Ali Cavasonza L, Ambrosi G, Arruda L, Attig N, Barao F, et al. The Alpha Magnetic Spectrometer (AMS) on the international space station: Part II — Results from the first seven years. Physics Reports. 2021;894:1-116. The Alpha Magnetic Spectrometer (AMS) on the International Space Station: Part II - Results from the First Seven Years. Available from: https://www.sciencedirect.com/science/article/pii/S0370157320303434.

[24] Lübelsmeyer K, Schultz von Dratzig A, Wlochal M, Ambrosi G, Azzarello P, Battiston R, et al. Upgrade of the Alpha Magnetic Spectrometer (AMS-02) for long term operation on the International Space Station (ISS). Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2011;654(1):639-48. Available from: https://www.sciencedirect.com/science/article/pii/S0168900211011752.

[25] Ambrosi G, Burger WJ, Oliva A. The AMS-02 silicon tracker: Recent results and current status. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2010;617(1):471-2. 11th Pisa Meeting on Advanced Detectors. Available from: https://www.sciencedirect.com/science/article/pii/S0168900209018919.

[26] Kirn T. The AMS-02 TRD on the international space station. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2013;706:43-7. TRDs for the Third Millenium. Available from: https://www.sciencedirect.com/science/article/pii/S0168900212004974.

[27] Bindi V, Chen GM, Chen HS, Choumilov E, Choutko V, Contin A, et al. Calibration and performance of the AMS-02 time of flight detector in space. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2014;743:22-9. Available from: https://www.sciencedirect.com/science/article/pii/S0168900214000102.

[28] Giovacchini F. The RICH detector of the AMS-02 experiment aboard the International Space Station. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2023;1055:168434. Available from: https://www.sciencedirect.com/science/article/pii/S0168900223004242.

[29] Adloff C, Basara L, Bigongiari G, Bosi F, Brun P, Cadoux F, et al. The AMS-02 lead-scintillating fibres Electromagnetic Calorimeter. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2013;714:147-54. Available from: https://www.sciencedirect.com/science/article/pii/S016890021300199X.

[30] Aguilar M, Aisa D, Alvino A, Ambrosi G, Andeen K, Arruda L, et al. Electron and Positron Fluxes in Primary Cosmic Rays Measured with the Alpha Magnetic Spectrometer on the International Space Station. Phys Rev Lett. 2014 Sep;113:121102. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.113.121102.

[31] Accardo L, Aguilar M, Aisa D, Alpat B, Alvino A, Ambrosi G, et al. High Statistics Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5–500 GeV with the Alpha Magnetic Spectrometer on the International Space Station. Phys Rev Lett. 2014 Sep;113:121101. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.113.121101.

[32] Aguilar M, Ali Cavasonza L, Ambrosi G, Arruda L, Attig N, Azzarello P, et al. Towards Understanding the Origin of Cosmic-Ray Positrons. Phys Rev Lett. 2019 Jan;122:041102. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.122.041102.

[33] Adriani O, Barbarino GC, Bazilevskaya GA, Bellotti R, Boezio M, Bogomolov EA, et al. An anomalous positron abundance in cosmic rays with energies 1.5–100 GeV. Nature. 2009 Apr;458(7238):607-9. Available from: https://doi.org/10.1038/nature07942.

[34] Ambrosi G. AMS-L0: upgrade status and prospects; 2023. Advances in Space AstroParticle Physics (ASAPP) conference. Available from: https://indico.cern.ch/event/1208314/contributions/5283387/.

[35] Aguilar M, Alberti G, Alpat B, Alvino A, Ambrosi G, Andeen K, et al. First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5–350

GeV. Phys Rev Lett. 2013 Apr;110:141102. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.110.141102.

[36] Aguilar M, Ali Cavasonza L, Alpat B, Ambrosi G, Arruda L, Attig N, et al. Towards Understanding the Origin of Cosmic-Ray Electrons. Phys Rev Lett. 2019 Mar;122:101101. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.122.101101.

[37] Abdo AA, Ackermann M, Ajello M, Atwood WB, Axelsson M, Baldini L, et al. Measurement of the Cosmic Ray $e^+ + e^-$ Spectrum from 20 GeV to 1 TeV with the Fermi Large Area Telescope. Phys Rev Lett. 2009 May;102:181101. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.102.181101.

