Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

Finley Alexander Quinton

Studying the arrival directions of ultra-high energy cosmic rays with deep neural networks

Master's thesis in Physics
Supervisor: Foteini Oikonomou
June 2023

NTNU
Norwegian University of
Science and Technology

Finley Alexander Quinton

Studying the arrival directions of ultra-high energy cosmic rays with deep neural networks

Master's thesis in Physics
Supervisor: Foteini Oikonomou
June 2023

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

NTNU
Norwegian University of
Science and Technology

Abstract

The origin of ultra-high energy cosmic rays (UHECRs) has remained a mystery for over 100 years and the search for their sources is one of the biggest ongoing challenges in astroparticle physics. As UHECRs are charged particles they are deflected by extragalactic and Galactic magnetic fields during propagation in the Universe, resulting in their arrival directions not aligning with their source directions. If however, a significant enough fraction of UHECRs originate from a common source, magnetically-induced patterns in their arrival directions are expected. These might lead us to the discovery of their sources.

This work presents a method on how deep neural networks can be used to search for such patterns in an all-sky analysis by using a simulated flux of the most energetic UHECRs. It is shown that by using variations of the Galactic magnetic field model for the creation of the simulated data, the neural network can be trained to predict source events exposed to a different inherently unknown deflection. The analysis is applied to different mass composition scenarios where it shows excellent performance for a small fraction of proton source events embedded in a large isotropic background. Despite the neural network performing worse on heavier nuclei, it retains a substantial detection ability for a higher multiplicity of source events. When applied to data taken by the Pierre Auger Observatory, two potential regions of interest are identified, showing a noticeably higher probability of arising in a powerful astrophysical source.

CONTENTS

List of Figures

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The existence of charged particles arriving at Earth from all directions has been known for over 100 years now. The discovery of these cosmic rays (CRs) in 1912 by Victor Hess [1] won him the Nobel Prize in 1936. The most energetic form of these particles above 1 EeV = 10^{18} eV are called ultra-high energy cosmic rays (UHECRs) and are the most energetic particles ever detected by mankind. When they hit the Earth's atmosphere, they create air showers consisting of a cascade of secondary particles. The world's largest ground-based air shower observatory, Pierre Auger (Auger), measures the particles at ground level to reconstruct the properties of the primary particle before interaction with the atmosphere. Yet, in spite of great efforts and progress in this field, it is to this day still unclear which sources could create particles of these energies and where these sources are located.

When studying UHECRs, their initial arrival directions are either compared to known source candidates or analysed for possible anisotropies. Although their arrival directions are highly isotropic at almost all energies, some small and large-scale anisotropies have been found at the highest energies. Since UHECRs are charged nuclei ranging up to iron, they are deflected in extragalactic magnetic fields, but predominantly by the Galactic Magnetic Field (GMF). Therefore, the arrival directions of UHECRs here on Earth do not directly point back to their origin. When studying the propagation of UHECRs through the Galactic Magnetic Field, rigidity, defined as the energy over charge ratio, is the key quantity to describe the magnitude of deflection. If a few sources would contribute an appreciable UHECR flux here on Earth, we would expect an energy ordering in the arrival direction of the UHECRs since magnetic deflections are inversely proportional to the rigidity of each particle. In such an instance, UHECRs of the highest energies would show the lowest order of deflection and a higher degree of correlation with their sources. Here, machine learning could help in finding such magnetically-induced patterns in their arrival directions.
In this thesis, deep neural networks are trained on simulations of UHECRs above an

energy of 32 EeV as they would be observed at Auger. For these simulations, the state-of-the-art software framework CRPropa3 [2] is used.

Since the exact nature of the GMF is not known several variations of the to-date best GMF model from Jansson and Farrar (JF12) [3] are created for the machine learning applications. Furthermore, the performance of the neural network (NN) is tested and compared for several different mass compositions and finally applied to actual data taken by Auger where it possibly identifies clusters of events that could plausibly be attributed to astrophysical sources.
This work begins with an introduction to the observed properties of UHECRs on Earth in Chapter 2, followed by a summary of the most important interaction during propagation, possible kinds of source candidates, and the modulation of the GMF in Chapter 3. The different steps for simulating an UHECR flux as could be observed on Earth and the creation of the different GMF models are explained in Chapter 4. A general overview of the relevant machine learning methodology needed to create functional deep neural networks for the created data set is introduced in Chapter 5. The results of their application for different mass compositions are presented and discussed in Chapter 6 with a final conclusion and outlook in Chapter 7.

CHAPTER 2

UHECRs at Earth

The flux of CRs reaching Earth is highly dependent on their energy. CRs with energy below $\sim 10^{15}$ eV can be directly measured by satellites [4] or high-altitude balloons, the latter being how they were discovered in 1912 by Victor Hess [1]. CRs on the high-energy side of the CR spectrum, on the other hand, need much larger ground-based instruments for a reasonable detection rate, due to their lower flux. In this chapter the observed properties of CRs observed here on Earth are briefly introduced, starting with their incident flux on Earth in Section 2.1, followed by the latest inferences on their composition in Section 2.2, and their arrival-direction distribution in Section 2.3. The basic concepts and equations of CR physics are based on reference [5].

2.1. Cosmic ray energy spectrum at Earth

Because CRs interact with particles in the Earth's atmosphere, creating vast air-showers of secondary particles in the process, ground-based detectors need to reconstruct the initial state particle energy and composition prior to the interactions [6]. The observed energy spectrum of the CRs can be well approximated by a power-law function with a differential flux

$$F(E) = \frac{\mathrm{d}N}{\mathrm{d}E\,\mathrm{d}t\,\mathrm{d}A\,\mathrm{d}\Omega} E = E^{-\gamma}, \tag{2.1}$$

where $\mathrm{d}N$ is the number of CRs passing though a surface element $\mathrm{d}A$ with solid angle $\mathrm{d}\Omega$ in a time interval $\mathrm{d}t$ and energy $\mathrm{d}E$. Measurements of several different air-shower experiments are shown Figure 2.1. Here the flux was multiplied with a factor of $E^{-2.6}$ to stress the spectral power law dependency of the distribution and highlight certain observed points where the power law is broken. These points are labelled as *knee*, *second knee*, and *ankle*.

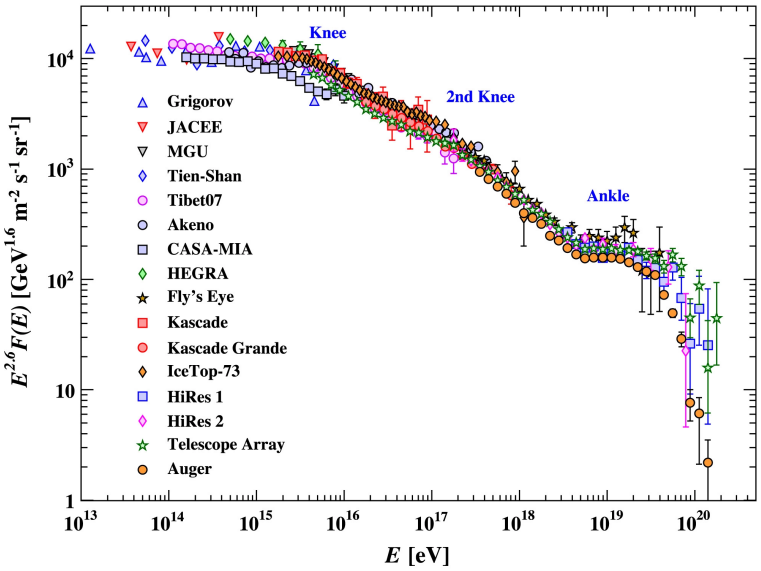
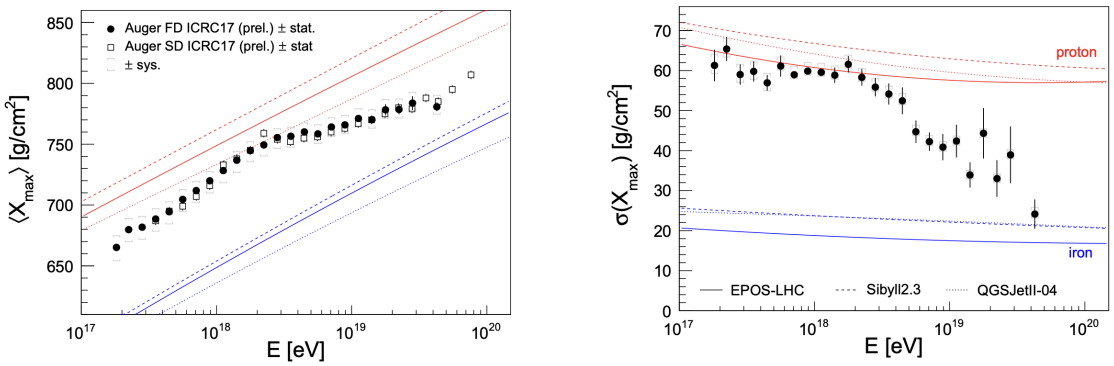


Figure 2.1.: Measurements of the CR spectrum from several different air-shower experiments as a function of the energy-per-nucleus, taken from [7]. In order to highlight the spectral power-law dependency the energy is multiplied with a factor of $E^{-2.6}$.

The nature of these points is still the subject of ongoing research and is still unresolved. The commonly accepted explanation for the knee is that it arises from a change from a Galactic CR origin for lower energies to an extragalactic one for higher energies [8]. The "second knee" could then be seen as the end of the Galactic origin of iron CRs [9]. The flattening at the "ankle" has been confirmed in [10] and may result from a dip structure of these highly energetic CRs and their interaction with the CMB during propagation, but this view has recently been disfavored [11]. A further steepening beyond 5×10^{19} EeV has also been confirmed in [10].

2.2. Composition of UHECRs

At the highest energies, the actual composition is not fully known yet, as the measurements are governed by many uncertainties when detecting the showers of secondary particles. However, advances have been made by extensive studies using e.g. deep learning to reconstruct the particle showers in the form of a PhD thesis in [12]. The composition of UHECRs is determined by measuring the atmospheric depth X_{max} at which a particle shower reaches the maximum number of constituents. As this depth scales $\propto \log(E/A)$ with energy E and atomic number A, lighter particles of the same energy penetrate deeper into the atmosphere than heavier ones. The reconstructed X_{max} values derived by measurement from fluorescence detectors (black dots) and the surface array (squares) at Auger are shown in Figure 2.2a together with the width $\sigma(X_{\mathrm{max}})$ of the distribution in Figure 2.2b.



(a) Atmospheric shower depth

(b) Standard deviation

Figure 2.2.: The mean of atmospheric shower depth $X_{\max}$ (a) together with the standard deviation (b) from measurements of the fluorescence detectors (black dots) and surface arrays (squares) at Auger. The simulated values for protons (red) and iron (blue) are shown as coloured lines. Taken from reference [13].

The black dots represent data points from Auger while the red line shows the prediction for a proton and the blue line for an iron-induced particle shower. It can be seen that the data suggests a heavier composition of the CRs towards higher energies. Just above an energy of 10^{18} eV, the composition changes from a proton-dominated composition towards Helium and Nitrogen which is in agreement with a possible observation of a Peters cycle [14]. If a source can accelerate CR protons up to a certain maximum energy, due to their lesser charge, protons will reach this limit before heavier nuclei with higher charge number Z. Thereby the energy spectrum from such a source on Earth would show a cutoff of protons first, then Helium and so on.

2.3. UHECR arrival directions

In addition to the energy and composition of UHECRs, their arrival directions are monitored and studied in order to find correlations with possible sources in the universe. Their distribution however is highly isotropic at almost all energies and shows no significant enough correlations with any of the assumed possible candidate sources. Searching for any anisotropies in the arrival-direction distribution is indispensable when looking for the origin of UHECRs.

2.3.1. Dipole anisotropy

Using an integrated exposure of $76,800\,\mathrm{km}^2$ a large-scale anisotropy has been discovered at energies beyond 8 EeV with a significance of 6.6 σ [15]. It can be described by a dipole with an amplitude of 6.5% at a Galactic longitude of -127° and latitude of -13° with an angular distance to the Galactic centre of 125°, shown in Figure 2.3. It is highly unlikely

that the sources at these energies are Galactic, as the magnetic field strength of the Milky Way could not displace the dipole by this amount even for heavier charged nuclei like iron.

If the UHECR sources have the same spatial distribution of galaxies, an alignment of the dipolar anisotropies derived from the 2 Micron All-Sky Redshift Survey (2MRS) [16] with the UHECR dipole is expected. It can be seen that the centre of the 2MRS measurements, denoted by a black diamond, indeed lies approximately within the direction of the UHECR observed dipole after accounting for the deflections in the GMF. This is indicated by the two arrows, representing a CR of energy 2 EeV and 5 EeV propagated using the JF12 magnetic field model, respectively. The deflection points towards the 68% and 95% confidence levels of the dipole are shown as black contours around the reconstructed centre of the UHECR dipole (black cross).

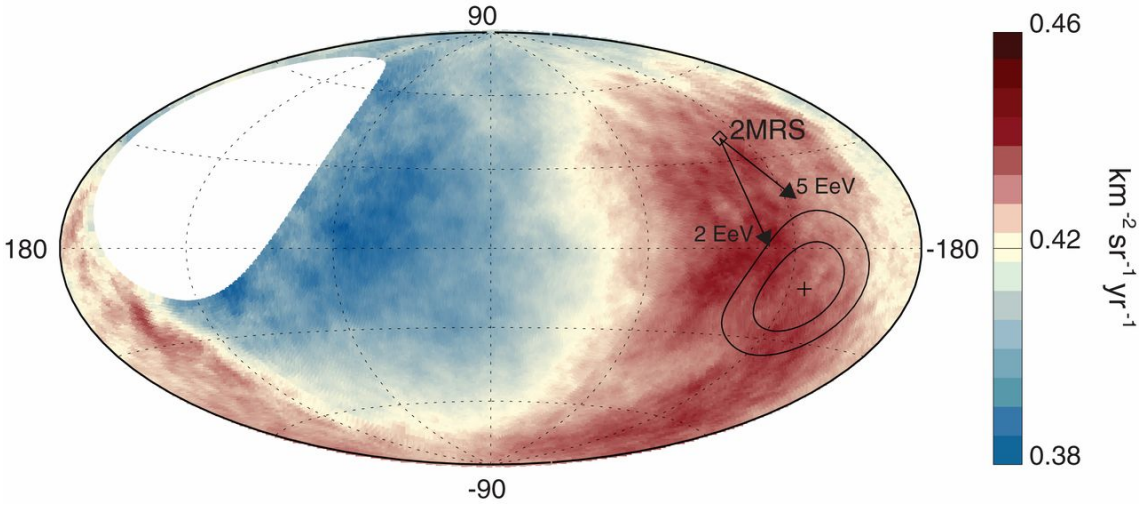


Figure 2.3.: UHECR flux in Galactic coordinates above an energy of 8 EeV and smoothed with a top-hat function of 45°. The black contours indicate the 68% and 95% confidence levels while the black cross marks the centre of the identified dipole. Also, the Two Micron All Sky Redshift Survey's (2MRS)[16] dipole of galaxies is shown together with the deflection in the Galactic magnetic field for UHECRs as a function for two different energies. Taken from [17].

Keeping in mind the large uncertainties in the magnetic field models [3], it is justified to assume that slight deviations are not in conflict with the galaxy origin interpretation. The explanation for this displacement is still unresolved. A contributor could be the movement of the Milky Way relative to that of the CR sources which is referred to as the Compton-Getting effect [18]. However, with an amplitude of 0.6%, the magnitude of this Doppler effect is significantly lower than the observed dipole. Other explanations are that it derives from an inhomogeneous source distribution [19, 20] or that the CRs originate from a dominant source and are diffused by extragalactic magnetic fields [21]. Nonetheless, all these theories and observations support an extragalactic origin of the sources above an energy of 8 EeV [17]. The dipole has been further studied in different

energy bins in [22, 23]. It was shown that the dipole becomes more visible towards higher energies. This is expected as CRs of higher energies experience a higher attenuation by cosmic photon fields (see Section 3.1.1) and thereby need to have a closer source origin. The measured dipole amplitude of the different energy bins is shown in the left panel of Figure 2.4 together with simulations of a uniform (grey) source distribution and one that follows the 2MRS (blue). The simulations show a high level of agreement with the data in both amplitude and energy dependence.

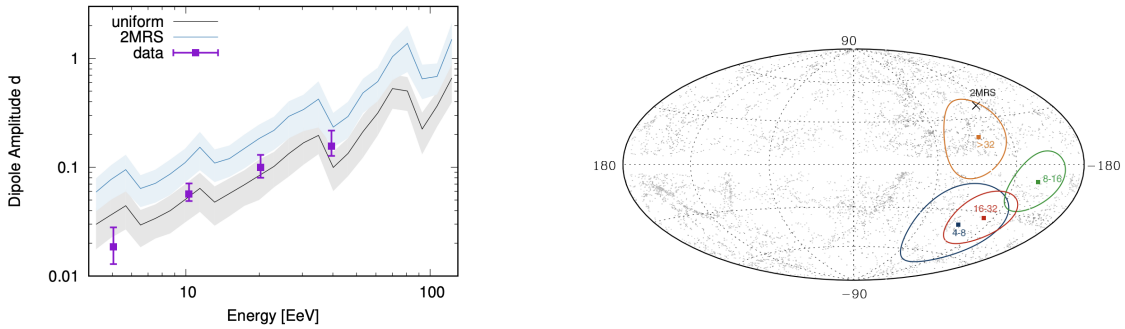


Figure 2.4.: The dipole amplitude as a function of energy above 4 EeV is shown together with predictions from [24] in the left panel. The right panel shows 2MRS galaxies within 100 Mpc distance with reconstructed dipole directions for different energy bins. Taken from [23].

Furthermore, in the right panel of Figure 2.4 the contours for the same four energy bins at 68% confidence level are shown. The grey dots correspond to local galaxies within 100 Mpc and the black cross to the direction of the weighted flux dipole of the 2MRS distribution of local galaxies. The reconstructed dipole maxima of the contours all point towards an extragalactic source origin. At the energy of the "ankle" at roughly 0.03 EeV it was shown in reference [25] that the maximum dipole points towards the Galactic centre, supporting the assumption of the transition of a Galactic to an extragalactic origin.

2.3.2. Smaller scale anisotropies

While large-scale anisotropies were discovered at energies between 4–32 EeV, smaller anisotropies become visible at even higher ones. At above 40 EeV, smaller angular anisotropies become more dominant, revealing a hot spot in the Centaurus region at Galactic coordinates $b = 19.4°$, $l = 309.5°$, see Figure 2.5.

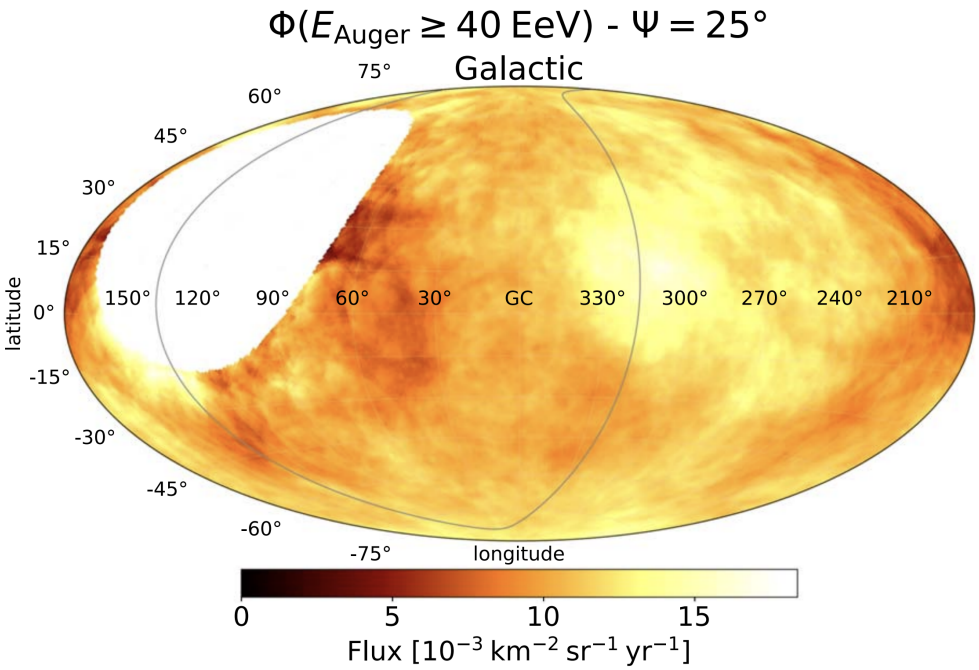


Figure 2.5.: UHECR flux above 40 EeV as seen from Auger in Galactic coordinates, taken from reference [26].

Nevertheless, one of the top candidates is the radio-loud galaxy Centaurus A (Cen A), which has been extensively studied for being the closest active Galactic nucleus (AGN) with a distance of $\sim$ 3.8 Mpc and as an AGN hosting possibly several possible acceleration sites, see Section 3.2.1, e.g. the different particle emission processes of Cen A were studied via simulations in [27]. It is often called the so-called "hot spot" and in e.g. [28, 29] excesses have been searched for in this region. A correlation with a significance of 3.9σ between the UHECR arrival directions and the direction of Cen A was found in [28]. Further, a significant deviation from isotropy at a level of 4.2σ is found for a sample of starburst galaxies [26].