[38] Atwood WB, Abdo AA, Ackermann M, Althouse W, Anderson B, Axelsson M, et al. The large area telescope on thefermi gamma-ray space telescopemission. Astrophys J. 2009 Jun;697(2):1071-102.

[39] Meegan C, Lichti G, Bhat PN, Bissaldi E, Briggs MS, Connaughton V, et al. The Fermi Gamma-ray Burst Monitor. apj. 2009 Sep;702(1):791-804.

[40] Ackermann M, Ajello M, Allafort A, Atwood WB, Baldini L, Barbiellini G, et al. Measurement of Separate Cosmic-Ray Electron and Positron Spectra with the Fermi Large Area Telescope. Phys Rev Lett. 2012 Jan;108:011103. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.108.011103.

[41] Akaike Y, Adriani O, Asano K, Asaoka Y, Berti E, Bigongiari G, et al. The Calorimetric Electron Telescope (CALET) on the International Space Station: Results from the first eight years on orbit. Advances in Space Research. 2024;74(9):4353-67. Progress in Astrophysics of Cosmic Rays. Available from: https://www.sciencedirect.com/science/article/pii/S027311772400382X.

[42] Torii S, Marrocchesi PS. The CALorimetric Electron Telescope (CALET) on the International Space Station. Advances in Space Research. 2019;64(12):2531-7. Advances in Cosmic-Ray Astrophysics and Related Areas. Available from: https://www.sciencedirect.com/science/article/pii/S027311771930273X.

[43] Adriani O, Akaike Y, Asano K, Asaoka Y, Bagliesi MG, Bigongiari G, et al. Energy Spectrum of Cosmic-Ray Electron and Positron from 10 GeV to 3 TeV Observed with the Calorimetric Electron Telescope on the International Space Station. Phys Rev Lett. 2017 Nov;119:181101. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.119.181101.

[44] Adriani O, Akaike Y, Asano K, Asaoka Y, Bagliesi MG, Berti E, 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. 2018 Jun;120:261102. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.120.261102.

[45] Adriani O, Akaike Y, Asano K, Asaoka Y, Berti E, Bigongiari G, et al. Direct Measurement of the Spectral Structure of Cosmic-Ray Electrons + Positrons in the TeV Region with CALET on the International Space Station. Phys Rev Lett. 2023 Nov;131:191001. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.131.191001.

[46] DAMPE Collaboration. Direct detection of a break in the teraelectronvolt cosmic-ray spectrum of electrons and positrons. Nature. 2017;552(7683):63-6.

[47] Perrina, Chiara. Progressing our understanding of cosmic rays with the HERD space-borne experiment. EPJ Web Conf. 2025;319:04003. Available from: https://doi.org/10.1051/epjconf/202531904003.

[48] Matthews J. A Heitler model of extensive air showers. Astroparticle Physics. 2005;22(5):387-97. Available from: https://www.sciencedirect.com/science/article/pii/S0927650504001598.

[49] Cherenkov PA. Visible luminescence of pure liquids under the influence of γ-radiation. Dokl Akad Nauk SSSR. 1934;2(8):451-4.

[50] de Naurois M. The H.E.S.S. experiment : current status and future prospects. PoS. 2019;ICRC2019:656.

[51] The Pierre Auger Cosmic Ray Observatory. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2015;798:172-213. Available from: https://www.sciencedirect.com/science/article/pii/S0168900215008086.

[52] Ma XH, Bi YJ, Cao Z, Chen MJ, Chen SZ, Cheng YD, et al. Chapter 1 LHAASO Instruments and Detector technology *. Chinese Physics C. 2022 mar;46(3):030001. Available from: https://dx.doi.org/10.1088/1674-1137/ac3fa6.

[53] Acharya BS, Actis M, Aghajani T, Agnetta G, Aguilar J, Aharonian F, et al. Introducing the CTA concept. Astroparticle Physics. 2013;43:3-18. Seeing the High-Energy Universe with the Cherenkov Telescope Array - The Science Explored with the CTA. Available from: https://www.sciencedirect.com/science/article/pii/S0927650513000169.

[54] Greisen K. End to the Cosmic-Ray Spectrum? Phys Rev Lett. 1966 Apr;16:748-50. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.16.748.

[55] Zatsepin GT, Kuz'min VA. Upper Limit of the Spectrum of Cosmic Rays. Soviet Journal of Experimental and Theoretical Physics Letters. 1966 Aug;4:78.

[56] Torres DF, Anchordoqui LA. Astrophysical origins of ultrahigh energy cosmic rays. Reports on Progress in Physics. 2004 jul;67(9):1663. Available from: https://dx.doi.org/10.1088/0034-4885/67/9/R03.

[57] Evoli C. The Cosmic-Ray Energy Spectrum. Zenodo; 2020. Available from: https://doi.org/10.5281/zenodo.7948212.

[58] Hillas AM. The Origin of Ultra-High-Energy Cosmic Rays. araa. 1984 jan;22:425-44.

[59] Thompson DJ, Baldini L, Uchiyama Y. Cosmic ray studies with the Fermi Gamma-ray Space Telescope Large Area Telescope. Astroparticle Physics. 2012;39-40:22-32. Cosmic Rays Topical Issue. Available from: https://www.sciencedirect.com/science/article/pii/S0927650511001964.

[60] Gabici S, Evoli C, Gaggero D, Lipari P, Mertsch P, Orlando E, et al. The origin of Galactic cosmic rays: Challenges to the standard paradigm. International Journal of Modern Physics D. 2019;28(15):1930022. Available from: https://doi.org/10.1142/S0218271819300222.

[61] Bethe HA. Energy Production in Stars. Phys Rev. 1939 Jan;55:103-3. Available from: https://link.aps.org/doi/10.1103/PhysRev.55.103.

[62] Burbidge EM, Burbidge GR, Fowler WA, Hoyle F. Synthesis of the Elements in Stars. Rev Mod Phys. 1957 Oct;29:547-650. Available from: https://link.aps.org/doi/10.1103/RevModPhys.29.547.

[63] Cerda-Duran P, Elias-Rosa N. In: Rezzolla L, Pizzochero P, Jones DI, Rea N, Vidaña I, editors. Neutron Stars Formation and Core Collapse Supernovae. Cham: Springer International Publishing; 2018. p. 1-56. Available from: https://doi.org/10.1007/978-3-319-97616-7_1.

[64] Hewish A, Bell SJ, Pilkington JDH, Scott PF, Collins RA. Observation of a Rapidly Pulsating Radio Source. nat. 1968 feb;217(5130):709-13.

[65] Völk, H J , Berezhko, E G , Ksenofontov, L T . Internal dynamics and particle acceleration in Tycho's SNR. A&A. 2008;483(2):529-35. Available from: https://doi.org/10.1051/0004-6361:20079337.

[66] Schroer B, Evoli C, Blasi P. TeV halos and the role of pulsar wind nebulae as sources of cosmic-ray positrons. Phys Rev D. 2023 Jun;107:123020. Available from: https://link.aps.org/doi/10.1103/PhysRevD.107.123020.

[67] Abraham J, Abreu P, Aglietta M, Aguirre C, Allard D, Allekotte I, et al. Correlation of the Highest-Energy Cosmic Rays with Nearby Extragalactic Objects. Science. 2007;318(5852):938-43. Available from: https://www.science.org/doi/abs/10.1126/science.1151124.

[68] Waxman E. Cosmological Gamma-Ray Bursts and the Highest Energy Cosmic Rays. Phys Rev Lett. 1995 Jul;75:386-9. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.75.386.

[69] Longair MS. High Energy Astrophysics. 3rd ed. Cambridge University Press; 2011.

[70] Fermi E. On the Origin of the Cosmic Radiation. Phys Rev. 1949 Apr;75:1169-74. Available from: https://link.aps.org/doi/10.1103/PhysRev.75.1169.

[71] Axford WI, Leer E, Skadron G. The Acceleration of Cosmic Rays by Shock Waves. In: International Cosmic Ray Conference. vol. 11 of International Cosmic Ray Conference; 1977. p. 132.

[72] Bell AR. The acceleration of cosmic rays in shock fronts – I. Monthly Notices of the Royal Astronomical Society. 1978 02;182(2):147-56. Available from: https://doi.org/10.1093/mnras/182.2.147.

[73] Arbutina B, Zeković V. Non-linear diffusive shock acceleration: A recipe for injection of electrons. Astroparticle Physics. 2021;127:102546. Available from: https://www.sciencedirect.com/science/article/pii/S0927650520301183.