CHAPTER 3

PROPAGATION AND ORIGIN OF UHECRS

When UHECRs propagate through the vastness of space they not only get deflected by the Galactic and extragalactic magnetic fields but also interact with different photon fields, losing energy in the process. This chapter will introduce the most relevant forms of energy losses in the studied energetic regime of this work in Section 3.1, followed by different acceleration mechanisms in Section 3.2 and possible candidate sources in Section 3.2.1. Finally, the influence on the arrival directions of UHECRs for different GMF models modulation will be explained in Section 3.3.

3.1. UHECR energy losses

The propagation of UHECRs through space alters their properties. While the magnetic fields are the origin of their deflection, interactions with the photon fields are the main cause of energy losses, setting limitations on the possible travel path of the most energetic particles in the UHECR spectrum.

3.1.1. Photon background fields

Even the seemingly most empty parts of the Universe are filled with diffuse radiation spanning across large ranges of the electromagnetic spectrum. The most influential ones for the CR propagation are shown in Figure 3.1. The interaction rate depends on the UHECR energy, the density of the photon field and the cross-section of the interaction process.

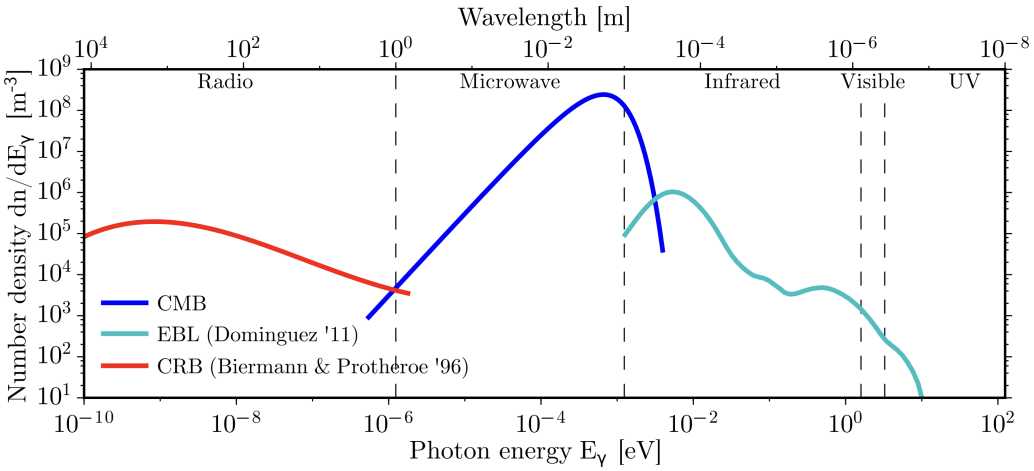


Figure 3.1.: Spectrum of the most relevant photon background radiation fields interacting with UHECR during their propagation through the Universe. Shown are the cosmic radio background (CRB) [30] in red, the cosmic microwave background (CMB) in blue and the extragalactic background light (EBL) [31] in turquoise. Taken from reference[32].

The least energetic one of the three is the **cosmic radio background (CRB)** radiation which dominates up to an energy of $\sim 10^{-6}$ eV. It is assumed that its origin is of extragalactic nature and its measurement thereby impeded by the stronger Galactic radio background signal present at Earth [30, 31]. Due to the very low energy of these photons, the interaction is only relevant for the most extremely energetic UHECRs and due to its lower number density than the CMB the interaction rate is significantly lower.

Of the three presented fields, the **cosmic microwave background (CMB)** has the highest number density with roughly 400 photons per cm^3 and is a remnant from the recombination epoch about 400,000 years after the Big Bang when the Universe first became transparent [33]. It can be described by an almost perfect thermal blackbody spectrum of temperature T = 2.726K [34] and is one of the most significant pieces of evidence for the Big Bang theory. The primary interaction is the GZK process [35] where a CR proton of $\geq 5 \times 10^{19}$ eV has sufficient enough energy to produce a pion with a CMB photon, see Section 3.1.2. For more insight into the different interactions between CRs and photons see the next section 3.1.2.
Third, the **extragalactic background light (EBL)** is made of the most energetic photons of the three presented background fields. It originates from stars and Active Galactic Nuclei (AGN) since the reionization phase of the Universe. Thereby, it is strongly correlated to the cosmological evolution which complicates its modelling as the indications of the evolution come mostly from measurements of the changing luminosities of galaxies populations at different redshifts [36]. The EBL shows a double peak in its spectrum which is caused by a large proportion of optical light being absorbed and re-emitted to the IGM like interstellar dust with a wavelength of 10^{-4}m.

3.1.2. Cosmic ray interactions with photons

GZK effect

In 1966 Greisen, Zatsepin and Kuzmin [35] postulated an effect, which describes the interaction between a UHECR proton of energy $\geq 5 \times 10^{19}$ eV with a CMB photon γ_{CMB} creating a pion via an extremely short-lived intermediate Δ^+ baryon. The two decay channels, which constitute approximately 100% of the total branching ratio, are shown in Equation (4.2).

$$p + \gamma_{\text{CMB}} \rightarrow \Delta^+ \rightarrow \begin{cases} p + \pi^0 \rightarrow p + 2\gamma \\ n + \pi^+ \rightarrow (p + e^- + \overline{\nu}_e) + (e^+ + \nu_e + \overline{\nu}_\mu) \end{cases} \tag{3.1}$$

It is assumed that the GZK effect is the cause of the strong cutoff in the UHECR flux and the end of the CR spectrum is therefore referred to as the GZK cutoff. However, this effect alone can not fully explain the UHECR data acquired by Auger [37]. This limits the possible propagation distance for the most energetic UHECR energy regime. Depending on the production rate, if the most energetic cosmic rays are of extragalactic nature, it is assumed that their sources must lie within a few hundred Mpc, see Figure 3.2.

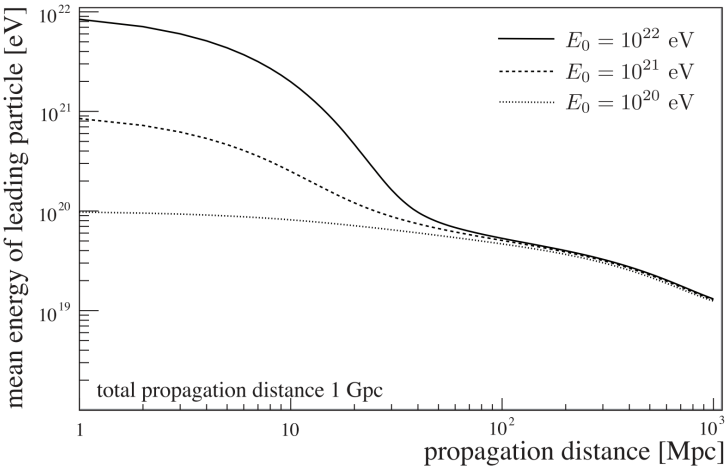


Figure 3.2.: Simulation of three different particles of different energies: 10^{20} eV (dotted line), 10^{21} eV (dashed line), and 10^{22} eV (solid line), as a function of propagated distance. Taken from reference [38].

Here the interaction of particles at three different energies was simulated as a function of propagated distance. It can be observed that all of them converge to the same mean energy of the leading particles after a travelled distance of just over 100 Mpc, making it impossible to distinguish from a particle of originally higher energy at even larger distances. This could also explain the pile-up at the ankle, seen in Figure 2.1.

Electron pair production

With the lowest energy threshold of the introduced processes, the electron pair production is dominant for UHECR proton energies of 10-50 $\times 10^{18}$ eV. The interaction can be written as

$$p + \gamma \rightarrow p + e^- + e^+ \tag{3.2}$$

and occurs when a photon γ converts into an electron–positron pair in the electromagnetic field of the proton.

Photon disintegration

When a nucleus X_A^Z with charge number Z and mass number A absorbs a photon γ, the photon can transform to another species by emitting a proton, neutron or even an α-particle. In case of proton emission, the process reads

$$X_A^Z + \gamma \rightarrow X_{A-1}^{Z-1} + p, \tag{3.3}$$

which is the dominant energy loss process for UHECR of higher mass.

3.2. Acceleration

So far, the different processes and interactions of UHECRs during propagation in space have been shown. It is still unclear, what mechanisms can accelerate these particles to such high energies initially. The acceleration methods can be divided into two main categories: top-down and bottom-up. The top-down structure corresponds to super-massive initial particles which create the UHECR due to their decay, while the bottom-up structure corresponds to charged CRs being accelerated within strong magnetic field regions, which can be further categorized in stochastic and direct acceleration mechanisms. One problem with the top-down method is that, if it is true, large photon fluxes will be expected, which is not in agreement with observations at these energies [39], resulting in a strong preference for the bottom-up scenario.

Direct acceleration

Direct acceleration considers the charged particles to be accelerated to their final energy via a "single shot". Only very few objects in the Universe could be powerful enough to create UHECRs of these energies in such a process. The most likely candidates are neutron stars, which not only have sufficiently strong magnetic fields but also rotation periods of only a few microseconds, thereby creating extremely powerful electric fields. They are assumed to be capable to accelerate protons to 10^{20} eV at young stages and nuclei like iron in later ones [40]. For more information on neutron stars, see Section 3.2.1.

Stochastic acceleration

The idea of charged particles being repeatedly deflected by magnetic clouds was originally proposed by Enrico Fermi in 1949 [41]. It is also referred to as the second-order Fermi acceleration as the energy gain of the particle is proportional to the square of the velocity of the magnetic cloud. These clouds move in random directions and can scatter elastically with charged particles causing them to either lose or gain energy through tail-on or head-on collisions, respectively, with each collision. As the head-on collisions are more probable the particle gains a continuous amount of energy $\frac{\langle \Delta E \rangle}{E} \propto \frac{4}{3}\beta^2$. The first-order Fermi acceleration occurs in shocks within the magnetic field. As the name states, the energy gain in this kind of acceleration is linear and given by $\frac{\langle \Delta E \rangle}{E} \propto \frac{4}{3}\beta$ [42, 43]. For the most energetic CRs, the first-order mechanism is assumed to be more likely as measurements of cloud densities do not fulfil the requirements for the second-order acceleration up to this energy regime [44]. After n interactions via either of the Fermi processes, a particle of initial energy E_0 will have been accelerated to an energy

$$E_n = \left(1 - \frac{\langle \Delta E \rangle}{E}\right)^n \times E_0. \tag{3.4}$$

With each encounter and thereby energy gain, the probability of the particle escaping the acceleration region p_{esc} increases, so that, for an energy $E_m > E_n$, the number of particles left can be estimated by

$$N(E_m > E_n) = N_0 \times \sum_{i=n}^{\infty} (1 - p_{esc})^i \propto \left(\frac{E_n}{E_o}\right)^{-\gamma}. \tag{3.5}$$

This clearly shows that the resulting spectrum has the form of a power-law function, falling off with an energy-dependent exponent of

$$\gamma = \frac{\log(1 - p_{\mathrm{esc}})}{\log(1 + \Delta E/E)} \approx \frac{p_{\mathrm{esc}}}{\langle \Delta E \rangle / E} \tag{3.6}$$

which has been derived via semi-analytical calculations and simulations to a value of $\gamma_{\mathrm{rel}} \approx 2.1 - 2.3$ for relativistic shocks [45, 46]. The acceleration limit is dependent on several parameters of the acceleration region. A particle of charge Ze within a magnetic field region of size R, magnetic field strength B, and velocity β of a scattering object can be accelerated to a maximum energy

$$E_{max}[EeV] = Ze \times B[\mu\mathrm{G}] \times R[\mathrm{kpc}] \times \beta. \tag{3.7}$$

This limitation on these source candidate parameters is called the **Hilas criterion** and was introduced by Anthony Michael Hillas in 1984 [47], shown in Equation (3.7) Using this equation together with the relatively large uncertainties on the different parameters the maximum reachable acceleration energy of different source candidates is shown in a Hillas diagram in Figure 3.3.

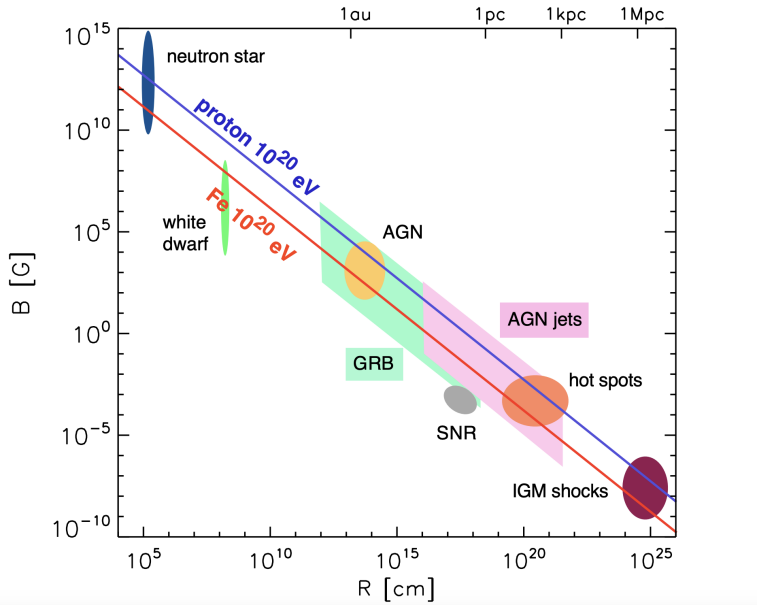


Figure 3.3.: Updated Hillas diagram showing different source candidates in a B-R phase space with their uncertainties on their given parameters. The red or blue line indicates the region in which iron or protons can be accelerated to 10^{20}eV. Taken from [48].

Here it can be seen that supernovae remnants do not have the sufficient capacity to accelerate UHECRs of any mass to energies of 10^{20}eV, while AGN might be able to accelerate both heavier charged particles like iron and even simple protons to this energy. When considering the parameters of the Milky Way with a radius of roughly 20 kpc and magnetic field strength in a range of μG, the UHECRs of the highest energy are unlikely to be of Galactic origin.

Although the Hillas criterion is a necessary criterion for any source candidate, it is definitely not a sufficient one. For instance, the energy loss the particle experiences via synchrotron radiation or photonuclear and hadronic interactions are not allowed to surpass the energy gain of the acceleration process.

3.2.1. UHECR candidate source classes

Although, for this work, only the source positions of four AGNs: Centaurus A (Cen A), Messier 87 (M87), Fornax A, Cygnus A (Cyg A), and two hot spots stated in [49] were considered as the positions of possible source candidates, a short description of all the different candidate classes seen in Figure 3.3 are given for completeness.

Active Galactic Nuclei

AGN are the most powerful known sources in the Universe and are also strong candidates for sources of the most energetic particles. Additionally, the AGN do not only

supply one but can have several possible acceleration sites as part of their structure. The AGN's source of power is a supermassive black hole (SMBH) at its centre creating a rotating accretion disk around it through its huge gravitational pull. The vicinity of this SMBH so-called vacuum gaps can be one possible accelerator of the UHECRs, but also the highly collimated twin jets could be their possible origin. Although not all AGNs create such a powerful jet and their exact nature is not fully understood, several regions within it could explain an acceleration of particles to the energies observed here on Earth. It is observed that these twin jets that outflow highly energetic matter in opposite directions consist of an inner and an outer larger scale jet, followed by large scale lobes [50] and a jet termination shock ("hot spot"). All of these regions could be the possible source of acceleration and some of them could also possibly work in tandem. Several theoretical acceleration mechanisms in these structures have been tested in [51, 52] Yet, AGNs and the structure of their jets need to be further studied in order to say which regions within them could be a possible source of the UHECRs and what exact acceleration method governs it. Especially multi-messenger astrophysics will play a key element in gaining a more detailed insight into these sources and their accelerating potential [53].

Gamma ray bursts

Gamma-ray bursts are highly energetic explosions created by an imploding massive star that becomes a black hole or a neutron star. They are named after their emitting wavelength of gamma rays and are the most luminous electromagnetic events since the Big Bang [54]. The burst can range from a few milliseconds to a couple of minutes, followed by an afterglow of less energetic radiation. They are rare events observed so far only outside our galaxy and could accelerate CRs via 2nd order Fermi acceleration to high enough energies. GRBs as candidate sources are extensively discussed in [55].

Neutron stars

Neutron stars are formed by the left-over dense core of a massive star which continues to collapse by its own gravitational pull. The core becomes so dense that protons coalesce with electrons and form neutrons, giving the neutron star its name. Most neutron stars rotate at extreme frequencies, those whose rotation we can observe are called pulsars, emitting radiation in regular intervals. A fraction of these pulsars may display a possible source for UHECRs, especially *magnetars* with huge magnetic and induced electric fields [56], releasing energy by emission of particle winds and radiation. In [40] it is argued that the young supernovae shell could prevent protons to escape the acceleration site up to energies of 10^{20} eV and iron nuclei even in later stages. As indicated in Figure 2.2, a heavier composition dominated by iron is assumed towards the high end of the UHECR spectrum, meaning a source would have to supply an injection of a large quantity of heavier nuclei from either a metal-rich environment or contain a large enough amount

of nucleosynthesis. While fireball-type GRBs most likely do not fulfil these requirements [57], AGNs can have super-solar metallicities, but not with a sufficient enough fraction above nitrogen [58]. However, a small proportion of iron-rich young neutron stars, that are embedded in specific supernovae, could explain the UHECR spectrum and composition observed on Earth [59]. Yet, due to their short lifetime, correlations between them and the UHECR arrival directions are not expected [48].

3.3. UHECRs in magnetic fields

When a ultra-relativistic particle of charge $q = Ze$ travels with velocity $\vec{v}$ within a magnetic field $\vec{B}$, it experiences a deflection perpendicular to both the magnetic field lines and the velocity. In the energy regime of UHECR, the rest mass can be neglected and the change of direction can be written as

$$\frac{d\vec{v}}{dt} = \frac{c^2}{R}(\vec{v} \times \vec{B}), \tag{3.8}$$

where $R = \frac{E}{Ze}$ is the rigidity of the particle. The rigidity is defined as the ratio of the energy over the charge and is thereby a measure of the magnitude of deflection of a CR in a magnetic field. Following this equation a CR of smaller rigidity experiences a larger deflection within a magnetic field and vice versa. Rigidity is the key quantity when describing the deflection of UHECRs in magnetic fields. For small deflections, the arrival direction of a charged particle Θ with an energy E can be approximated [60] with its initial source direction Θ_S as

$$\Theta = \Theta_S + \frac{Ze}{E}\int_0^L \mathrm{d}\vec{l} \times \vec{B}(\vec{l}). \tag{3.9}$$

An example of the deflection of particles of different rigidities between 5–100 EV originating from the same source is shown in Figure 3.4. We note that for simplicity no random magnetic field components, introduced in the next section, are used here, resulting in a linear deflection distribution.

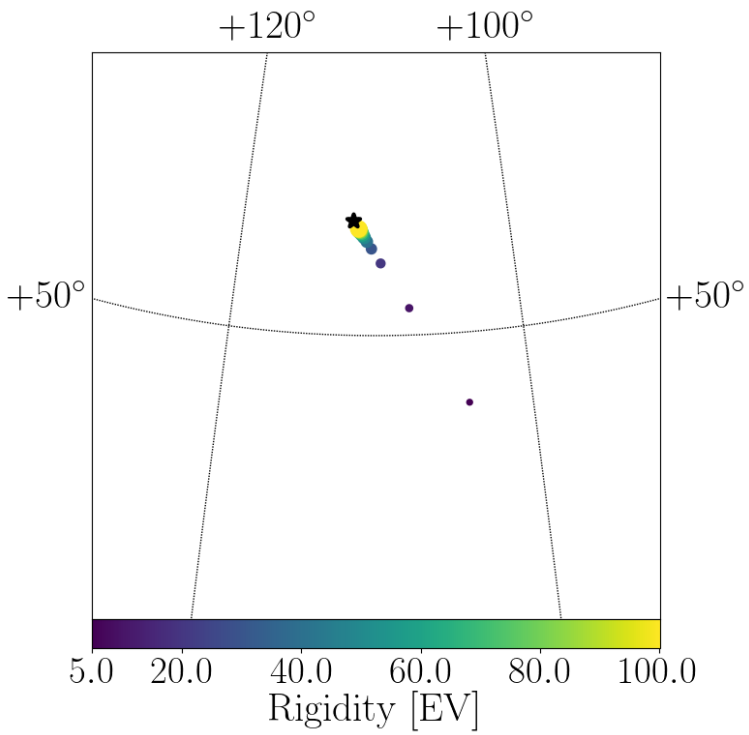


Figure 3.4.: Arrival direction of UHECRs of different rigidities from 5–100 EV propagated through a simplified JF12 GMF model without any random field components, indicated by the size and colour of the dots. The original source direction is indicated by the black star.

Although there are Galactic and extragalactic magnetic fields present in the Universe the latter is assumed to be subdominant to the Galactic one. Therefore, for simplicity, only the Galactic deflections are often considered, as we have also done for this work. To study the propagation and arrival directions of UHECRs, a GMF model is needed, see Section 3.3.2. However, since the exact GMF is not known, varying the different parameters of the JF12 model allows us to study characteristic deflections of UHECRs.
While the magnetic field structure of other galaxies can be measured from many different angles the measurement of our own is aggravated from the inside. However, some possible measurements were made that set some constraints on certain features that can be compared to observations of other similar spiral galaxies. The most known components of the GMF are the large-scale regular fields and the small-scale random fields, reviewed in [61, 62]. By measuring **Farraday Rotation Measures (RMs)** and **polarized synchrotron emissions** constraints on the GMF could be derived.