[74] Gaisser TK, Engel R, Resconi E. Cosmic Rays and Particle Physics. 2nd ed. Cambridge University Press; 2016.

[75] Jones FC. Inverse Compton Scattering of Cosmic-Ray Electrons. Phys Rev. 1965 Mar;137:B1306-11. Available from: https://link.aps.org/doi/10.1103/PhysRev.137.B1306.

[76] Berezinskij VS, Ginzburg VL, Berezinskiĭ VS, Ginzburg VL. Astrophysics of cosmic rays. Amsterdam: North-Holland; 1990.

[77] Syrovatskii SI. The Distribution of Relativistic Electrons in the Galaxy and the Spectrum of Synchrotron Radio Emission. sovast. 1959;3:22.

[78] CODATA Value: electron mass — physics.nist.gov;. https://physics.nist.gov/cgi-bin/cuu/Value?me.

[79] Aguilar M, Aisa D, Alvino A, Ambrosi G, Andeen K, Arruda L, et al. Electron and Positron Fluxes in Primary Cosmic Rays Measured with the Alpha Magnetic Spectrometer on the International Space Station. Phys Rev Lett. 2014 Sep;113:121102. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.113.121102.

[80] Aharonian F, Benkhali FA, Aschersleben J, Ashkar H, Backes M, Martins VB, et al. High-Statistics Measurement of the Cosmic-Ray Electron Spectrum with H.E.S.S. Phys Rev Lett. 2024 Nov;133:221001. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.133.221001.

[81] Aguilar M, Aisa D, Alpat B, Alvino A, Ambrosi G, Andeen K, et al. Precision Measurement of the $(e^+ + e^-)$ Flux in Primary Cosmic Rays from 0.5 GeV to 1 TeV with the Alpha Magnetic Spectrometer on the International Space Station. Phys Rev Lett. 2014 Nov;113:221102. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.113.221102.

[82] Abdollahi S, Ackermann M, Ajello M, Atwood WB, Baldini L, Barbiellini G, et al. Cosmic-ray electron-positron spectrum from 7 GeV to 2 TeV with the Fermi Large Area Telescope. Phys Rev D. 2017 Apr;95:082007. Available from: https://link.aps.org/doi/10.1103/PhysRevD.95.082007.

[83] Adriani O, Barbarino GC, Bazilevskaya GA, Bellotti R, Bianco A, Boezio M, et al. Cosmic-Ray Positron Energy Spectrum Measured by PAMELA. Phys Rev Lett. 2013 Aug;111:081102. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.111.081102.

[84] Grimani C, Stephens SA, Cafagna FS, Basini G, Bellotti R, Brunetti MT, et al. Measurements of the absolute energy spectra of cosmic-ray positrons and electrons above 7 GeV. aap. 2002 Sep;392:287-94.

[85] Boezio M, Barbiellini G, Bonvicini V, Schiavon P, Vacchi A, Zampa N, et al. Measurements of cosmic-ray electrons and positrons by the Wizard/CAPRICE collaboration. Advances in Space Research. 2001;27(4):669-74. Available from: https://www.sciencedirect.com/science/article/pii/S0273117701001089.

[86] Aguilar M, Alcaraz J, Allaby J, Alpat B, Ambrosi G, Anderhub H, et al. Cosmic-ray positron fraction measurement from 1 to 30 GeV with AMS-01. Physics Letters B. 2007;646(4):145-54. Available from: https://www.sciencedirect.com/science/article/pii/S0370269307001104.

[87] DuVernois MA, Barwick SW, Beatty JJ, Bhattacharyya A, Bower CR, Chaput CJ, et al. Cosmic-Ray Electrons and Positrons from 1 to 100 GeV: Measurements with HEAT and Their Interpretation. The Astrophysical Journal. 2001 sep;559(1):296. Available from: https://dx.doi.org/10.1086/322324.

[88] Boulares A. The Nature of the Cosmic-Ray Electron Spectrum, and Supernova Remnant Contributions. apj. 1989 Jul;342:807.

[89] Blasi P. The origin of galactic cosmic rays. aapr. 2013 Nov;21:70.

[90] Komori Y. Propagation of cosmic-ray electrons in the Galaxy. Advances in Space Research. 2006;37(10):1918-22. Astrophysics. Available from: https://www.sciencedirect.com/science/article/pii/S0273117705009543.

[91] Di Mauro M, Donato F, Manconi S. Novel interpretation of the latest AMS-02 cosmic-ray electron spectrum. Phys Rev D. 2021 Oct;104:083012. Available from: https://link.aps.org/doi/10.1103/PhysRevD.104.083012.

[92] Zwicky F. Republication of: The redshift of extragalactic nebulae. General Relativity and Gravitation. 2009 Jan;41(1):207-24.

[93] Hubble E, Humason ML. The Velocity-Distance Relation among Extra-Galactic Nebulae. apj. 1931 Jul;74:43.

[94] Rubin VC, Ford WK Jr, Thonnard N. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc). apj. 1980 Jun;238:471-87.

[95] Bertone G, Hooper D. History of dark matter. Rev Mod Phys. 2018 Oct;90:045002. Available from: https://link.aps.org/doi/10.1103/RevModPhys.90.045002.

[96] Chen JY, Kolb EW, Wang LT. Dark matter coupling to electroweak gauge and Higgs bosons: An effective field theory approach. Physics of the Dark Universe. 2013;2(4):200-18. Available from: https://www.sciencedirect.com/science/article/pii/S2212686413000356.

[97] Workman RL, Others. Review of Particle Physics. PTEP. 2022;2022:083C01.

[98] Bozek B, Marsh DJE, Silk J, Wyse RFG. Galaxy UV-luminosity function and reionization constraints on axion dark matter. Monthly Notices of the Royal Astronomical Society. 2015 04;450(1):209-22. Available from: https://doi.org/10.1093/mnras/stv624.

[99] Monroy-Rodríguez MA, Allen C. The end of the macho era, revisited: New limits on macho masses from halo wide binaries. Astrophys J. 2014 Jul;790(2):159.

[100] De Baerdemaeker S. Method-Driven Experiments and the Search for Dark Matter. Philosophy of Science. 2021;88(1):124–144.

[101] De Roeck A. Dark matter searches at accelerators. Nuclear Physics B. 2024;1003:116480. Special Issue of Nobel Symposium 182 on Dark Matter. Available from: https://www.sciencedirect.com/science/article/pii/S0550321324000464.

[102] Aprile E, Aalbers J, Abe K, Maouloud SA, Althueser L, Andrieu B, et al. The XENONnT dark matter experiment. The European Physical Journal C. 2024;84(8):784. Available from: https://doi.org/10.1140/epjc/s10052-024-12982-5.

[103] Aprile E, Aalbers J, Agostini F, Alfonsi M, Althueser L, Amaro FD, et al. Excess electronic recoil events in XENON1T. Phys Rev D. 2020 Oct;102:072004. Available from: https://link.aps.org/doi/10.1103/PhysRevD.102.072004.

[104] Shen ZQ, Xia ZQ, Fan YZ. Search for line-like and box-shaped spectral features from nearby galaxy clusters with 11.4 years of Fermi large area telescope data. Astrophys J. 2021 Oct;920(1):1.

[105] Alemanno F, An Q, Azzarello P, Tiziana Barbato FC, Bernardini P, Bi XJ, et al. Search for gamma-ray spectral lines with the DArk Matter Particle Explorer. Science Bulletin. 2022;67(7):679-84. Available from: https://www.sciencedirect.com/science/article/pii/S2095927321007672.

[106] Stecker FW, Tylka AJ. The Cosmic-Ray Antiproton Spectrum from Dark Matter Annihilation and Its Astrophysical Implications: A New Look. apjl. 1989 Jan;336:L51.

[107] Cholis I, Goodenough L, Hooper D, Simet M, Weiner N. High energy positrons from annihilating dark matter. Phys Rev D. 2009 Dec;80:123511. Available from: https://link.aps.org/doi/10.1103/PhysRevD.80.123511.

[108] Pohl M. Cosmic-ray electron signatures of dark matter. Phys Rev D. 2009 Feb;79:041301. Available from: https://link.aps.org/doi/10.1103/PhysRevD.79.041301.

[109] Kopp J. Constraints on dark matter annihilation from AMS-02 results. Phys Rev D. 2013 Oct;88:076013. Available from: https://link.aps.org/doi/10.1103/PhysRevD.88.076013.