3.3.1. Magnetic field constraints

When polarized radiation travels true a magnetic field, it experiences a rotation of the polarisation plane by an angle

$$\theta = \lambda^2 \mathrm{RM}, \tag{3.10}$$

depending on its wavelength λ and the rotation measure (RM). The RMs can be derived as

$$\mathrm{RM} = 0.81 \int_0^L \left(\frac{n_\mathrm{e}(l)}{\mathrm{cm}^{-3}}\right)\left(\frac{B_\parallel(l)}{\mu\mathrm{G}}\right)\left(\frac{\mathrm{d}l}{\mathrm{pc}}\right) \tag{3.11}$$

with an electron density of the magnetized plasma n_e, magnetic field strength B and path element dl along the line of sight [63]. However, the electron density is not directly measurable and therefore the RMs are calculated by comparing radiation from the same source with different wavelengths λ, so that

$$RM = \frac{\theta_{1,2}}{\lambda_1^2 - \lambda_2^2}. \tag{3.12}$$

By deriving the RMs, constraints on the magnetic field can be obtained. In a survey from 2009 this has been done for almost 40,000 extragalactic sources [64], finding several coherent large-scale structures. Generally, the Galactic centre displays a large magnitude of the RMs and lessens towards both the poles and away from the centre. Also, the, in other galaxies observed spiral arm structure, is supported by the alternating negative and positive RM values within the Galactic disk [62]. In addition to the RMs, measurements of polarized synchrotron radiation caused by electrons of relativistic energy scattering within the GMF have been used to further model the magnetic field structure. It gives information on the perpendicular GMF component with respect to the line of sight and has been extensively studied in [65]. Although more observations helped improve the understanding of the structure of the GMF, the RMs and the polarized synchrotron radiation were the most influential for the improvement of the latest JF12 [3] magnetic field model used in this work.

3.3.2. Galactic Magnetic Field Models

Despite great efforts in modelling the GMF and measuring its parameters and components, many aspects of it still stay eluded. The most known ones on the other hand are the large-scale regular fields and the small-scale random fields, reviewed in [61, 62]. Several modulations have been created over the years, yet many differ quite severely in their structure. In Figure 3.5 a comparison of the deflection of a 60 EeV proton of three different GMF models is shown including the JF12 (left) [3], the 2010 model from Sun and Reich (SR10) [66] (middle) and the 2011 Pshirkov, Tinyakov and Kronberg (PTK11)[67] model (right). It can be seen that the SR10 model only considers a very low level of angular deflection of the arrival direction while the PT11 model assumes it to be several times higher in the Galactic plane. The JF12 on the other hand has a far more complex structure and of all three the largest degrees of deflection with the maximum around the Galactic

centre.

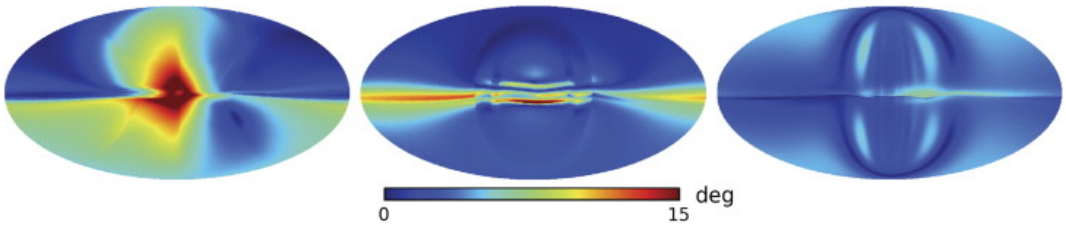


Figure 3.5.: Comparison of the magnitude of deflection of a 60 EeV proton for three different modulations of the GMF. The previously already mentioned JF12 is shown on the left, the SR10 in the middle and the PTK11 on the right. Taken from reference [68].

3.3.3. Jansson-Farrar (JF12) Model

The Jansson-Farrar JF12 model [3] is the so far most complete analytic description of the global magnetic field [69] fitted to polarized synchrotron emission taken by the WMAP7 [70] and over 40,000 RMs [64, 71]. It contains **three** components: **the coherent large scale field** which includes the halo, the disk, and the out-of-plane components, **the random turbulent field**, defined by its rms field strength, and the **striated field** accounting for the stretching or compressing of the coherent and/or fully random field [68].

Coherent Field

The coherent field has a logarithmic-spiral geometry with a ring of 3–5 kpc in its centre as shown in the left panel of Figure 3.6. The arms are individually fitted to the RM data [64].

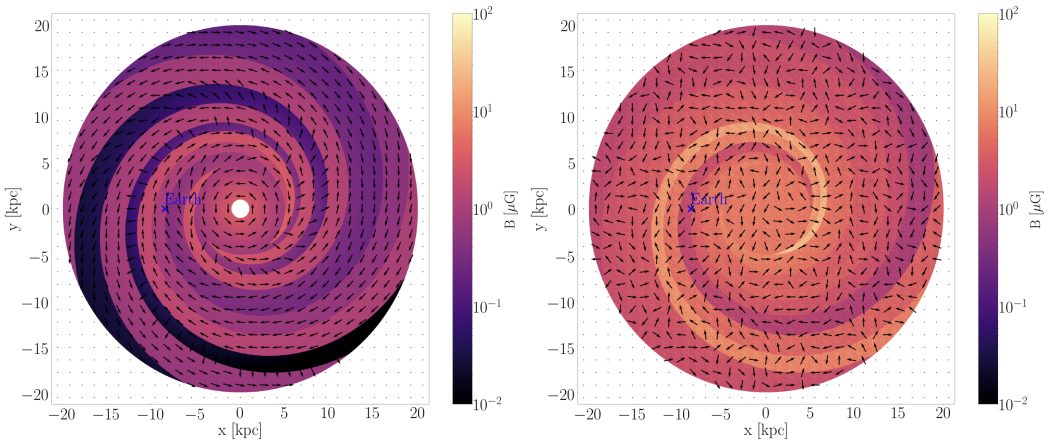


Figure 3.6.: Top-down onto the plane of the JF12 magnetic field model. The coherent field is shown left and the random turbulent is on the right. The magnetic field strength is illustrated by the colour scale while the arrows indicate the magnetic field directions. The Earth's position is shown as a blue cross at a distance of 8.5 kpcs from the Galactic centre. Taken from reference [72].

Additionally, an out-off-plane halo component often referred to as "X-field" was added, which was motivated by the observed X-shaped field structures in radio measurements of different edge-on galaxies [73].

Random turbulent field

Similar to the coherent field the fully random field also has spiral arms but with independently fitted rms strengths. The disk component has been superpositioned with a smooth halo component without any azimuthal structure, illustrated in the right panel in Figure 3.6. The coherence length Lc of this field can be written as

$$L_c = \frac{1}{5} L_{\text{max}} \frac{1 - (L_{\text{min}}/L_{\text{max}})^{5/3}}{1 - (L_{\text{min}}/L_{\text{max}})^{2/3}} \tag{3.13}$$

with L_{min} and L_{max} corresponding to the minimum and maximum variation in the turbulent magnetic field [74]. As the characteristic properties of the turbulent magnetic field are governed by many uncertainties we varied L_{min} and thereby the coherence length L_c to create different variations of the JF12 model.

Striated field

The striated field allows a random field that averages to zero on large scales to locally align in some preferred direction. This would arise if a fully random field got stretched or compressed and this has been the initial intention of including it. Also, a supernova would be able to locally compress or stretch a coherent field causing a striated field to arise.

Although the JF12 model showed a great improvement from previous modulations it still not describes the true structure of the GMF and has faults within it. For example in [75] it was indicated that the synchrotron emission map used for JF12 might be off by a factor of $\approx 1/3$. Therefore, as the exact structure is still unknown, this has to be considered in the data creation for the machine learning analysis in Chapter 5.
Additionally, to the variations of L_{min} and therefore the coherence length, we also change the random seed of both the turbulent and striated field component to impede the neural network to learn just a specific realisation of the GMF model. Similar variations of this kind were performed in [76].

CHAPTER 4

METHODS

For many unsolved questions regarding the nature and propagation of UHECRs, simulations enable a way of testing theories, running different "what if?" scenarios and visualizing them when real-life measurements are not possible, too expensive or not scalable. Yet caution has to be taken when interpreting the results as these simulations are limited by the knowledge they were created with, so their result could be unrepresentative of the real underlining physics. Nevertheless, simulations have shown great success in many fields of physics and science in general, e.g. the Pythia [77] simulation program for high-energy collisions at accelerators like the LHC. Here, simulations have had not only a large influence on how the detectors should be built but also in the testing of theories of physics beyond the Standard Model (BSM) and the finding of the Higgs boson in 2012 [78]. For this work, we create simulations of UHECRs as could be observed at Auger to train a neural network (NN) to detect a fraction of particles originating from a few sources embedded in a larger isotropic background distribution. We consider several different variations of the JF12 GMF model and mass compositions of the UHECRs before we finally apply the tuned model to data taken by Auger.

This chapter begins with an introduction of the simulation program and the modules (Section 4.1) that are used to create the simulated data for this analysis, followed by the creation of the different variations of the JF12 GMF model Section 4.2. Next, we show the necessary steps to create a UHECR flux as could be observed at Auger in Section 4.3. Finally, we show how the different mass compositions can be calculated as a function of the power-law index in Section 4.4.

4.1. CRPropa 3

CRPropa 3 is the third generation of the open-source simulation framework created to study the propagation of UHECRs through (extra)Galactic space [2]. For this work, the fully implemented Python steering of the program was used. For detailed information on the installation and all its functionalities see the documentation page [79].

Since CRPropa 3.0 it is possible to attain the entire information of each CR at any propagation step. This information is stored in the so-called Candidate class on which each module acts that are within this class. An overview of the modular structure of CRPropa 3 is shown in Figure 4.1.

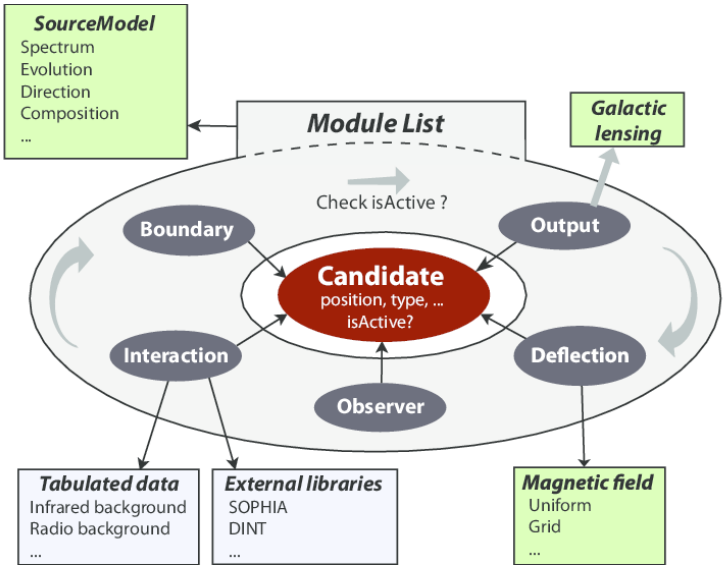


Figure 4.1.: Modular structure of CRPropa 3. The Candidate class is acted on by each of the modules within the class. After each propagation step the Candidate is checked if it is still active.

4.1.1. Propagation modules

The propagation through Galactic and extragalactic space is simulated by steps, where for each step the new location, energy and composition of each propagating particle is calculated based on the stochastic Monte Carlo approach. For each step, a CR thereby

has a certain probability to interact in one of the ways introduced in Chapter 3. The chosen step size differs for different propagation modules. For this work, the "Simple Propagation" was used for the 1D extragalactic (see **??**) and "Propagation CK" for the Galactic simulations through the GMF. The latter is based on the Cash–Karp (CK) algorithm [80] and handles the deflection of CRs by the GMF with a dynamic step size control. The "Simple Propagation" on the other hand uses a simple rectilinear propagation with equally binned step sizes.

4.1.2. Magnetic lenses

Computationally it is far more efficient to inverse the propagation of the CRs through the Galaxy as the majority of particles would miss Earth and thereby their detection, resulting in unnecessary computations. Also, the actual amount of detected CRs on Earth would statistically vary with each simulation. Therefore, particles are usually "backtracked" from Earth to the source by just reversing their charge, changing a proton for example to an anti-proton and switching the Galactic border with the observer. Although this method is computationally a lot more efficient, when the particle number is increased to a realistic value, this process still takes a significant time, dependent on the chosen propagation step size. As the simulation of the Galactic propagation in this method is basically just a one-to-one projection from the Galactic border to the observer Earth, this process can be replaced by large matrix multiplication [81]. By binning each point in the sky and multiplying it with a matrix that maps each point at the Galactic border on a point of the observer's sky as a function of rigidity, the use of computational resources can be decreased by a large factor. This matrix, which represents the mapping from the Galactic border to the observer, is referred to as a "magnetic lens" and is used throughout this thesis. To create a lens of a resolution similar to the angular resolution of Pierre Auger, 256 particles are backtracked for each rigidity bin and each pixel. The entire sphere corresponds to 49152 pixels, which is close to the resolution seen by Auger. For a chosen rigidity binning of $\Delta E = 10^{0.02}$ the computation is very expensive, adding up to $n \times 256 \times 49152$ simulated particles for m propagation steps. As this work only looks at energies from 32–200 EeV, the simulation can be narrowed down to the corresponding rigidity range. To cover the deflection up to iron for these energies the lenses are created ranging from 1–200 EV, resulting in 117 rigidity bins.

When using a magnetic lens as a propagation module the setup of the simulation can be summarized as follows

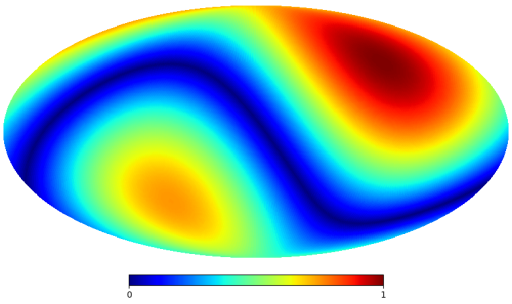
1. An empty so-called "ParticleMapContainer" is created that can store the entire particle information
2. Particles are added to the map, following the desired distribution, direction and composition at the Galactic border.
3. The energy spectrum can be input directly as an injected power law, with a certain energy range and power law index γ.
4. The lens is applied to this "map" at the Galactic border
5. Random particles are extracted from the "lensed" map following the input given to it. The amount of extracted particles from the "map" should not exceed the amount of extracted ones.

The arrival distribution of the extracted particles will then follow one deflected by the GMF model.

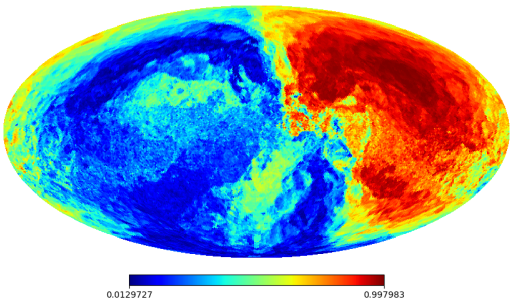
All maps are created with UHECRs ranging from 32–200 EeV following a power law of $\gamma = -4.2$, which is motivated by the observed diffuse spectrum above the spectral break at roughly 43 EeV from Pierre Auger.

4.2. Variations of the Galactic magnetic field model

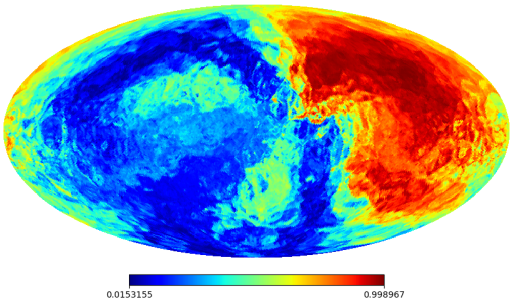
To ensure that the neural network does not learn just the specific GMF, we create several variations of the JF12 model. At first, we change the random seed of the random turbulent and striated magnetic field component. Each of the seed variations is used to create a different data partition for the neural network, explained in Section 5.2. Also, we create further variations by stepwise increasing the coherence length L_c of the random turbulent magnetic field from a default value of $Lc \approx 5$ to $12.5, 25, 37.5$ and 50 parsec. To visualize these variations, the magnitude of deflection of a particle of rigidity 5 EeV is illustrated in Figure 4.2 by applying the different GMF lense to a sample healpy [82, 83] heatmap. Shown are the effects on the heatmap for the default JF12 magnetic field lens, a seed variation and that of an increased coherence length of $L_c = 50$ pc.



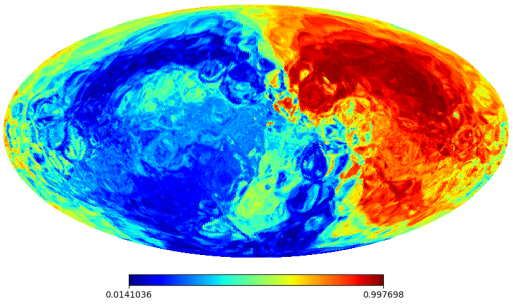
(a) Sample heatmap



(b) Default JF12 magnetic field lens



(c) Seed variation of the random turbulent and striated magnetic field components of the JF12 model



(d) Variation of the coherence length $L_c = 50$ of random turbulent magnetic field component of the JF12 magnetic field model.

Figure 4.2.: Comparison of three different variations of the JF12 magnetic field model lenses applied to a sample heatmap (a). The magnitude of deflection corresponds to a particle rigidity of 5 EeV.

Although the overall structure is similar, several significant small-scale changes can be observed for both shown variations. These differences present a crucial aspect to ensure and validate that the NN does not just learn how to reverse the Galactic propagation.

4.3. Geometrical exposure of the Pierre Auger Observatory

As Auger is a ground-based observatory, the detection of UHECRs is bound by the Earth's rotation. Therefore, an exposure function is built into the simulation in the form of an acceptance or rejection condition. Auger is located in Argentina at a declination of –35° so that there is a fraction of the Northern sky it can not observe at any given time. The exposure of an observatory can be written as a function of the declination only measured in $\mathrm{km}^2 \times$ years, resulting in an effective time-integrated area of the CR flux. The absolute number of CRs that are measured from each point of the sky is not only dependent on the actual celestial anisotropies but also depends on the observatory's relative exposure. Also, the detectors at Auger are assumed to be first fully efficient at zenith angles below $\theta_m = 80°$, resulting in a gradually increasing exposure as seen in equatorial coordinates

in Figure 4.3a and transformed into Galactic coordinates in Figure 4.3b. Here, the relative exposure as stated in reference [84] is used, which is a function of declination δ and detector efficiency θ_m only:

$$\omega \propto \cos(a_o)\cos(\delta)\sin(a_m) + a_m \sin(a_0)\sin(\delta), \tag{4.1}$$

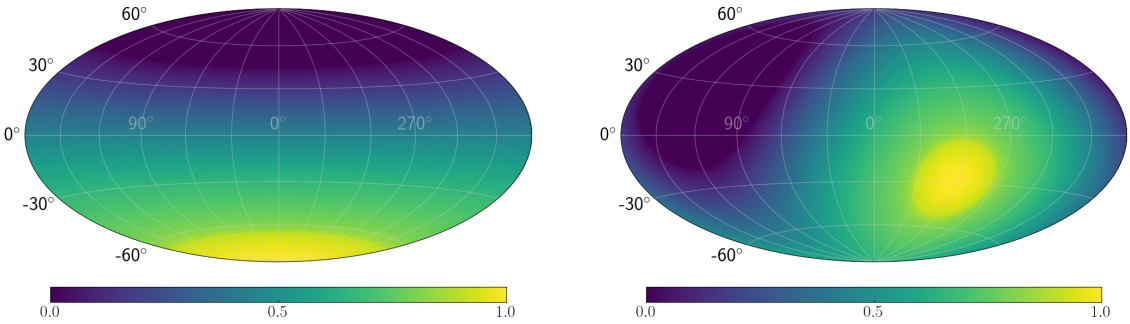
where

$$a_m = \begin{cases} 0 & \text{if} \quad \xi > 1 \\ \pi & \text{if} \quad \xi < -1 \\ \cos^{-1}(\xi) & \text{otherwise} \end{cases} \tag{4.2}$$

and

$$\xi \equiv \frac{\cos(\theta_m) - \sin(a_0)\sin(\delta)}{\cos(a_0)\cos(\delta)}. \tag{4.3}$$

The Pierre Auger observatory is located at a latitude of $a_0 = -35°$ and we use a detector efficiency $\theta_m = 80°$ throughout all our simulations.

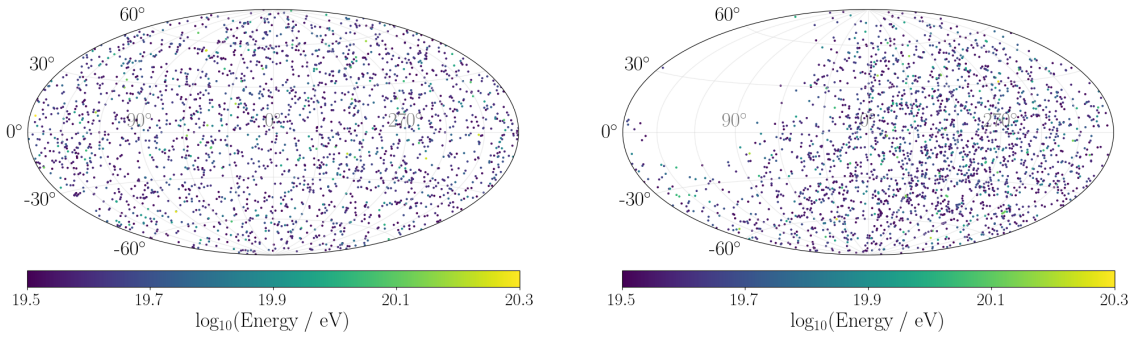


(a) Equatorial coordinates

(b) Galactic coordinates

Figure 4.3.: Exposure as a function on the celestial sphere (a) in equatorial coordinates (b) in Galactic coordinates.