[110] Chang J, Ambrosi G, An Q, Asfandiyarov R, Azzarello P, Bernardini P, et al. The DArk Matter Particle Explorer mission. Astroparticle Physics. 2017;95:6-24.

[111] Alemanno F, Altomare C, An Q, Azzarello P, Barbato FCT, Bernardini P, et al. Measurement of the cosmic $p + \mathrm{He}$ energy spectrum from 50 GeV to 0.5 PeV with the DAMPE space mission. Phys Rev D. 2024 Jun;109:L121101. Available from: https://link.aps.org/doi/10.1103/PhysRevD.109.L121101.

[112] DAMPE Collaboration. Detection of spectral hardenings in cosmic-ray boron-to-carbon and boron-to-oxygen flux ratios with DAMPE. Science Bulletin. 2022;67(21):2162-6. Available from: https://www.sciencedirect.com/science/article/pii/S2095927322004492.

[113] Collaboration D, An Q, Asfandiyarov R, Azzarello P, Bernardini P, Bi XJ, et al. Measurement of the cosmic ray proton spectrum from 40 GeV to 100 TeV with the DAMPE satellite. Science Advances. 2019;5(9):eaax3793. Available from: https://www.science.org/doi/abs/10.1126/sciadv.aax3793.

[114] Collaboration D, Alemanno F, Altomare C, An Q, Azzarello P, Barbato FCT, et al.. Observation of a spectral hardening in cosmic ray boron spectrum with the DAMPE space mission; 2024. Available from: https://arxiv.org/abs/2412.11460.

[115] Duan KK, Jiang W, Li X, Shen ZQ, Xu ZL, Alemanno F, et al. Observations of gamma-ray sources with DAMPE. PoS. 2021;ICRC2021:631.

[116] Yu Y, Sun Z, Su H, Yang Y, Liu J, Kong J, et al. The plastic scintillator detector for DAMPE. Astroparticle Physics. 2017;94:1-10. Available from: https://www.sciencedirect.com/science/article/pii/S0927650516302031.

[117] Azzarello P, Ambrosi G, Asfandiyarov R, Bernardini P, Bertucci B, Bolognini A, et al. The DAMPE silicon–tungsten tracker. Nuclear Instruments and Methods in

Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2016;831:378-84. Proceedings of the 10th International "Hiroshima" Symposium on the Development and Application of Semiconductor Tracking Detectors.

[118] Zhang Z, Wang C, Dong J, Wei Y, Wen S, Zhang Y, et al. The calibration and electron energy reconstruction of the BGO ECAL of the DAMPE detector. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2016;836:98-104.

[119] Huang YY, Ma T, Yue C, Zhang Y, Cai MS, Chang J, et al. Calibration and performance of the neutron detector onboard of the DAMPE mission. Research in Astronomy and Astrophysics. 2020 sep;20(9):153.

[120] Ding M, Zhang YP, Zhang YJ, Wang YP, Dong TK, Benedittis AD, et al. Calibration of the DAMPE Plastic Scintillator Detector and its on-orbit performance. Research in Astronomy and Astrophysics. 2019 mar;19(3):047.

[121] Dong T, Zhang Y, Ma P, Zhang Y, Bernardini P, Ding M, et al. Charge measurement of cosmic ray nuclei with the plastic scintillator detector of DAMPE. Astroparticle Physics. 2019;105:31-6. Available from: https://www.sciencedirect.com/science/article/pii/S0927650518301786.

[122] Atwood WB, Bagagli R, Baldini L, Bellazzini R, Barbiellini G, Belli F, et al. Design and initial tests of the Tracker-converter of the Gamma-ray Large Area Space Telescope. Astroparticle Physics. 2007;28(4):422-34. Available from: https://www.sciencedirect.com/science/article/pii/S0927650507001302.

[123] Catanzani E, Ambrosi G, Azzarello P, Perrina C, Ionica M, Tykhonov A, et al. Study of the performances of the DAMPE silicon-tungsten tracker after five years of mission. Journal of Physics: Conference Series. 2022 nov;2374(1):012067. Available from: https://dx.doi.org/10.1088/1742-6596/2374/1/012067.

[124] Turchetta R. Spatial resolution of silicon microstrip detectors. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 1993;335(1):44-58. Available from: https://www.sciencedirect.com/science/article/pii/016890029390255G.

[125] Qiao R, Peng WX, Guo DY, Zhao H, Wang HY, Gong K, et al. Charge reconstruction of the DAMPE Silicon–Tungsten Tracker: A preliminary study with ion beams. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2018;886:48-52. Available from: https://www.sciencedirect.com/science/article/pii/S0168900218300093.

[126] Viehhauser G, Weidberg T. Detectors in particle physics. Boca Raton: CRC Press; 2024.

[127] Wei Y, Zhang Y, Zhang Z, Wu L, Wen S, Dai H, et al. Performance of the DAMPE BGO calorimeter on the ion beam test. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2019;922:177-84. Available from: https://www.sciencedirect.com/science/article/pii/S0168900218318333.

[128] Zhang YQ, Guo JH, Liu Y, Feng CQ, Zhang YL, Dong TK, et al. Design and on-orbit status of the trigger system for the DAMPE mission. Research in Astronomy and Astrophysics. 2019 sep;19(9):123. Available from: https://dx.doi.org/10.1088/1674-4527/19/9/123.

[129] Agostinelli S, Allison J, Amako K, Apostolakis J, Araujo H, Arce P, et al. Geant4—a simulation toolkit. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2003;506(3):250-303. Available from: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[130] Kaidalov AB, Ter-Martirosyan KA. Pomeron as quark-gluon strings and multiple hadron production at SPS-Collider energies. Physics Letters B. 1982;117(3):247-51. Available from: https://www.sciencedirect.com/science/article/pii/0370269382905561.

[131] Wright DH, Kelsey MH. The Geant4 Bertini Cascade. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2015;804:175-88. Available from: https://www.sciencedirect.com/science/article/pii/S0168900215011134.

[132] Hussein MT, El-Harby N, Hassan MH. High Energy Nucleus—Nucleus Collisions in View of the Multi-Peripheral Model. Acta Physica Hungarica A) Heavy Ion Physics. 2001 Jun;13(4):277-88. Available from: https://doi.org/10.1556/APH.13.2001.4.10.

[133] Pierog T, Karpenko I, Katzy JM, Yatsenko E, Werner K. EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider. Phys Rev C. 2015 Sep;92:034906. Available from: https://link.aps.org/doi/10.1103/PhysRevC.92.034906.

[134] Ulrich R, Pierog T, Baus C. Cosmic Ray Monte Carlo package, CRMC. Zenodo; 2021.

[135] Tykhonov A, Droz D, Yue C, Cui M, Li X, Wu X. TeV–PeV hadronic simulations with DAMPE. PoS. 2019;ICRC2019:143.

[136] Albertsson K, Altoe P, Anderson D, Andrews M, Araque Espinosa JP, Aurisano A, et al. Machine Learning in High Energy Physics Community White Paper. J Phys: Conf Ser. 2018;1085:022008.

[137] Xu Z, Li X, Cui M, Yue C, Jiang W, Li W, et al. An Unsupervised Machine Learning Method for Electron–Proton Discrimination of the DAMPE Experiment. Prepint. 2022;8(11):570.

[138] Droz D, Tykhonov A, Wu X, Alemanno F, Ambrosi G, Catanzani E, et al. A neural network classifier for electron identification on the DAMPE experiment. J Inst. 2021;16(07):P07036.

[139] Tykhonov A, Kotenko A, Coppin P, Deliyergiyev M, Droz D, Frieden JM, et al. A deep learning method for the trajectory reconstruction of cosmic rays with the DAMPE mission. Astroparticle Physics. 2023;146:102795.

[140] Stolpovskiy M, Wu X, Tykhonov A, Deliyergiyev M, Perrina C, Salinas MM, et al. Machine learning-based method of calorimeter saturation correction for helium flux analysis with DAMPE experiment. Journal of Instrumentation. 2022;17(06):P06031.

[141] Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016. Available from: `http://www.deeplearningbook.org`.

[142] Dettmers T. Understanding Convolution in Deep Learning;. Available from: `https://timdettmers.com/2015/03/26/convolution-deep-learning/`.

[143] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research. 2014;15(56):1929-58. Available from: `http://jmlr.org/papers/v15/srivastava14a.html`.

[144] Bottou L. Stochastic Learning. In: Bousquet O, von Luxburg U, editors. Advanced Lectures on Machine Learning. Lecture Notes in Artificial Intelligence, LNAI 3176. Berlin: Springer Verlag; 2004. p. 146-68. Available from: `http://leon.bottou.org/papers/bottou-mlss-2004`.