The arrival direction of 32-200 EeV isotropically distributed protons at the Galactic border is shown in Figure 4.4a. The same distribution after the JF12 magnetic field lens and the Auger exposure is applied is in Figure 4.4b. After the propagation, the arrival-direction distribution is visually still highly isotropic and the particle density can be seen to increase towards the "hot spot" shown in Figure 4.3b. However, some GMF-induced small-scale clustering can be observed. The task of the neural network will be to distinguish this kind of clustering from one that arises from several particles arriving from the same source.



(a) Isotropic distribution

(b) Applied lens and exposure

Figure 4.4.: Simulation of an isotropic CR proton distribution (a) at the Galactic border and (b) propagated via the JF12 magnetic field model with the applied Auger exposure.

4.4. Composition calculation

For simplicity, we choose as starting point to simulate a proton-only composition, however, this is not representative of the actual observations made by Auger as stated in Section 2.2. The actual composition is not yet totally resolved but tends to become heavier towards higher energies. In the second run of simulations, we create a pure nitrogen configuration to view the performance of the NN for still just a single element, but with a seventh of the rigidity. Finally, we create a mixed composition of several elements motivated by the actual observations, however, instead of taking all the occurring elements, a selection of representatives is commonly chosen. Often these elements are P, He, N, Si, and Fe, as we also do in this work. Their abundance is created based on the elemental injection fractions f_A^R as stated in table 1.4 of [5]. These fractions are given as a relative ratio at the same rigidity and can be transformed to the ratio at the same energy for a spectral index γ via

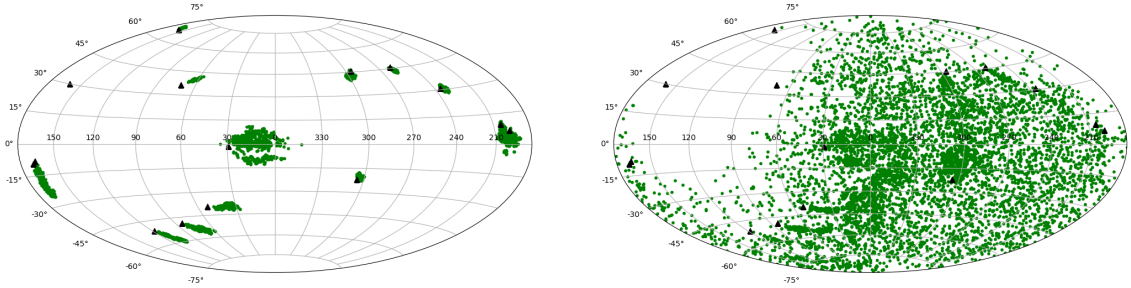
$$f_R R^{-\gamma} = \Phi(R(E)) = \frac{dN}{dR}\left|\frac{dE}{dR}\right| = \Phi(E)|Z| = f_E E^{-\gamma} Z. \tag{4.4}$$

Multiplying both sides with R^γ and keeping in mind that $E = RZ$ an expression can be obtained for the conversion of the composition fraction at the same energy f_E and same rigidity [1]

$$f_R = f_E Z^{-\gamma+1}. \tag{4.5}$$

For an index of $\gamma = -4.2$ we acquire an iron-dominated composition with roughly P $\approx$ 2.6%, He $\approx$ 3.7%, N $\approx$ 17.5%, Si $\approx$ 13.3% and Fe $\approx$ 62.9%. While the deflection of protons of energy above 32 EeV is of an order of a few degrees as can be seen in Figure 6.14a, the deflection by the GMF of heavier nuclei becomes so large that the particles can not be visually assigned to their sources anymore as shown in Figure 6.14b.

[1] This method was motivated in private communication with D.Ehlert [85].



(a) Proton-only composition with $\gamma = -4.2$ **(b)** Mixed composition with $\gamma = -4.2$

Figure 4.5.: Arrival directions of 7500 UHECRs (green dots) propagated through the default JF12 magnetic field model originating from 15 sources indicated as black triangles. Shown are (a) a proton-only composition and (b) a mixed iron-dominated composition calculated via Equation (4.4).

For the mixed composition, some clustering can still be seen which corresponds to the fraction of lighter nuclei while for the heavier nuclei, the deflections become so much larger that they too arrive from almost all directions.

CHAPTER 5

MACHINE LEARNING SETUP

Mankind lives in a data-driven world, producing 97 zettabytes of data in 2022 [86]. In physics alone, experiments like e.g. the LHC, over 380 petabytes of data were saved from experiments since its first run [87]. To handle this vast amount of data and shine a light on all the information hidden within it, large computational resources and efficient algorithms are needed. Here machine learning and deep learning in particular have shown omens success in finding the underlining truths and patterns in many fields of physics, which would be incomprehensible to scientists without. For example, the usage of cutting-edge machine learning applications at the LHC was a major contributor to finding the Higgs Boson in 2012 [78]. Although the actual data is often limited, simulations can be used to upscale the data size to possibly find physical patterns using deep neural networks.
At first, an introduction to the methodology of deep learning is given Section 5.1 followed by a detailed walkthrough of the data pre-processing and features extensions in Section 5.2. The most important hyperparameters are introduced in Section 5.3 and the different performance measures used for this work in Section 5.4. The final neural network structure is given in Section 5.5.

5.1. Structure of deep neural networks

A neural network (NN), takes a set of input variables or so-called features x_1, x_2, x_3, x_4..., and outputs an array of predictions, called the output layer. All layers in between are named hidden layers as their outputs are not directly observed. A NN which consists of at least two or more hidden layers is referred to as a deep neural network. A layer consists of several nodes that are each connected to all other nodes in the previous and next layer via so-called weights. An illustration of the structure of a simple deep neural network is shown in Figure 5.1.

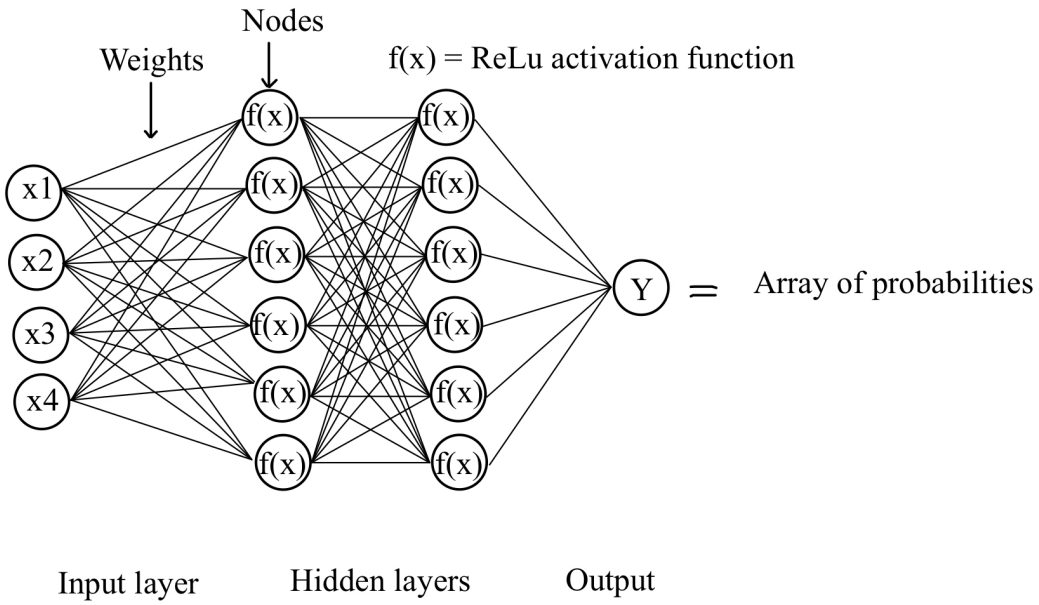


Figure 5.1.: Illustration of the structure of a deep neural network with 2 hidden layers.

The value of a certain node, also called feature or neuron (as in the brain), is calculated via the weighted sum of all m incoming connected inputs

$$Feature = \sum_{i}^{m}(w_i \times x_i) + b. \tag{5.1}$$

Here, w_i is the weight of the corresponding input x_i and b is a constant to shift the activation threshold of that node. If a node is activated or not, and if it is, how much influence it has on the next layer, is decided by an activation function. The most commonly used activation function is the ReLU function [88] which is zero for all negative values and x for positive ones. If thereby, a node has a weighted sum of below zero the node is deactivated and does not influence the next layer. This procedure is similar to the brain and is also the reason it is called a neuron, which can or can not be switched on. The power of the neural network lies now in updating the weights in such a way that the final error on the output is minimized. The error is calculated via the sum over a cost function measuring the difference between all predictions and true values

$$J(\omega) = \frac{1}{m}\sum_{i}^{m}\text{Cost}(h_\omega(x_i), y_i). \tag{5.2}$$

For this work, we create a binary classifier and thus we use the **Binary Cross-Entropy** [89] cost function which is given as

$$\text{Cost}(h_\omega(x), y) = \begin{cases} -\log(h_\omega(x)) & \text{if} \quad y = 1 \\ -\log(1 - h_\omega(x)) & \text{if} \quad y = 0. \end{cases} \tag{5.3}$$

This can be rewritten in one line as

$$J(\omega) = -\frac{1}{m}\sum_{i}^{m}[y_i \log(h_\omega(x_i)) + (1-y_i)(\log(h_\omega(x_i))]. \tag{5.4}$$

Hereby, the condition

$$0 \leq h_\omega \leq 1 \tag{5.5}$$

has to be satisfied via a function A

$$h_\omega = A(\omega^T x). \tag{5.6}$$

A function that can acquire this is the Sigmoid activation function

$$A(x) = \frac{1}{1+e^{-x}}, \tag{5.7}$$

used for binary classifiers, as it is also done in this work.

By using a process called **gradient decent** the error can be backpropagated through the network by calculating the partial derivative of the cost function with respect to each weight and updating the weights repeatedly as follows

$$\omega_j := \omega_j - \eta \frac{\partial}{\partial \omega_j} J(\omega). \tag{5.8}$$

Here η is the step size of each weight update and is referred to as the learning rate, one of the most important tunable hyperparameters of the NN.

The entire training process can be summarized as follows

0. The very first set of weights is most commonly randomly initialized

1. The values of each node of each layer are calculated by applying an activation function to the weighted sum (Equation (5.1)) of all incoming connections.

2. For the output layer, the Sigmoid activation function is used to obtain a prediction value between zero and one.

3. The difference between the predictions and the true labels is calculated via the cost function Equation (5.2). The goal of the network is to minimize this error.

4. Via gradient descent, the weights are updated by calculating the partial derivative of the cost function with respect to each weight (Equation (5.8)).

5. Steps 1-4 are repeated for each new data point until the error stops decreasing/ the prediction ability of the NN stops improving. Hereby, an iteration over the entire data is referred to as one epoch.

Finally, a certain threshold on the output node is chosen to distinguish one class from another. By default, $y \leq 0.5$ would correspond in this application for a predicted source event, while $y < 0.5$ to an isotropic background event. This threshold can be adjusted though, if it is advantageous to the general performance of the NN. To speed up the training process and smoothen the learning curve the weights are often not updated for each single data example but instead for a so-called "batch" of the data consisting of a certain amount of samples. This too is a tuneable hyperparameter that has to be adjusted in order to get the best performance. The most important hyperparameters, and how they can be changed, are shortly explained in Section 5.3.

5.2. Data pre-processing

Before any training is applied the data set is split into three fractions: the training, the validation and the test data set, see Figure 5.2. The training data set gets the majority of the examples (often 70% or more) as it is the data used to update the weights and actually trains the model. The training set can contain certain so-called "artefacts" which are specific features that only apply to this fraction of the data. Such learned artefacts are not desired as the NN should be robust and perform well on any data created by the same physical laws e.g. in this work the NN should predict if a UHECR originated from a source or the isotropic background. It is not desired to just learn the source positions of the training set but to learn the magnetically-induced patterns independent of their exact arrival direction. Therefore the validation set is used to compare the accuracy and loss of the new weights on a not trained fraction of the data set. If the loss on the training set keeps decreasing but not for the validation set, it is a sign that the model is overfitting and only learning something specific for the training set. The accuracy and loss of a clearly overfitting NN trained on a pure proton data set are shown in Figure 5.3a. It can be seen that the loss continues to decrease for the training data set but not for the validation data set. The NN is learning properties that only apply to the specific training fraction of the data. Although the validation set is not used to update the weights, it is used to evaluate the model after each epoch. Different implementations to the network structure and changes to the hyperparameters as oriented to the learning curve of this data fraction. It thereby is in a way involved in the training process of the NN and thus a third fraction, the test data set, is needed. The test data set is used to check the final performance of the NN. As it has not been seen at any point during the training it can give an unbiased evaluation of the model. The accuracy on this fraction gives the actual performance of the NN.

We state here again that to generalize the training of the NN as well as possible, we use a different random set of source positions and variations of the GMF model for each fraction. We can thereby evaluate the accuracy and loss during the training process with the validation set and use it to implement an "early stopping" condition explained in Sec-

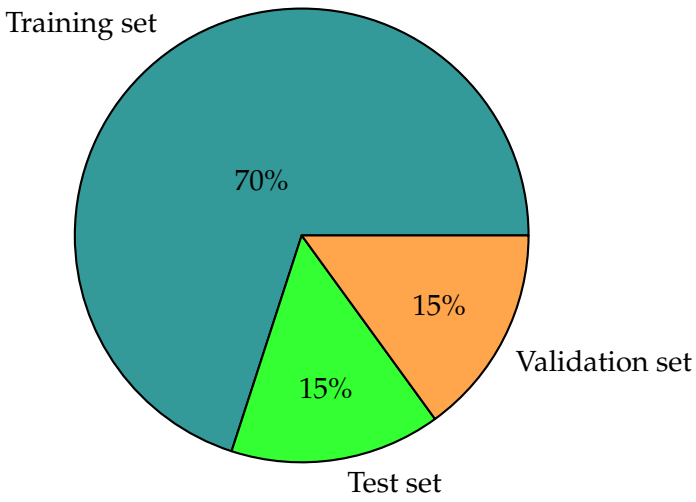


Figure 5.2.: Used split of input data into training, test and validation set.

tion 5.3. Also while at first, we only use one seed for each data fraction, we from hereon out use three different seeds for the training data sets to directly include the uncertainty of the random components in the training process.

Data normalization

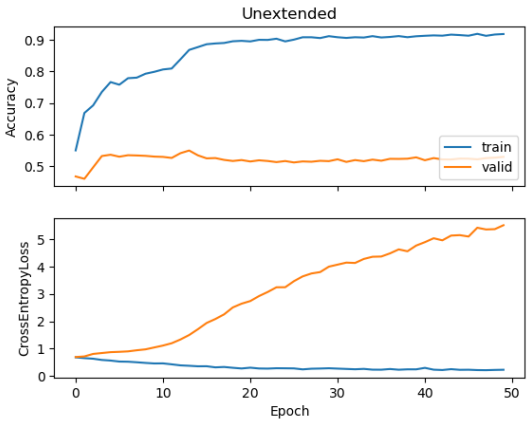
When training a NN the normalization of the data set with respect to each input feature is very important as a feature that contains significantly higher values than others can become dominant in the training process [90]. Such a feature will acquire a much higher loss by default without being of actual higher importance to the model's performance. As the data set contains on the one hand values up to 2×10^{20} for the energy and on the other hand coordinates that only reach a value of 2π for the longitude, the training would be extremely biased. Therefore each input feature has to be normalized to a range $[0, 1]$ for unbiased training to be acquired.

Feature selection

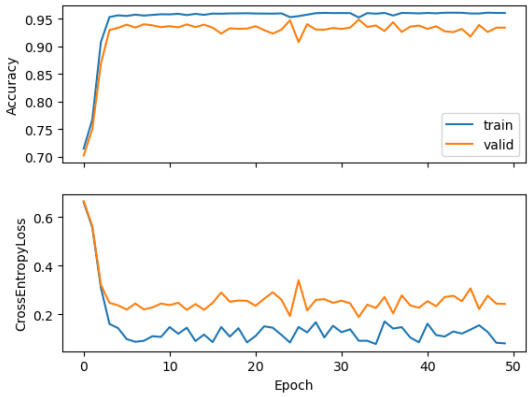
When training NNs the results can only be as good as the data trained on, therefore a careful data selection is key. At first, we replace the energy information with its logarithm, which gives us improved results. As the energy follows a power-law distribution the majority will have relatively small values and a few relatively big ones. Taking the logarithms naturally reduces the skewness of the distribution while a normalization to values between zero and one would result in the majority of values being very close to zero. Training on the logarithm stresses the relative multiplicity in the data by linearizing the relationship of the power-law distribution and therefore shows to be beneficial towards the NN's performance.

The UHECR data is limited to their energy and arrival direction in Galactic coordinates only, as the charge is considered an unknown quantity. This is a very small range of usable data to train on. However, more information can be highlighted for the NN to gain better performance by combining already existing features in order to create new ones that represent known relations or correlations of the data set. This is a commonly used method in machine learning called "feature crossing". At first, we expanded the data set by adding the distance of each event to its N nearest neighbours to stress the clustering in the arrival directions of the UHECRs. Also by adding relative spatial information, we are able to teach the NN to some extent spherical data. Without this information, the NN has no way to learn that an event just below a longitude of 2π (the maximum value of this feature) may lie directly next to one at longitude zero (the minimum value of this feature).

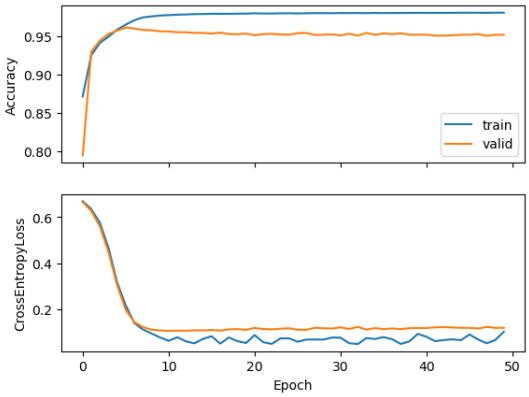
We calculate the angular distance to all other particles in each of the three data set fractions together with the corresponding difference in energy to these particles. At first, we only consider the closest 5 neighbours and add them to the existing three features, resulting in a total of 13 features. The influence of this data extension is shown in Figure 5.3 by monitoring the training and validation loss and accuracy of each epoch. It can be observed that the NN is not able to learn the clustering of the UHECR arrival distribution just from the energy and arrival coordinates alone. While the loss on the training data is decreasing, the validation loss is increasing. The model is just learning the specific source positions of the training set and thereby shows almost no prediction ability for the validation data with different source positions and a different magnetic field model. After the feature extension of the 5 nearest neighbours, not only does the performance on the training set increase but also the validation performance increases to almost the level of the training one, see Figure 5.3b. To see if the performance can be further improved for both fractions, we further extend the data to the 10 nearest neighbours, adding 20 new features instead of 10, Figure 5.3c. It can be seen that the performance increases even more, however, a slight overfit is still observed. A further feature extension to more nearest neighbours did not show any improvement. We want to note here that the hyperparameters of the network have already been adjusted accordingly to some extent. A detailed introduction to how this is done and the final best performance is achieved is shown in the next section.



(a) Unextended data set consisting of the energy and Galactic arrival coordinates. The NN is learning the source positions of the training data and therefore performs poorly on the validation set that contains different source positions.

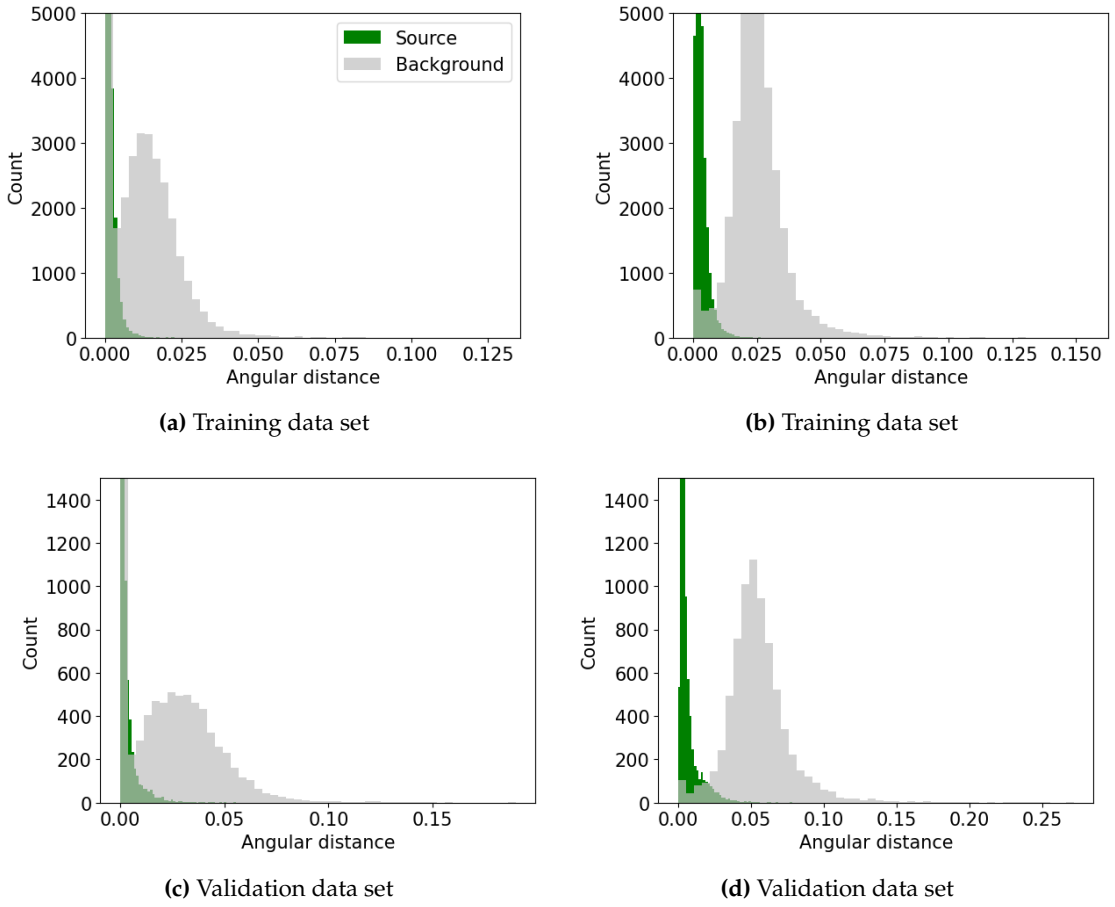


(b) The data was extended with the angular distance and energy difference to the 5 nearest neighbours of each UHECR. The network can learn the clustering and thereby performs similarly well on the validation set.



(c) The data was extended with the angular distance and energy difference to the 10 nearest neighbours of each UHECR.

Figure 5.3.: Loss and accuracy for each epoch with a threshold at 0.5 on the output node. The NN is trained on 70000 proton UHECRs propagated through the default JF12 magnetic field model. The blue and orange lines correspond to the training and validation data sets respectively.



(a) Training data set

(b) Training data set

(c) Validation data set

(d) Validation data set

Figure 5.4.: Comparison between the distribution of source and background UHECRs of their angular distance of the nearest (left) and 10th nearest neighbour (right) before normalization.

Furthermore, we try a feature extension motivated by the analytical multiplet search stated in [91]. Here the so-called multiplets for selected source candidates were analysed by deriving the correlation coefficient and thrust ratio for certain preselected regions of interest (ROIs), however, they only considered protons and Helium deflections. Also, they focused on a "targeted search" meaning they considered each source candidate separately. We on the other hand are applying an all-sky search for unknown source positions and aim towards a heavier nuclei composition as it is suggested by the observations from Auger. Although they did also apply an all-sky search on Auger data, they had to create a grid of different ROIs from which they selected those with the highest p-value significance. For this work, such a division into different ROIs would have to be done for each data set fraction individually whereas the optimal size of the ROIs would have to be also calculated for the different compositions analysed. As this kind of data pre-processing would take considerable time and effort, with no guarantee of success for heavier nuclei, it is not included in the scope of this thesis. However, we take their used method as motivation for a feature extension and derive for a subsample N of UHECRs the correlation

coefficient

$$C(u, 1/E) = \frac{\mathrm{Cov}(\beta, 1/E)}{\sqrt{\mathrm{Var}(u)\mathrm{Var}(1/E)}} \tag{5.9}$$

with variance $\mathrm{Var(x)} = \langle (x - \langle x \rangle)^2 \rangle$ and angular distance β. We try out several different variations of subsamples N, all based on nearest neighbour calculations for each UHECR. For example, we select the $10^{th} - 20^{th}$, then the $20^{th} - 30^{th}$ and calculate the correlation coefficient as stated in Equation (5.9) between the angular distance and the inverse energy in the hope to capture a correlation with source particles with a higher degree of deflection. Although some of these features capture local correlations well, it does not improve the neural network's performance for higher deflections, due to the large amount of background particles being included. We note though, that training on this extension alone does show a decent training ability but does not surpass or improve as additional features the overall performance of our NN and we thereby chose to discard it. Motivated by the calculation of the thrust ratio, we added the corresponding exposure weight of each UHECR as a further feature, visualized in Figure 5.5, yet we observed no change in performance by this addition either. We did not derive a closely related thrust ratio feature, as ROIs would be needed for the calculation. Nonetheless, we do not rule out that a closer combination of our analysis and the thrust ratio method used in [91] could show to be beneficial in detecting some larger deflected source events.

To study the correlation between the new features, the correlation matrix of the extended data set is shown in Figure 5.6. It can be seen, that the energy difference to the nearest neighbours shows a decreasing correlation with the corresponding angular distance towards particles further apart. The inverse energy difference is also tested as a feature, however, neither shows any improvement of the network when compared to a model trained without. Therefore we dropped it which did increase the performance by a slight margin. This could be possibly explained by the model learning the energy-dependent clustering structure with the total energy of each UHECR together with the angular distance and thereby the added information would be redundant. This is supported by the large correlation with the energy feature itself. Redundant, highly correlated or useless features can be harmful towards the network's performance [92]. After dropping the energy difference the accuracy increased by 1% on both data set fractions. The angular distance on the other hand shows a negative correlation with the class and also quite some correlation with itself for obvious reasons. This feature showed the largest improvement for the network by giving the model the ability to capture the clustering structure of the NN. The apparent difference in the distribution between the source and the background UHECRs can be seen in Figure 5.4 where the nearest and the 10th nearest neighbour are shown for the training and validation data set of the pure proton composition.

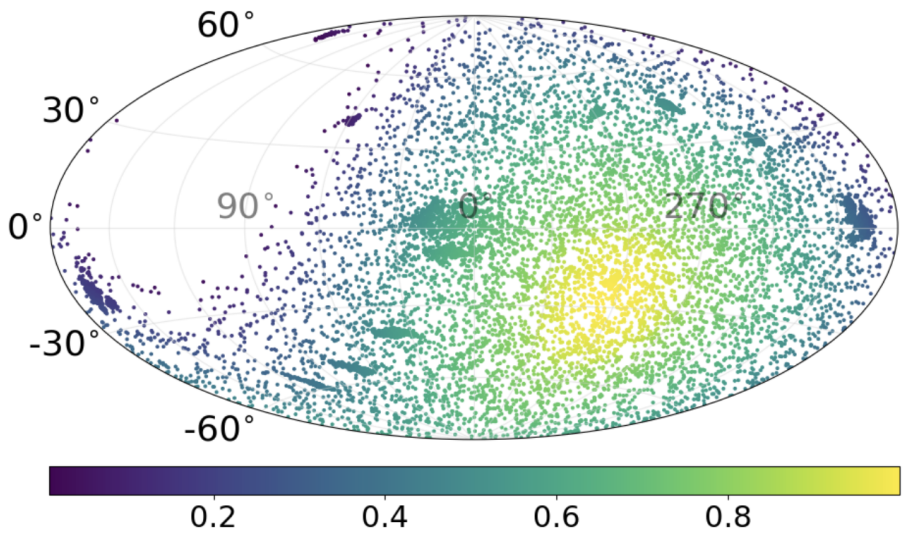


Figure 5.5.: Exposure weight for each UHECR's arrival direction of the validation data set.

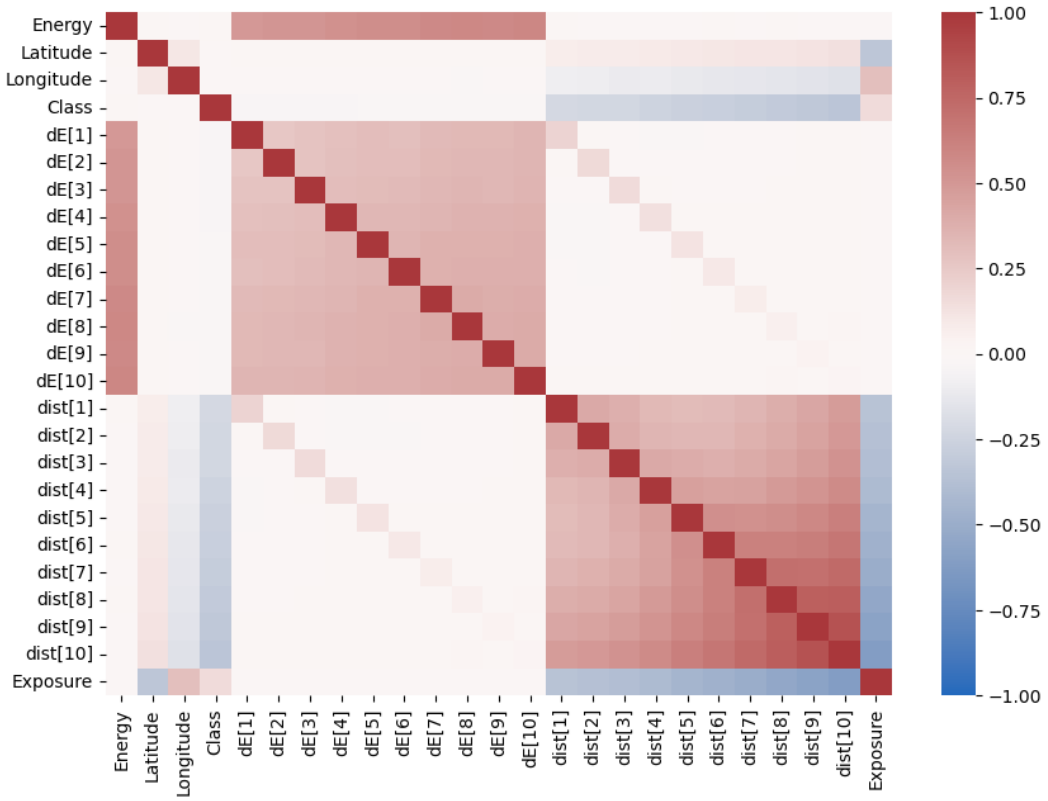


Figure 5.6.: Correlation matrix of the extended data set. The numbers in the brackets refer to the energy and distance to the N_{th} nearest particle respectively. Red corresponds to a positive and blue to a negative correlation respectively, while white indicates no correlation between the features.

5.3. Hyperparameters

Parameters of the NN that have to be set before the training process begins, are considered hyperparameters. It is one of the main challenges towards getting a good performance of the NN and they have to be often severely "tuned". An overview of the most important ones and their functionality are listed in Table 5.1.

Table 5.1.: Selection of some of the most important hyperparameters of an NN.

Hyperparameter	Functionality
Layers	Total amount of layers
Nodes	Number of nodes per layer
Learning rate	Step size of each weight update of the gradient decent process
Batch size	Number of training samples seen to update the weights
Activation function	Decides if/how much a decision impact the neuron has
Epoch	Training run over the entire data set
Weight decay	Decreases learning rate for higher iterations
Dropout	Probability of a node connection dropping out

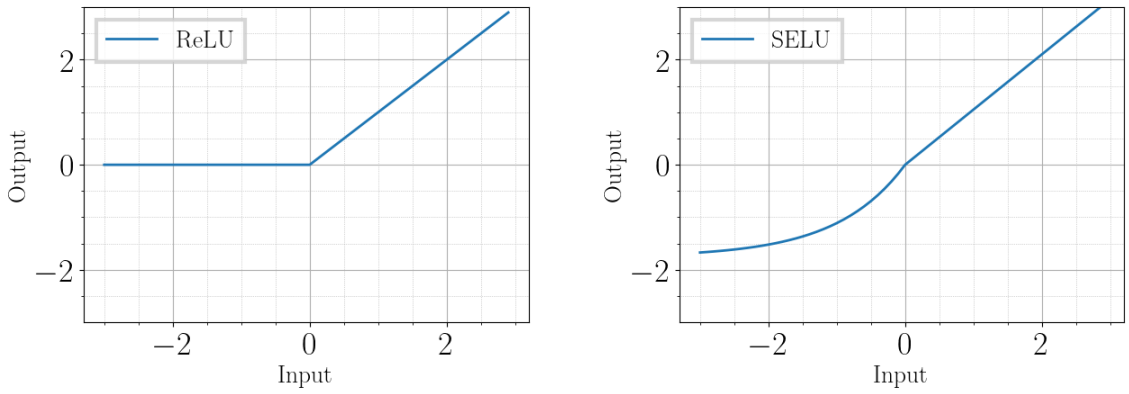
In order to tune these hyperparameters the loss and accuracy of the training and validation data set are monitored for each epoch as was already done for the feature selection. Of the hyperparameters shown in Table 5.1 the learning rate is arguably the most relevant as it dictates the step size of each weight update. If the learning rate is too large the neural network will continuously jump over the optimum and if chosen too small, it will take too long or will never reach the minimum loss value. On the other hand, decreasing the batch size leads to more weight updates within one epoch as an epoch is an iteration over the entire data set. This is just one example of how two hyperparameters are connected with each other and there is no common default setting. There are some loss functions that tend to work better with higher learning rates and smaller batch sizes and vice versa. The exact setting has to be found for each problem individually by comparing the validation and training loss for each epoch. Due to the counterplay of many hyperparameters, it is often better to only change one at a time, however, some only show an improvement in tandem and thereby finding an optimal setting is no easy task. The following hyperparameters are added to the network structure.

Batch normalization

Additional to the initial data set normalization, a so-called batch normalization should be implemented after each activation function. Hereby, during the training process, each batch is normalized to its zero mean and unit variance after each layer's activation function. It enables setting a higher learning rate, accelerating the training process and also serves as a regularization technique against overfitting [93]. However, for self-normalizing activation functions, this is not needed (see next section).

Activation functions

Over the years many different activation functions have been designed, each better suited for certain applications. The already previously introduced and most commonly used one is the ReLU activation function [88] shown in Figure 5.7a. Although it is often the default function of use, the ReLU activation function can get trapped in a dead state. This is the case when the change of the weights is so huge that the output of one layer becomes extremely low (negative). In this case, it can stay stuck in the negative regime for the next layers too. As a result, the affected cell can not contribute to the training of the network. If this occurs more frequently, the performance of the network will never reach its optimal capacity.



(a) Rectified linear activation function (ReLU)

(b) Scaled Exponential Linear Units function (SELU)

Figure 5.7.: Comparison of the ReLU and the in this work used SELU activation function.

The Scaled Exponential Linear Units (SELU) activation function on the other hand avoids this problem by enabling negative output values and thus can not deactivate any nodes. Simultaneously, just like the ReLU, it does not have the problem of vanishing gradients. For deep neural networks, the gradients can become so vanishingly small, that they are not able not to change the weights significantly enough. As both the SELU also has a derivative ≤ 1 for positive input values, this is not a problem. Furthermore,

the SELU activation function is self normalizing, as a result, batch normalization is not necessarily needed. The SELU is a relatively new activation function and therefore not fully explored yet in many scenarios, however, we experience consistently better results with it and therefore use it in our model structure. For more information on the SELU activation function, we refer to reference [94].

Dropout

Dropout is a regularization technique [95] which tackles the overfitting of a NN by adding a probability of forgetting random connections between two layers for one training iteration. The procedure is illustrated in Figure 5.8. It therefore can help to avoid getting stuck in a local minimum during training by preventing the model from just memorising certain insignificant random fluctuations. We added it between every layer, however, we note that the magnitude has to be adjusted for the different mass composition scenarios.

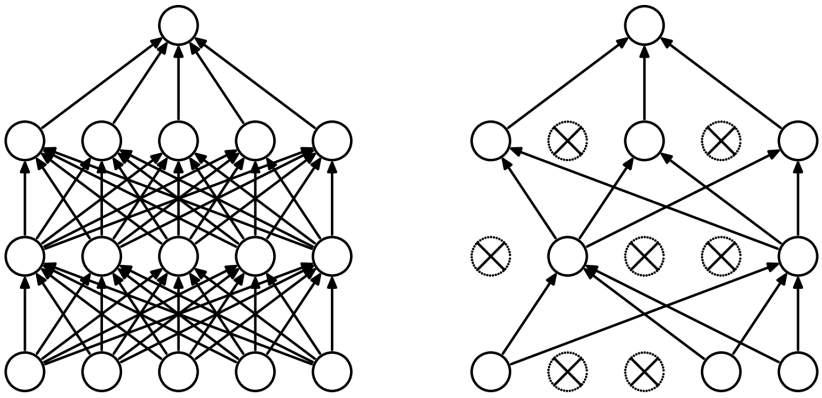


Figure 5.8.: Illustration of a dropout applied to a deep neural network structure. A probability is implemented of "forgetting" a connection between two nodes for one training iteration. Shown is a fully connected deep neural network (left) and the same network with an applied dropout probability after each layer (right). Taken from [95].

L2 regularization

L2 regularization [96] is another regularization technique. By adding a penalty term to the loss function that is proportional to the sum of the squared weights it reduces overfitting for higher epochs and is often referred to as weight decay as it continuously reduces the weight updates in the training process. The setting of it is highly related to the size of the learning rate and has to be tuned accordingly.

Early stopping

To secure the best set of weights and thereby the best performance of the NN, the model is saved after each epoch where the error on the validation set has decreased. Additively,

if the loss value has not decreased for a certain number of epochs in a row the training is stopped, as generally, the loss will tend to further increase from thereon out. This way not only will the best model be used but a large amount of unnecessarily spent computational time can be saved. By fixing this condition to the validation set we can ensure that only performance improvements that are independent of the exact GMF structure and source positions are kept.

5.4. Performance measures

There are many different ways to evaluate the performance of a machine learning model. The measure of choice is highly dependent on the application in use though. While the previously introduced accuracy is often the beginner's default, there are many downsides to it as it does not capture all the aspects of the performance of a NN. It is highly influenced by an imbalance in the distribution of the classes and the chosen threshold on the prediction output. For example, if a binary data set consists of a class imbalance of 90:10 and the model predicts all data points as the majority class, the accuracy would still give a value of 90% without any actual prediction ability. Also, for a fixed chosen threshold of e.g. ≤ 0.5, the accuracy does not account for how well each class is predicted. Meaning, a model that outputs a prediction note with values for the background very close to zero, and the positive class very close to one, could yield almost the same accuracy as a model where the predictions are distributed around 0.4 and 0.6 respectively. This is why other measures than accuracy are needed to capture the full performance of a NN.

Receiver Operating Characteristic

The Receiver Operating characteristic (ROC) is a threshold-independent performance measure [97]. It compares the true positive rate (TPR) and the false positive Rate (FPR) at all thresholds on the output node. The TPR, which is also commonly called **recall**, is defined as

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{5.10}$$

Here, FN corresponds to the number of false negatives and TP to the number of true positives. The recall is also referred to as sensitivity and is here the fraction of correctly predicted source events with respect to all possible predictable source UHECRs in the entire data set.

The FPR on the other hand is defined as

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{FN}}, \tag{5.11}$$

where FP are the false positive predictions. The ROC curve is created by plotting the recall and the FPR values for different thresholds on the output node, shown in Figure 5.9.

The area underneath the curve (AUC) is then a measure of the performance of the NN that is independent of the threshold, where a value of one is a perfect score and 0.5 is a random 50:50 classifier.

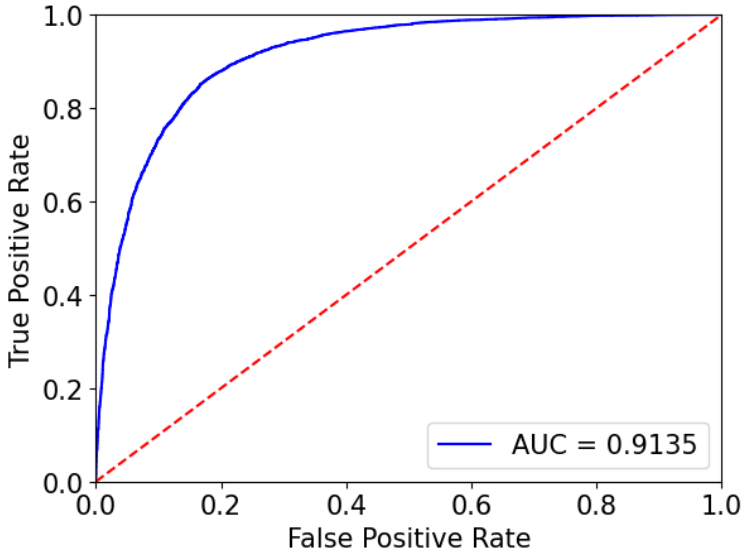


Figure 5.9.: ROC curve of the test data set of a mixed composition originating from Cen A alone. The AUC is shown in the bottom left corner and corresponds to the area underneath the blue curve. The red line indicates a random 50:50 classifier with no prediction ability.

The AUC captures the entire performance of the NN as it also measures how well the predictions are ranked. It is well suited for a balanced data set between both classes yet in the case of high imbalance, it becomes unrepresentative as in the regime of only a few TPs and FPs, small changes can have a large influence on the ROC.