[145] Putti-Garcia E, Tykhonov A, Kotenko A, Boutin H, Li M, Coppin P, et al.. Energy Reconstruction of Non-fiducial Electron-Positron Events in the DAMPE Experiment Using Convolutional Neural Networks; 2025. Submitted to JINST.

[146] Evoli C, Amato E, Blasi P, Aloisio R. Galactic factories of cosmic-ray electrons and positrons. Phys Rev D. 2021;103:083010.

[147] Turner MS, Wilczek F. Positron line radiation as a signature of particle dark matter in the halo. Phys Rev D. 1990;42:1001-7.

[148] Conrad J, Reimer O. Indirect dark matter searches in gamma and cosmic rays. Nature Physics. 2017 Mar;13(3):224-31. Available from: `https://doi.org/10.1038/nphys4049`.

[149] Yue C, Zang J, Dong T, Li X, Zhang Z, Zimmer S, et al. A parameterized energy correction method for electromagnetic showers in BGO-ECAL of DAMPE. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2017;856:11-6. Available from: https://www.sciencedirect.com/science/article/pii/S0168900217303406.

[150] Droz D. Electron Identification with Neural Networks and Measurement of the TeV Cosmic Electrons Flux with the DAMPE Experiment [PhD thesis]. Université de Genève; 2021.

[151] Carleo G, Cirac I, Cranmer K, Daudet L, Schuld M, Tishby N, et al. Machine learning and the physical sciences. Reviews of Modern Physics. 2019 Dec;91:045002. Available from: https://link.aps.org/doi/10.1103/RevModPhys.91.045002.

[152] Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al.. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems; 2015. Available from: https://www.tensorflow.org/.

[153] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In: International Conference on Learning Representations; 2015. .

[154] Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. In: Pereira F, Burges CJ, Bottou L, Weinberger KQ, editors. Advances in Neural Information Processing Systems. vol. 25. Curran Associates, Inc.; 2012. Available from: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

[155] Kingma DP, Ba J. Adam: A Method for Stochastic Optimization; 2017.

[156] Navas S, et al. Review of particle physics. Phys Rev D. 2024;110(3):030001.

[157] Bourbeau J, Hampel-Arias Z. PyUnfold: A Python package for iterative unfolding. The Journal of Open Source Software. 2018 Jun;3(26):741. Available from: https://doi.org/10.21105/joss.00741.

[158] Azzarello P, Ambrosi G, Asfandiyarov R, Bernardini P, Bertucci B, Bolognini A, et al. The DAMPE silicon–tungsten tracker. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2016;831:378-84. Proceedings of the 10th International "Hiroshima" Symposium on the Development and Application of Semiconductor Tracking Detectors. Available from: https://www.sciencedirect.com/science/article/pii/S0168900216002539.

[159] Putti-Garcia E, Li M, Tykhonov A, Kotenko A, Boutin H, Coppin P, et al.. Convolutional Method for Identification of Non-fiducial electrons with the DAMPE experiment; 2025. To be submitted to JINST.

[160] Aharonian F, Akhperjanian AG, Barres de Almeida U, Bazer-Bachi AR, Becherini Y, Behera B, et al. Energy Spectrum of Cosmic-Ray Electrons at TeV Energies. Phys Rev Lett. 2008 Dec;101:261104. Available from: https://link.aps.org/doi/10.1103/PhysRevLett.101.261104.

[161] Archer A, Benbow W, Bird R, Brose R, Buchovecky M, Buckley JH, et al. Measurement of cosmic-ray electrons at TeV energies by VERITAS. Phys Rev D. 2018 Sep;98:062004. Available from: https://link.aps.org/doi/10.1103/PhysRevD.98.062004.

[162] Ackermann M, Ajello M, Allafort A, Baldini L, Ballet J, Barbiellini G, et al. Detection of the Characteristic Pion-Decay Signature in Supernova Remnants. Science. 2013;339(6121):807-11. Available from: https://www.science.org/doi/abs/10.1126/science.1231160.

[163] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is All you Need. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, et al., editors. Advances in Neural Information Processing Systems. vol. 30. Curran Associates, Inc.; 2017. Available from: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.