Precision-Recall-Curve

The Precision-Recall-Curve (PRC) is a performance measure that is created by calculating the precision and the recall for different thresholds on the output node. The precision is defined as

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{5.12}$$

where TP and FP are the true and false positive predictions respectively. Here, the precision is the fraction of the correctly predicted source UHECRs with respect to the total number of predicted source events, those that are true, and those that are not. Just like for the ROC curve the precision and the recall are calculated for different cuts on the output node. An example of a PRC is shown in Figure 5.10

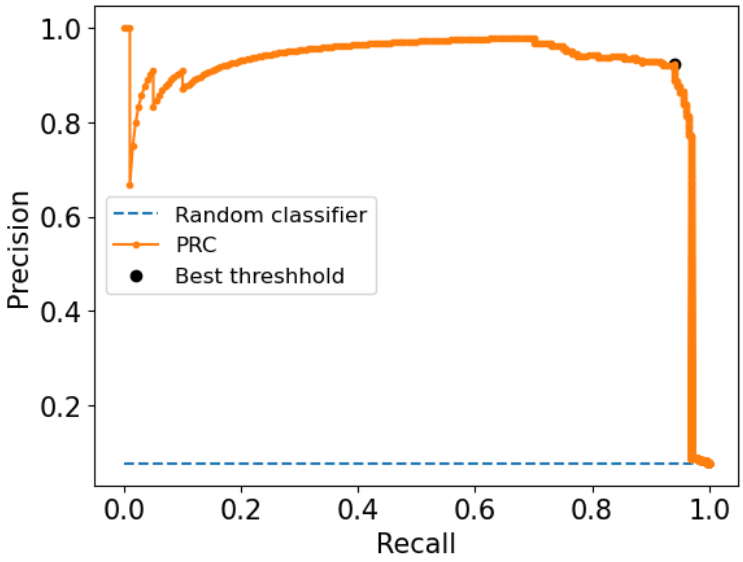


Figure 5.10.: The Precision-Recall-Curve of a proton test data set consisting of 2400 background CRs and 200 source CRs. This curve has an F-score score of 0.931 and the value closest to the top right corner corresponds to the best threshold and thereby the best performance.

The PRC focuses on the prediction ability of the positive class while being resistant to a high-class imbalance towards the background. It is the favoured performance measure for highly unbalanced data sets as it focuses on minority positive class. When testing the NN for different ratios of source events embedded in a large isotropic background it is a well-suited performance measure. By calculating the maximum of the harmonic mean of the recall and precision, the so-called F-score [98]

$$\mathrm{F-score} = \frac{2 \times \mathrm{Precision} \times \mathrm{Recal}}{\mathrm{Precision} + \mathrm{Recall}}, \tag{5.13}$$

the results of the PRC curve can be compared. An F-score of 1.0 is optimal while the worst possible one changes with the distribution of the data set indicated in Figure 5.10 as the blue line. We consistently used the ROC-AUC when evaluating a balanced and the PRC when evaluating an imbalanced test data set.

Furthermore, both the AUC and PCR can be used to derive the best possible threshold on the output node by determining the point in the curve that encloses the largest area with respect to both axes. This is shown in Figure 5.10 as the black dot. We thus use the best threshold to accomplish the best possible prediction performance of the classifier.

5.5. Deep neural network structure

The final structure of the constructed deep neural network is shown in Table 5.2. We chose a batch size of 64 resulting in the dimension of the input layer being (64, N) where

N corresponds to the number of input features discussed in the Section 5.2. For the implementation of different functions and hyperparameters we used PyTorch [99], which is a deep-learning Python library. A deeper and more complex structure showed no performance improvements. For the gradient descent task, we used the AdamW optimizer

Table 5.2.: Deep neural network structure

Layer	Layer Type	Number of Hidden Units	Activation Function
1	Linear	32	SELU
1	Dropout(0.3)	-	-
2	Linear	64	SELU
2	Dropout(0.3)	-	-
3	Linear	128	SELU
3	Dropout(0.3)	-	-
4	Linear	64	SELU
4	Dropout(0.3)	-	-
5	Linear	32	Sigmoid

[100]. We tried out several different optimizers, activation functions and combinations of hyperparameters however, we consistently observe better or at least equal results with the stated selection. Yet, some slight adjustments have to be made to this structure for the different data sets of the analysis e.g. with more complex deflection patterns we experience better results for higher dropout probabilities.

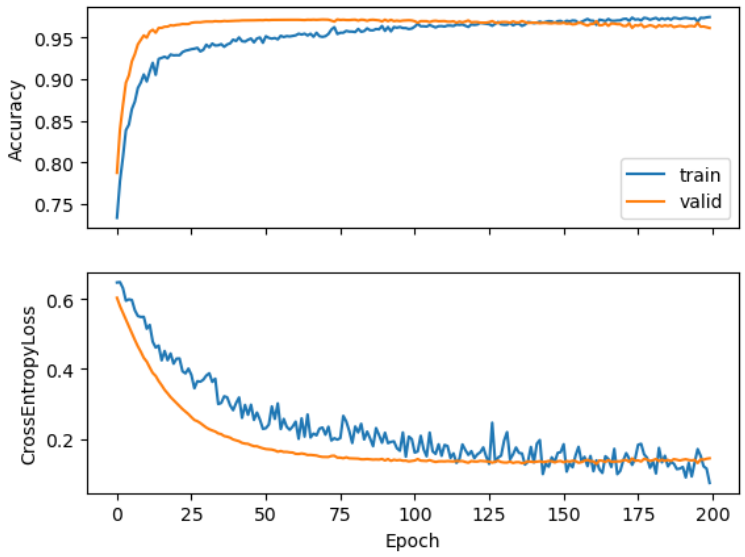


Figure 5.11.: Accuracy and loss of final tuned neural network trained on a pure proton composition. The blue and orange curve corresponds to the training and validation data set respectively.

Algorithm	Type	Reference
XGBoost	Boosted Decision Tree	[101]
LightGBM	Boosted Decision Tree	[102]
SGDClassifier	Logistic Regressor	[103]
ElasticNet	Logistic Regressor	[103]

5.5.1. Combined neural network

Additionally to the single deep neural network approach, we test extending the method by making use of several different pre-designed machine learning algorithms and adding their prediction values to the data set. We thereby hope that other algorithms like random forests might catch certain structures in the data set better than our deep neural network. A list of the machine algorithms used is given in Section 5.5.1. We train each algorithm individually and then run a linear regression on the data set consisting of all prediction nodes of the different machine learning models used, including that of our deep neural network. This is a commonly used method in machine learning to make use of the specific different strengths of the different algorithms. We then optimize the hyperparameters of each machine-learning model using a hyperparameter optimization framework called Optuna [104]. However, as the algorithms perform worse than our deep neural network on all of the different data sets of Chapter 6, no performance increase could be observed. This could be due to the very limited information given. For more complex data sets with a large number of input features with very different characteristics, we would expect an improvement, however, due to only having the spatial information plus the energy it is possible that the deep neural network alone is able to catch the full information.

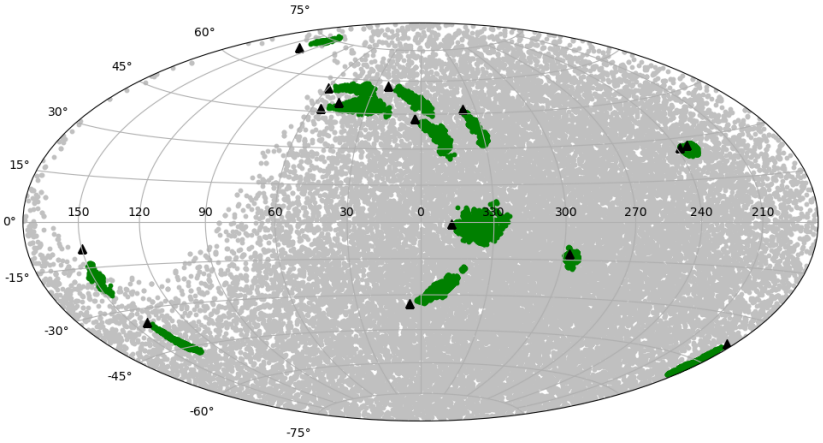
CHAPTER 6

SOURCE EVENT DETECTION USING DEEP LEARNING

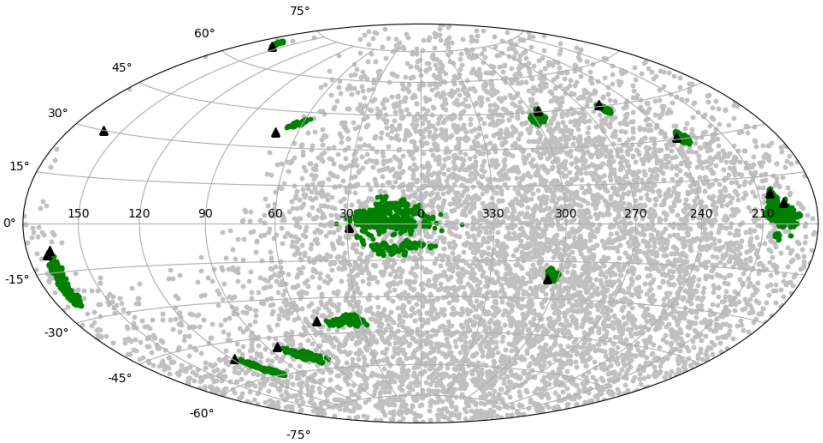
In this chapter, we study the performance of the NN for several different scenarios. In each case, we compare the results in both the limit of infinite statistics for randomly distributed sources and for different ratios of source-to-background UHECRs for a few selected candidate sources. At first, we choose a pure proton composition, simplifying the distinction for the NN due to more local clustering (Section 6.1) followed by a pure Nitrogen composition in Section 6.2. In both of these pure composition scenarios, we also analyse the feature importance of the trained networks. Afterwards, we consider a mixed composition of 5 representative nuclei for the same few candidate sources in and for Cen A alone in Section 6.3.1. Afterwards, we apply the differently trained models to actual data taken by Auger to look for possible source clustering events in Section 6.4. Finally, we shortly compare our work to other methods used to search for magnetically-induced patterns in the UHECR arrival directions in Section 6.5.

6.1. Pure proton composition

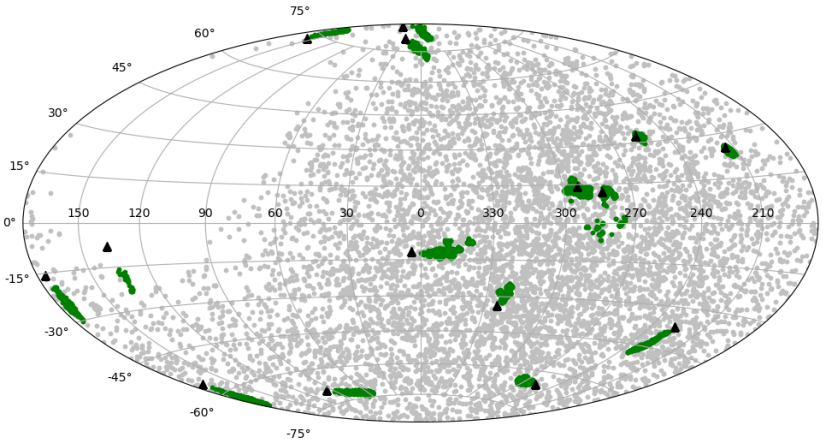
As already stated, we create each data fraction with different seeds for the random turbulent and striated magnetic field component and different random source positions. However, we use three seeds to create the training sets to avoid the NN learning just one realisation. The proton data set is shown in Figure 6.1. It can be seen that in this simplified scenario, the deflection is only a few degrees. The neural network is able to separate the isotropic background from the source UHECRs with an AUC score of 0.9958 on a balanced data set of 15000 particles. The prediction output of the proton test set is shown in Figure 6.2 where it can be seen that the network can clearly distinguish the two classes from each other. Next, we look at the performance for a variation of the coherence



(a) Training data



(b) Validation data



(c) Test data

Figure 6.1.: The three data partitions of the proton-only simulations. The training set (a) consists of 35000 samples per class and was created with two different realisations of the JF12 model. The validation (b) and the test (c) set were also created with different seeds and source positions, consisting of 7500 UHECRs of each class. The green and grey dots show the arrival directions of the source and isotropic background UHECRs respectively while the black crosses indicate the source direction.

Table 6.1.: Comparison of the performance of the proton composition NN applied to different test sets containing 7500 background and 7500 source UHECRs originating from 15 sources. The UHECRs were propagated with different coherence lengths of the random turbulent magnetic field component.

Lc [pc]	12.5	25	37.5	50
AUC	0.9952	0.9941	0.9948	0.9941

length of the random turbulent field. Table 6.1 shows the AUC for increasing coherence lengths for a data set of 7500 UHECRs originating from 15 random source positions embedded in 7500 isotropic background particles. The distinction ability does not worsen as the small deviation can be acquainted with just training fluctuations. Due to the relatively small deflection of protons at these energies and the large source particle count, the neural network is not much influenced by a significantly increased coherence length.

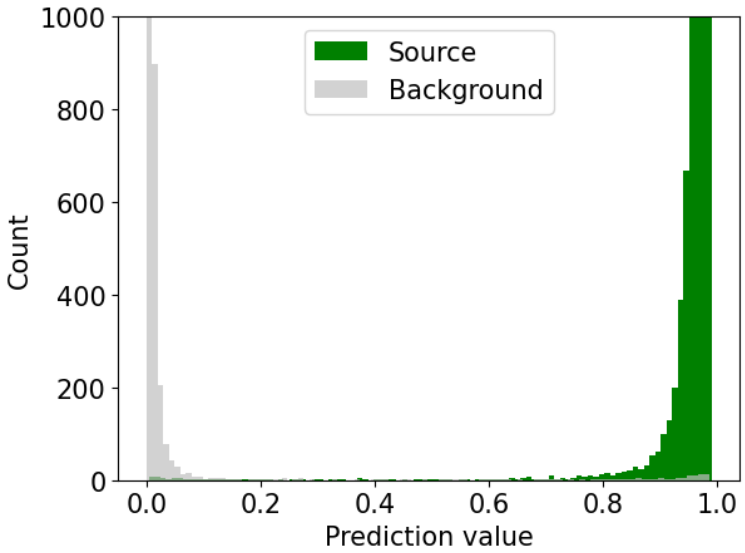


Figure 6.2.: Ouput node of NN applied to test data set. The NN was trained and tested on a proton-only composition. The true class of sources and background UHECRs is presented as green and grey respectively. The NN can clearly distinguish both classes.

This limit of infinite statistics where each source receives such a large amount of UHECRs is of course highly unrepresentative, yet we analyse the performance here just from a machine learning perspective. The data taken by Auger does clearly not show this degree of clustering. As we do not know if and how much single sources might contribute to the UHECR flux observed here on Earth, we test the performance of the NN for different source-to-background particle ratios. For this we keep the total amount of UHECRs at 2600, motivated by the selected Auger data from [105], and compare the performance of the NN more precisely for 4 different candidate sources and the coordinates of two hot spots (HS), as stated in [49]. Although the hot spots are no actual sources we choose to use their positions as additional origins for the simulations. Due to the very small dis-

tance between Cen A and HS1, we can test the NN when two sources are this close to each other and with the HS2 we get another statistic for the lower half of the Galactic plane. The Galactic coordinates of these candidates and hot spots are listed in Table 6.2 together with the numbering scheme used for them in the skymaps.

Table 6.2.: Galactic coordinates of candidate sources and two hot spot regions used as origin for the simulations of the source UHECR.

Source	Longitude	Latitude	Number
Cen A	309.51	19.42	0
Cyg A	76.19	5.76	1
Fornax A	240.16°	56.69°	2
M87	283.78	74.49	3
HS1	308°	26°	4
HS2	275°	-75°	5

Due to the increasingly high level of imbalance between the two classes, we use the F-score, see Equation (5.13), as the performance measure. The results are shown in Table 6.3.

Table 6.3.: The performance of the NN after obtaining the best threshold via the F-score for different source-to-background ratios of 2600 UHECR protons originating from 6 candidate source positions, as stated in [49].

Source CRs	F-score	Errors	TP	TN	FP	FN
1000	0.978	45	964/1000	1587/1600	14	31
500	0.954	47	476/500	2077/2100	23	24
200	0.931	28	188/200	2384/2400	16	12
100	0.830	34	83/100	2480/2500	17	17
50	0.653	34	32/50	2529/2500	16	18

It can be seen that the performance only decreases relatively little from 1000 to 200 total source particles. First at 100 and especially at 50 total source events the F-score drops significantly to a value of 0.653. In this case, each source received on average less than 10 particles causing the detection ability of the NN to stagnate. However, after calculating the best threshold on the output node we are still able to correctly find 32/50 particles with 16 FPs and 18 FNs. When considering 100 source protons at a ratio of 1:25, the NN can find 83/100. For this ratio, we show the full skymap as a function of the predicted output of the NN Figure 6.3. Here, the grey dots and green dots correspond to a background and source prediction respectively while a red point on top indicates if this prediction is wrong, so that all dots without correspond to correct predictions. The initial source positions are numbered and again indicated by black triangles. The corresponding source to the numbering scheme can be found in Table 6.2.

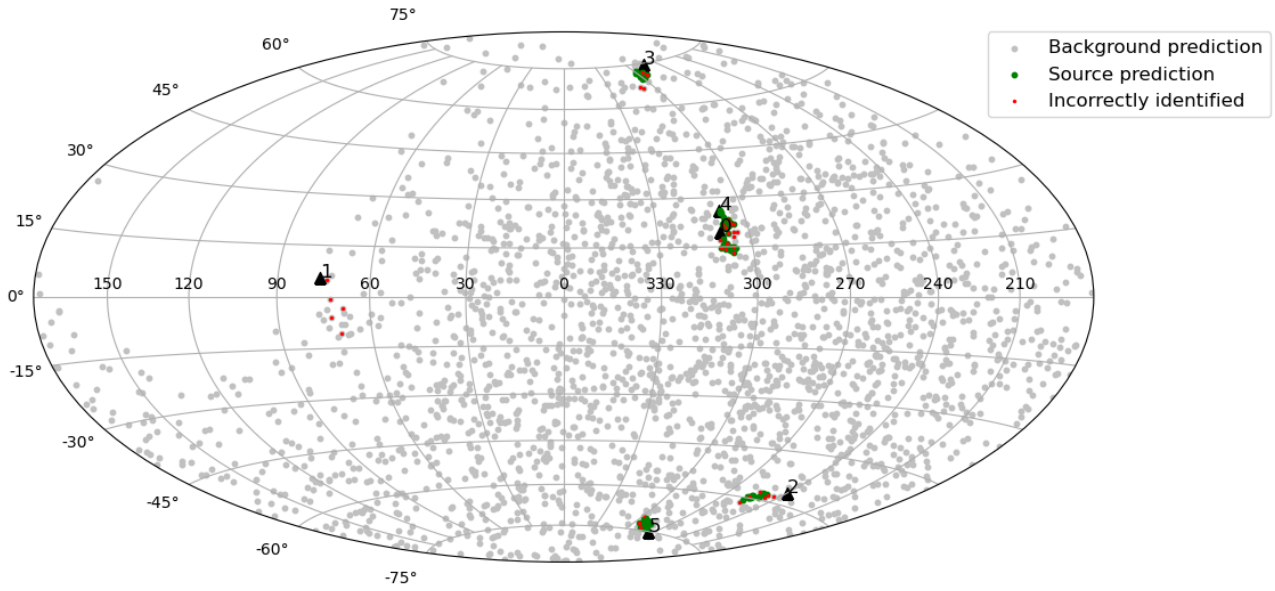


Figure 6.3.: Skymap in Galactic coordinates of 100 source events from 6 candidates sources embedded in 2500 isotropic background UHECRs. The grey and the green data points correspond to the NN-predicted background and source particles respectively. The red dots shown on top indicates a false prediction.

We further analyse the detection ability of the NN for this scenario for the sources individually. In Figure 6.4 a zoomed-in sky segment is shown of 12 UHECR protons (indicated with circles) originating from M87 (black triangle). The neural network detected 10/12 source particles correctly and has no false predictions for a threshold of 0.3 on the output node. However, the detection ability of the NN differs for different regions in the Galactic skymap. Due to the magnetic field being strongest in the Galactic plane, the particles experience a stronger deflection resulting in the performance of the NN worsening towards and in the plane. While 10/12 particles are found for M87 with no false predictions the FP and FN are significantly higher in the region of Cen A and Cyg A which lie close and in the Galactic plane.

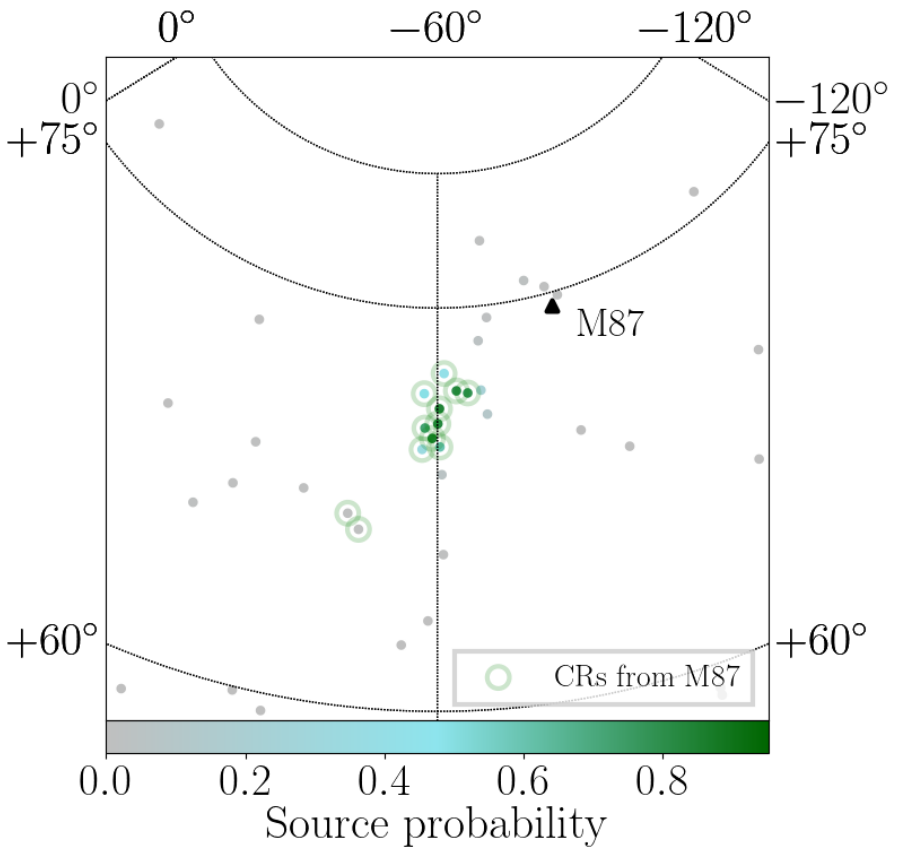
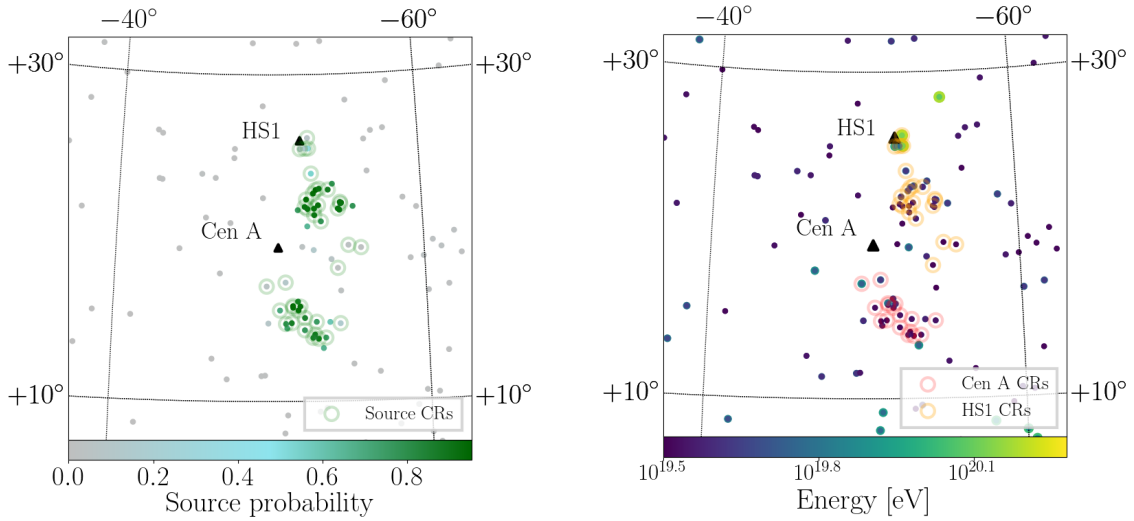


Figure 6.4.: Prediction values of a deep neural network trained on the arrival direction of UHECRs and tested on 2500 isotropic background and 100 source protons originating from 6 candidate sources. Shown is a sky segment in Galactic coordinates of the region around M87 (black triangle) with 12 UHECRs coming from that source (shown as circles). The colour indicates the probability of a particle being identified as a source. A value close to zero (grey) corresponds to a background and one closer towards one to a source prediction

When we look at the region of Cen A and HS1, we can see, that although these sources got a larger fraction of source particles due to the higher exposure in this part of the sky, the performance of the NN worsens. This could be explained by the nature of the two sources being so close to each other making it more difficult for the NN to find the deflection structure as it has no actual knowledge that there are two sources in this region. Furthermore, the magnetic field strength becomes increasingly stronger toward the plane and also thereby the effect of the random fields. We can observe a slight curvature for the particles from Cen A and a ring-like structure from HS1. In comparison, the UHECRs from M87 follow a more linear deflection as can be seen in Figure 6.4.



(a) Prediction output of the NN.

(b) Particle energy of the same sky segment shown in (a).

Figure 6.5.: Sky segment of Cen A and HS1 in Galactic coordinates of 100 source CRs from the 6 candidate sources embedded in an isotropic background of 2500 particles. Shown are both (a) the prediction value of the neural network and (b) the particle energy of the same region. The true source particles are marked by coloured circles.

As we are trying to distinguish a fully random isotropic arrival-direction distribution from one that shows a certain degree of correlation in the deflection, any effect that induces a higher degree of randomness to the latter will lower the distinction ability of the NN. For particles of lower rigidity, this effect will be amplified. We train the neural several times to compare the number of source particles N_s needed for each source to correctly identify a certain fraction of UHECRs after implementing the best threshold via the highest F-score. However, we note that the optimal threshold differs when looking at the performance of each source individually. While a threshold of 0.3 is optimal for M87, Fornax A needs one of 0.1 for the best performance in this region. This on the other side will cause M87 to get an additional two FN predictions. Thereby for such an analysis, it is crucial to test each source for different quantities of emitted source particles individually and averaging the results over several training runs. However, this could not be conducted in the scope of this work. Nevertheless, we can conclude by comparing the individual performance of several runs together with the results shown in Table 6.3 that for M87, HS2 and HS1 approximately $N_s = 10$ source particles were needed for a TP rate of above 80% while for Fornax A and Cen A 15 and 20 particles where needed respectively to not acquire an equal amount of FPs as TPs. The detection of Cyg A shows consistently by far the worse results which can be explained by its location directly in the Galactic plane and on the border of the observable sky of Auger. With such a low exposure together with lying in the plane, the UHECRs from Cyg A showed a significantly lower source probability value than any other source and a threshold below 0.05

with $N_s = 30$ was needed to detect half the particles. Additionally, we applied an input variable analysis of the trained NN using the Python SHAP library [106] to determine the influence of the different given features. In Figure 6.6 we illustrate the feature importance ranked from top to bottom of the proton neural network. The colour indicates the value of the feature and the SHAP value, the influence of this feature on the prediction output of that example. In this case, a negative SHAP value will correspond to a shift towards a background prediction and a positive SHAP value towards a source prediction.

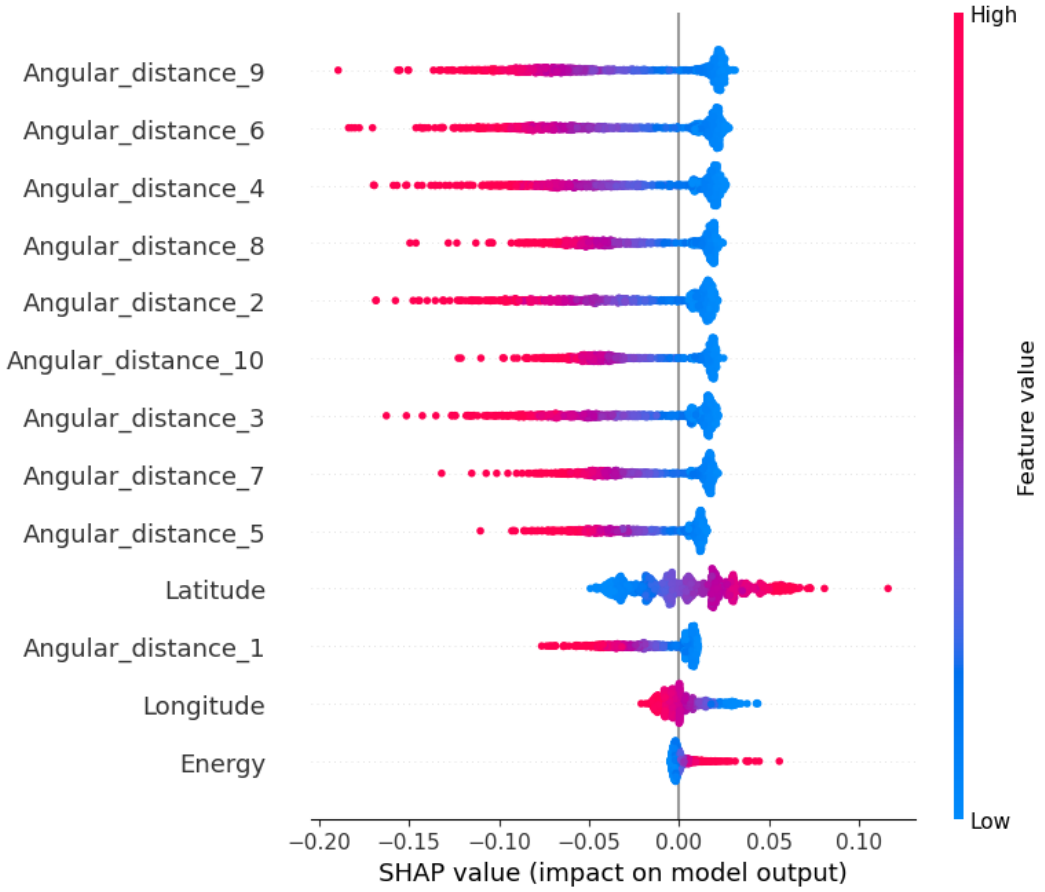


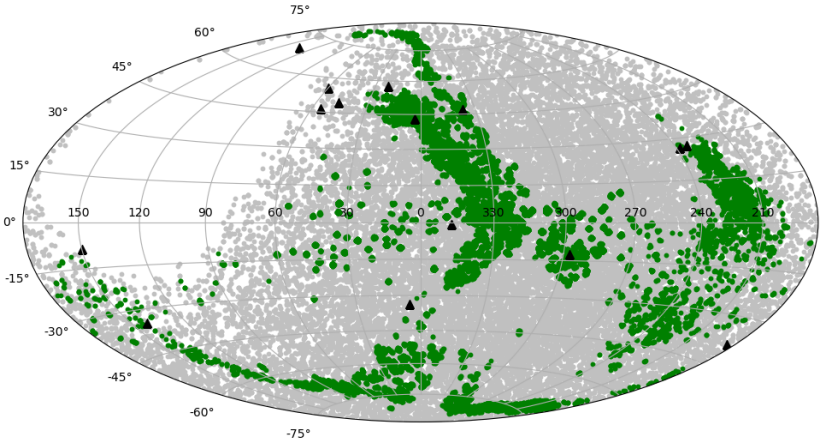
Figure 6.6.: Shap summary plot of the proton neural network. The features are ordered by their influence from top (high influence) to bottom (low influence) on the distinction. The SHAP value indicates how the effect of a feature value corresponds to the class. A negative SHAP value of a feature will shift the prediction towards the background class and a positive towards the source class

We can see that the angular distance to the nearest neighbours has the largest influence, whereas a large distance shifts the prediction value towards the background and a small value towards the source prediction. The energy of the particles on the other hand ranks lowest in importance, however, large energy values have a positive influence on the prediction value as these experience less deflection. The reason why the energy ranks so low could be explained by the NN learning the local clustering of protons from the spatial information alone and thereby the energy information is less needed. We expect

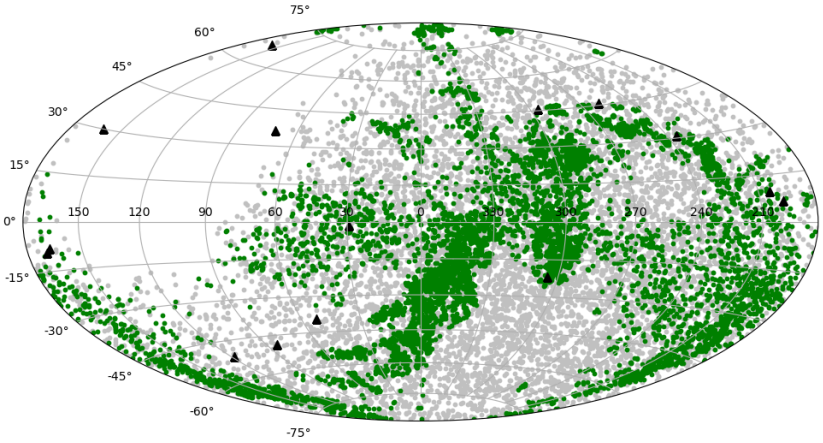
the importance of the energy information to increase for the Nitrogen deflections. Although we expect the latitude to have a higher influence on the NN than the longitude due to the stronger deflection within the plane, the clear tendency of higher latitudes shifting the prediction towards the source class and vice versa is unexpected. By further analysing the data set we realise that this can be an effect of different ratios of source to background events in the upper and lower half of the Galactic sky. The source-to-background events ratio is 0.85:1 in the lower and 1.27:1 in the upper half, thereby the NN could assume this anisotropy to be a physical property. This is not desired and an equal source-to-background ratio should be applied, if the latitude is kept as a feature and such a behaviour is not wanted. We tried replacing, the latitude and longitude with the exposure value as spatial information and excluding the latitude altogether, however, none showed an improvement on an unbiased test data set. As this was found in the later stages of the analysis a total recreation of the simulations was not possible. Furthermore, the influence of latitude is subdominant to most other features thereby we argue that although this is an undesired training artefact the imbalance is not too large and it does not alter the results significantly. We can observe a similar tendency for the longitude, however, its importance is even lower. To avoid such a training artefact in future analysis, we propose for an all-sky search to set the sources in the form of a grid over the entire observable sky.

6.2. Pure Nitrogen composition

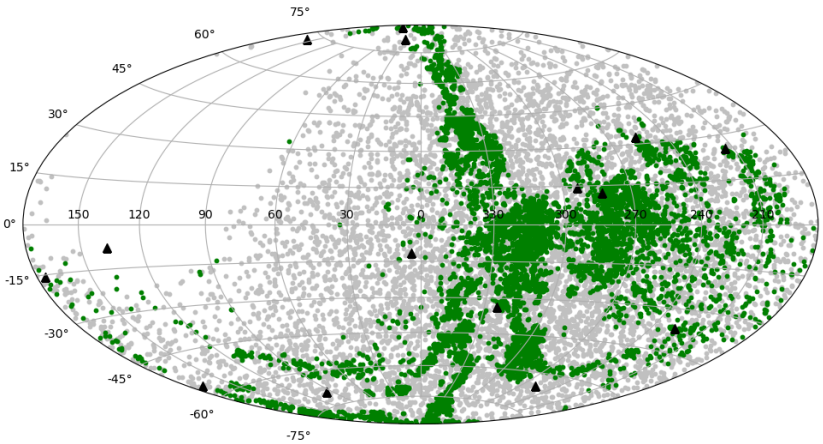
Next, we study the deflection of a pure Nitrogen composition. We chose the same source positions and data sizes for the three data fractions as in Figure 6.1. With a charge of $Z = 7$, Nitrogen has a significantly lower rigidity at energy above 32 EeV than protons and thereby a larger degree of deflection. The three data fractions are shown in Figure 6.7. The NN performs on the presented test data set with an AUC of 0.8775 which is a decrease of ≈ 0.14 to the pure proton composition. Again we tested the AUC score for the different coherence length variations of the turbulent magnetic field component. The results are shown in Table 6.4. Although all scores are slightly lower than the default coherence length of 5 pc, with 50 pc one scoring the lowest, the differences are insignificant. This indicates that the NN is able to distinguish the source from the background UHECRs in the limit of infinite statistics independently of the exact nature of the random turbulent magnetic field component.



(a) Training data



(b) Validation data



(c) Test data

Figure 6.7.: The three data partitions of the pure Nitrogen simulations. The training set (a) consists of 35000 samples per class and was created with two different realisations of the JF12 model. The validation (b) and the test (c) set were also created with different seeds and source positions, consisting of 7500 CRs of each class. The green and grey dots show the arrival directions of the source and isotropic background UHECRs respectively while the black crosses indicate the source direction.

Table 6.4.: Comparison of the performance of a NN trained on a pure Nitrogen composition and applied to test sets generated with variations of the coherence length of the random turbulent component of the JF12 GMF model.

Lc [pc]	12.5	25	37.5	50
AUC	0.8626	0.8751	0.8651	0.8590

Previously, the deflection of the protons with a minimum rigidity of 32 EV were only a few degrees. A 32 EeV Nitrogen UHECR on the other hand, would have a corresponding rigidity of ≈ 4.7 EV which is consistent with the degree of deflection shown in Figure 4.2. As at these rigidities, the deflection can differ by several degrees for the different variations we conclude that as desired the NN does not try to learn the precise GMF lens, but can handle a UHECR data set exposed to a different random turbulent and striated magnetic field component. As these are some of the most unknown aspects of the GMF we hereby present a method to avoid a certain degree of uncertainty in their characteristic properties. However, this behaviour would have to be tested in order to see if it holds for even stronger variations of these components.
As for the protons, we also test the NN for different background-to-source particle ratios. The results are shown in Table 6.5.

Table 6.5.: The performance of the NN after obtaining the best threshold via the F-score for different source-to-background ratios of 2600 Nitrogen UHECRs originating from 6 candidate source positions, as stated in [49].

Source CRs	F-score	Errors	TP	TN	FP	FN
1000	0.783	425	763/1000	1412/1600	188	237
500	0.595	425	312/500	1863/2100	237	188
200	0.347	385	101/200	2114/2400	286	99
100	0.163	339	33/100	2228/2500	272	67
50	0.142	122	9/50	2469/2550	81	41

The performance is significantly lower than for the pure proton composition, scoring only an F-Score of 0.142 when only 50 particles originate from sources. For 200 source particles, the NN detects approximately half correctly but at the cost of a high FP rate. These results are again gained after we apply the optimal threshold derived from the best F-score, meaning the cut for scoring the best precision while also obtaining the best recall. As the source events are clustered less local the nearest neighbour features are less likely to catch other source particles for a lesser abundance. Here, a better analytical pre-processing of the data set would be needed to capture these particles, yet none of our methods shown in Section 5.2 show any increase in performance. For heavier nuclei with a low abundance of source events, random alignments of background and source particles will become more of an issue and magnetically-induced patterns less dominant. However, when we increase the threshold on the output node, the FP rate can be signifi-

cantly decreased and we thereby look at the events with the highest prediction values. In the case of 500 source Nitrogen particles 60 source UHECRs can be detected with only 3 FPs, thereby scoring a precision of 0.95. For the 200 source particle scenario, the 9 and in the 50 source particle scenario the 4 UHECRs with the highest prediction value are TPs with no FPs. Thereby, although the NN might output less reliable results with respect to the entire data set, the highest prediction values show to be mostly correct. When repeating the training process several times we receive consistently similar results. We also show the feature importance of the Nitrogen-trained network in Figure 6.8 where we can see that the energy importance increased significantly in comparison to the proton-trained one. Also, the 7th, 8th, 9th and 10th nearest neighbour distances show the highest feature importance showing that the NN has captured higher deflections to a certain extent. Furthermore, the importance of the energy information has increased notably.

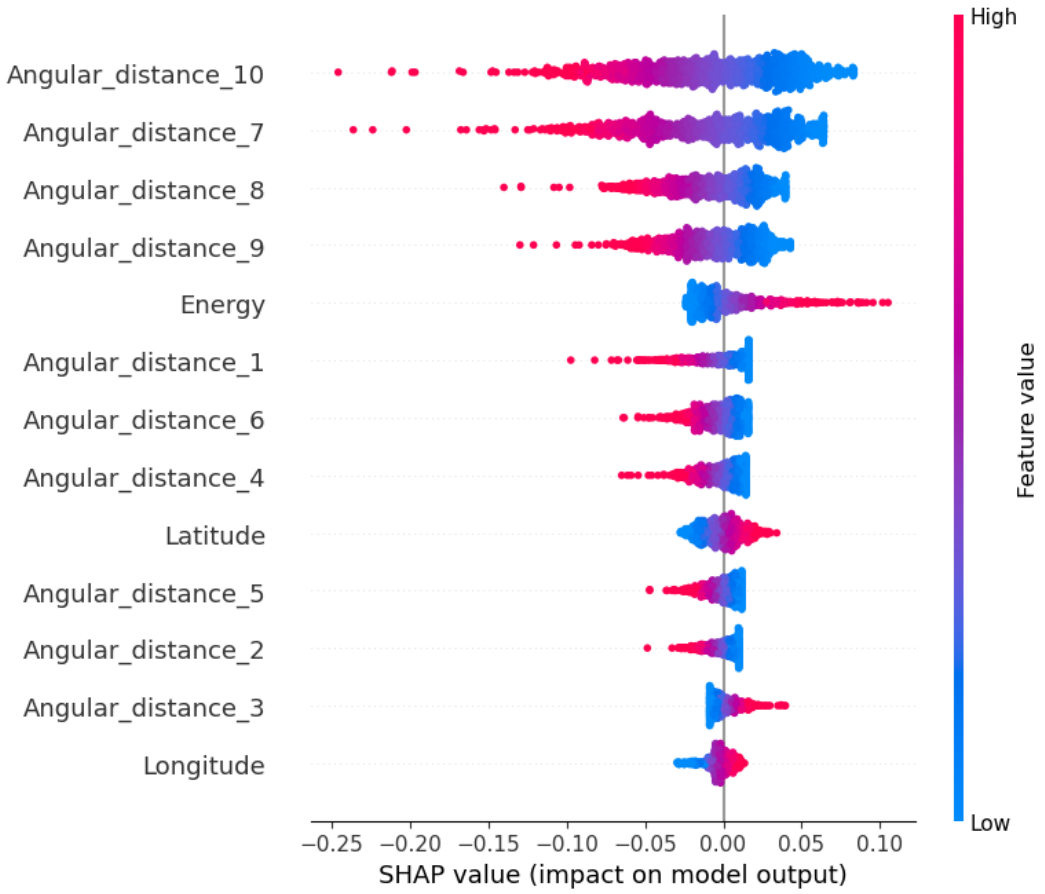


Figure 6.8.: Shap summary plot of the Nitrogen neural network. The features are ordered by their influence from top (high influence) to bottom (low influence) on the distinction. The SHAP value indicates how the effect of a feature value corresponds to the class. A negative SHAP value of a feature will shift the prediction towards the background class and a positive towards the source class.

Both the longitude and the angular distance to the 3rd nearest neighbour show a flipped correlation compared to the proton model. As the nearest neighbours will less

likely be of the same source due to the significantly lower rigidity of the Nitrogen UHECRs, the closest neighbours will more commonly be background particles too. The orientation for the longitude changes for the different subsamples used to calculate the SHAP values. We calculate the SHAP values by selecting 1000 examples of an unbiased test data set and observe these changes in behaviour for different selections. For the nitrogen training data, the source-to-background ratio distribution is almost perfectly balanced at 1:1, thus the NN should only find relevant information in correlation with other features. Therefore a switch of orientation is possible and expected for different subsamples. We can safely say that the presented input variable analysis clearly indicates a higher importance of the correlation of the angular distances and the energy of the deflected source events.

6.3. Mixed composition

Next, we train and apply the NN on a mixed, 62.9% iron-dominated composition as stated in Section 4.4. The data set partitions are shown in Figure 6.10 where we can already visually observe that the source particles now also stretch over the entire observable sky. Some clustering of the lighter elements can still be seen near the source position. As the data now consists of several elements that immensely differ in their deflection behaviour but no information on the particle's charge is given, the distinction is severely complicated.

The NN reaches an AUC score of 0.740 on the 15000 equally distributed test set and thereby a value of roughly 0.14 lower than for the pure Nitrogen scenario. Both the missing charge information and the increasingly larger influence of the random magnetic field components on heavier nuclei result in a higher probability of random alignments of background and source events.

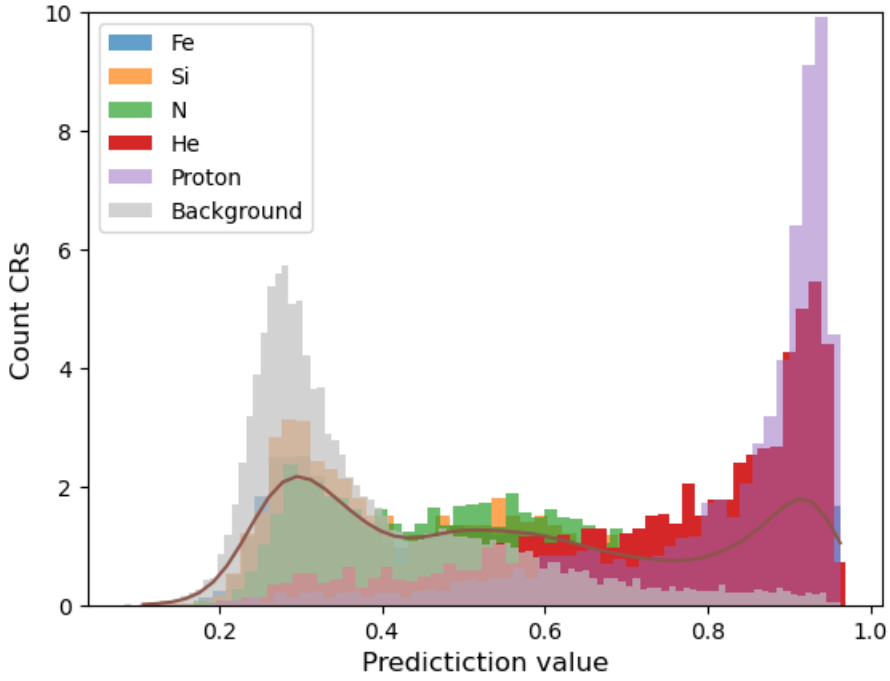
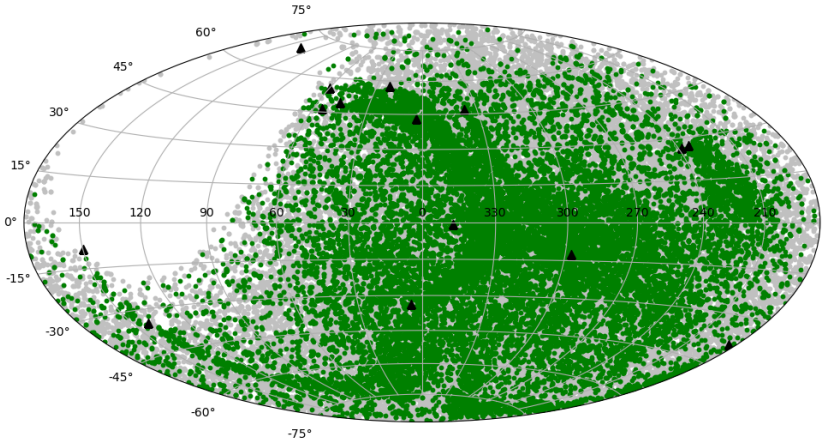
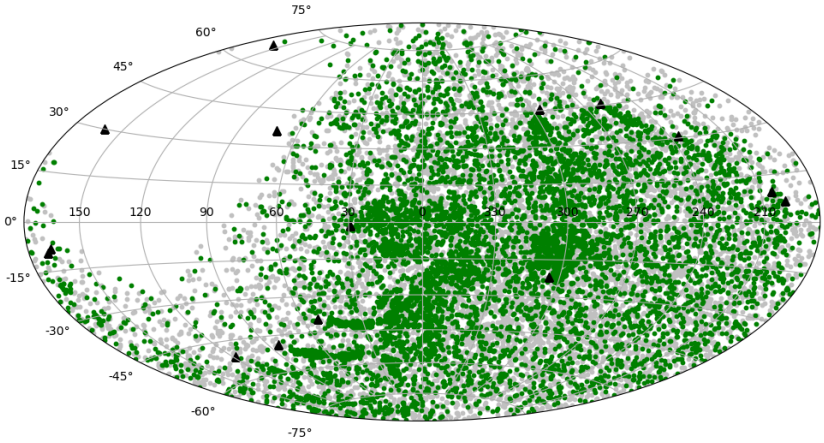


Figure 6.9.: Denisty histogram of the output node of the NN trained on a mixed composition. Shown are the normalized prediction values for each element of true source UHECRs individually. For comparison, the background prediction distribution is given in grey and the overall distribution is indicated by the brown curve. The NN performs best on protons, then helium and worst on Silicon and Nitrogen.

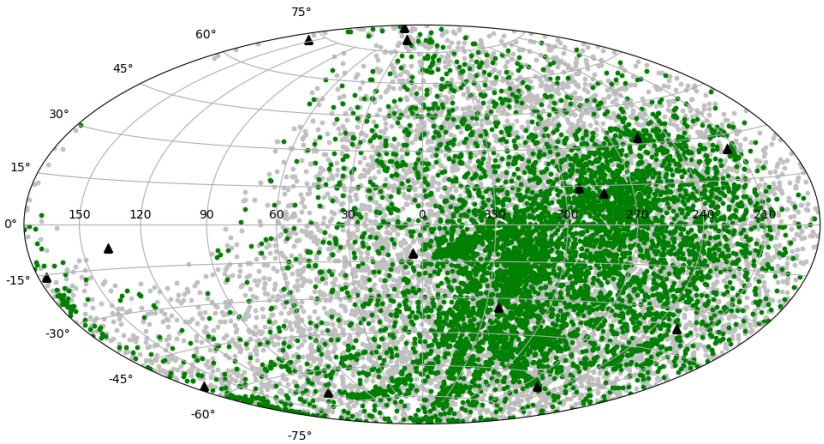
We look closer at the prediction output in the limit of infinite statistics of the source UHECRs for each element individually, shown in Figure 6.9. Due to the different degrees of deflection, a tendency of performance lessening for lower rigidities would be expected. However, we can observe that the NN performs best on protons, then Helium, but worst on Silicon and Nitrogen and not Iron.



(a) Training data



(b) Validation data



(c) Test data

Figure 6.10.: The three data partitions of the iron-dominated mixed composition simulations. The training set (a) consists of 35000 samples per class and was created with two different realisations of the JF12 model. The validation (b) and the test (c) set were also created with different seeds and source positions, consisting of 7500 CRs of each class. The green and grey dots show the arrival directions of the source and isotropic background UHECRs respectively while the black crosses indicate the source direction.

This could be explained by the severely higher number of iron examples in the data set, resulting in a far denser energy binning. With 62.9% of the UHECRs being iron particles this is almost 4 times the fraction of Silicon, thus the iron UHECRs will on average have smaller energy differences, creating possibly better detectable patterns.
Just as for the proton and Nitrogen simulations, the AUC for the different coherence length variations remains almost unaffected, again indicating that the NN has not learned a specific realisation of the GMF model, see Table 6.6. However, the model performs significantly worse on highly deflected heavy nuclei which are subject to the largest influence of changes in the turbulent and striated magnetic field components. Therefore it has to be considered that the AUC comparison may be less representative due to the lower performance.

Table 6.6.: Comparison of the performance of a NN trained on an iron-dominated mixed composition and applied to test sets generated with variations of the coherence length of the random turbulent component of the JF12 GMF model.

Lc [pc]	12.5	25	37.5	50
AUC	0.7394	0.7402	0.7337	0.7293

Table 6.7.: The performance of the NN after obtaining the best threshold via the F-score for different source-to-background ratios of 2600 UHECRs of 5 representative elements originating from 6 candidate source positions, as stated in [49].

Source CRs	F-score	Errors	TP	TN	FP	FN
1000	0.675	737	766/1000	1097/1600	503	234
500	0.486	608	286/500	1706/2100	394	214
200	0.334	324	80/200	2196/2400	204	120
100	0.250	114	19/100	2467/2500	33	81
50	0.088	63	2/50	2535/2550	15	48

Looking at the performance of the ratio comparison, shown in Table 6.7, the NN performs noticeably worse on almost all ratios. The prediction skymap for the 500 source particle scenario is shown in Figure 6.11, where each source except for Cyg A contributed roughly 90 events of the 500 events. Although some of the more moderate deflections can be caught by the network it is not able to catch almost any higher-deflected ones. A significant fraction of the UHECRs arrive alone or only a very few other source UHECRs within the next 10 nearest neighbours. The NN thereby is not able to find a significant enough correlation with other source particles in order to detect them. Due to the significantly more complex structure, the high uncertainties and the magnitude of deflections in this scenario, we tried to extend the data set by considering more nearest neighbours and several variations of the correlation coefficient extension introduced in Section 5.2. However, none showed any improvement in capturing these events. In the case of only 50 source particles, the NN shows basically no prediction ability anymore and for both 100 and 200 source events it is only able to catch the lighter nuclei and some of the most

energetic heavier ones. We thereby conclude that the NN fails in the iron-dominated mixed composition scenario and that both higher statistics and better knowledge of the GMF deflections that have to be integrated into feature selection are needed. In the rigidity regime of a few EVs, the difference in the degree of randomness between the isotropic background and the source event is just too insignificant.

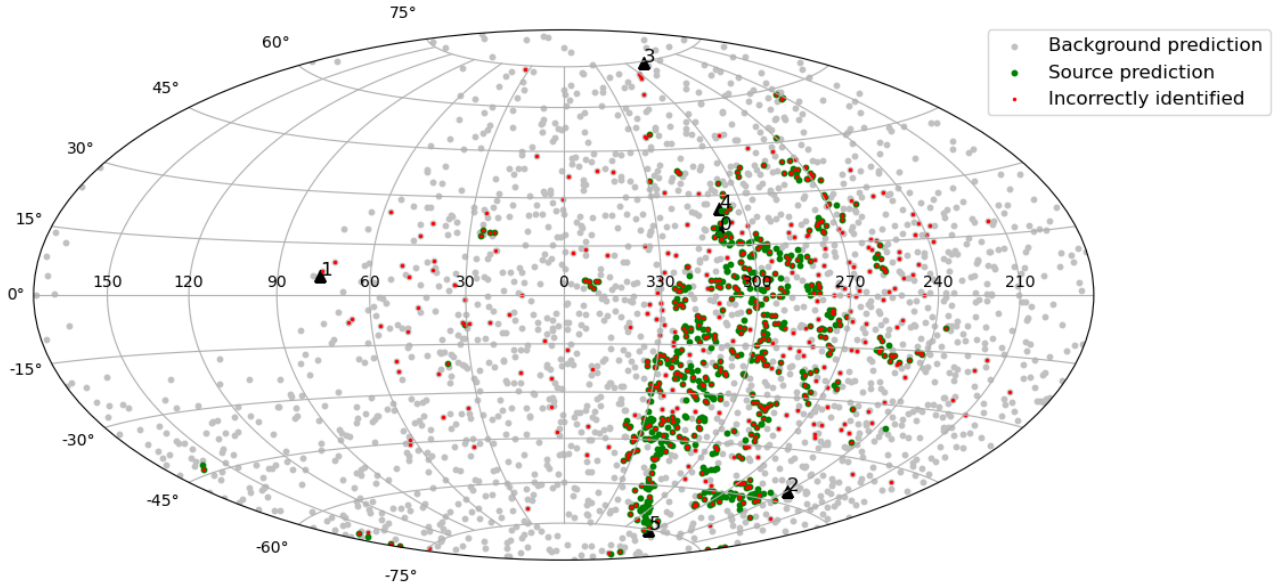


Figure 6.11.: Skymap in Galactic coordinates of 500 source events from 6 candidates sources embedded in 2100 isotropic background UHECRs. The grey and the green data points correspond to the NN-predicted background and source particles respectively. The red dots shown on top indicates a false prediction.

6.3.1. Cen A

As the NN fails for the previous scenario, we exclude one source of randomness and train on a single source instead of a set of randomly distributed ones. For this, we select Cen A as it is one of the most probable source candidates to date. We train the neural network again on 70000 equally distributed examples and scores an AUC of 0.9135 on the test set which is an increase of 0.17 then for the random sources. The prediction output after deriving the best threshold for the test set is shown in Figure 6.12 Although this is a highly unrepresentative scenario we show here that the NN is able to detect correctly a fraction of UHECRs that are largely deflected from Cen A with angular distances above 100°. In this scenario, the neural network will receive no "contamination" of other source events originating from different sources. Also, Cen A receives the full fraction of source events resulting in a smaller energy binning of the emitted UHECRs.

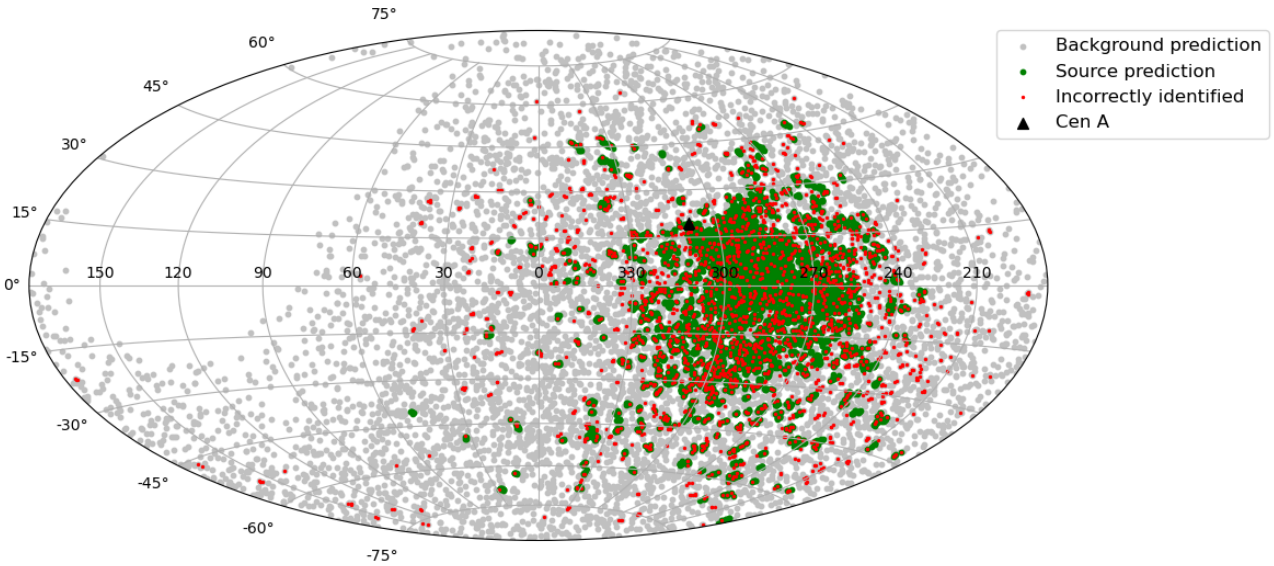


Figure 6.12.: Skymap in Galactic coordinates of the test data set consisting of 7500 background and 7500 particles from Cen A. The grey and the green data points correspond to the NN-predicted background and source particles respectively. The red dots shown on top indicates a false prediction.

Table 6.8.: The performance of the NN after obtaining the best threshold via the F-score for different source-to-background ratios of 2600 UHECRs of 5 representative elements originating from Cen A alone.

Source CRs	F-score	Errors	TP	TN	FP	FN
1000	0.914	178	944/1000	1478/1600	122	56
500	0.793	249	439/500	1932/2100	168	61
200	0.614	172	137/200	2291/2400	109	63
100	0.325	257	62/100	2281/2500	219	38
50	0.161	73	7/50	2520/2550	30	43

In several regions of the sky, the NN is able to find largely deflected small multiplets consisting of 7 iron UHECR events or more. However, these quantities and this energy binning are not present for the more realistic background-to-source ratios presented in Table 6.8. Although a slight performance improvement can be seen, only 7/50 source particles are correctly detected in the 50 source event scenario. With such few events and several different elements, no significant enough patterns can be found even when just considering a single source. We conclude here that the current information that we have given to the NN is not sufficient enough to capture magnetically-induced patterns of a mixed and heavier-nuclei-dominated UHECR composition. We want to note here that for a "target search" for events from source candidates the energy spectrum could be derived individually for each source. As the spectrum here on Earth is dependent on the distance travelled alone we simulated a 1D extragalactic propagation for the four

presented candidate sources, see Appendix A Section A.1. These derived energy spectra could be implemented into the simulations. However, it has to be ensured that the total energy spectra of both the isotropic background and the source candidates follow the one observed at Auger. With this, the neural network would be able to better distinguish the two classes as the energy information will not anymore be the same.

6.4. Auger data

Finally, we apply the differently trained neural networks on data taken by Auger consisting of 2635 UHECRs above and energy of 32 EeV, taken from [105]. The data is shown in Figure 6.13 and is visually in a high level of agreement with several of our simulations with just a few source particles. Some small-scale clustering can be observed, however, this is consistent with background fluctuations.

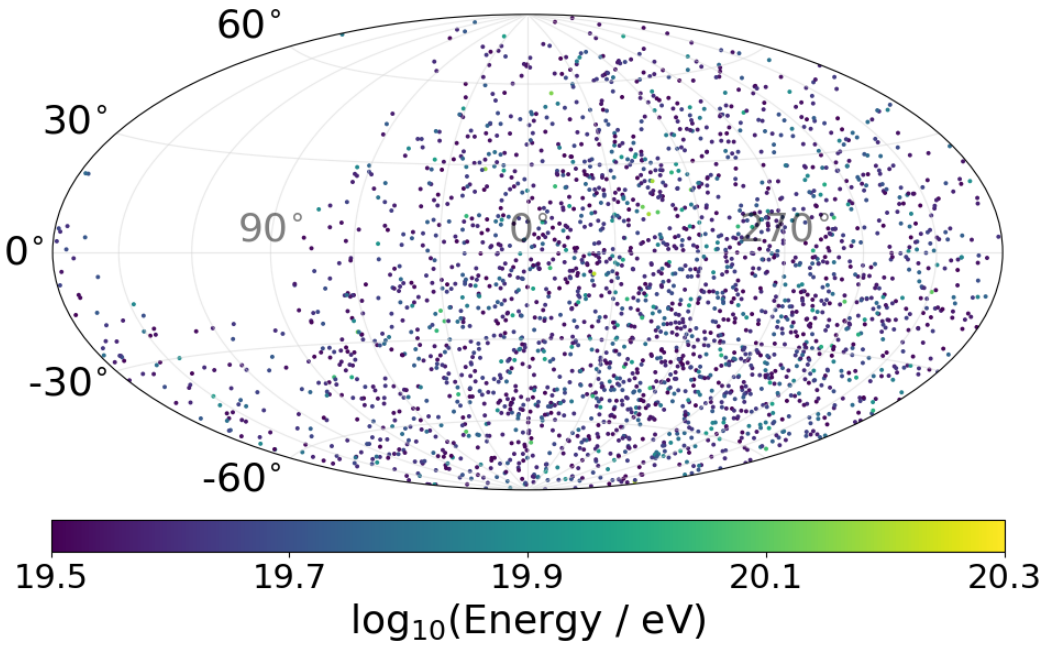
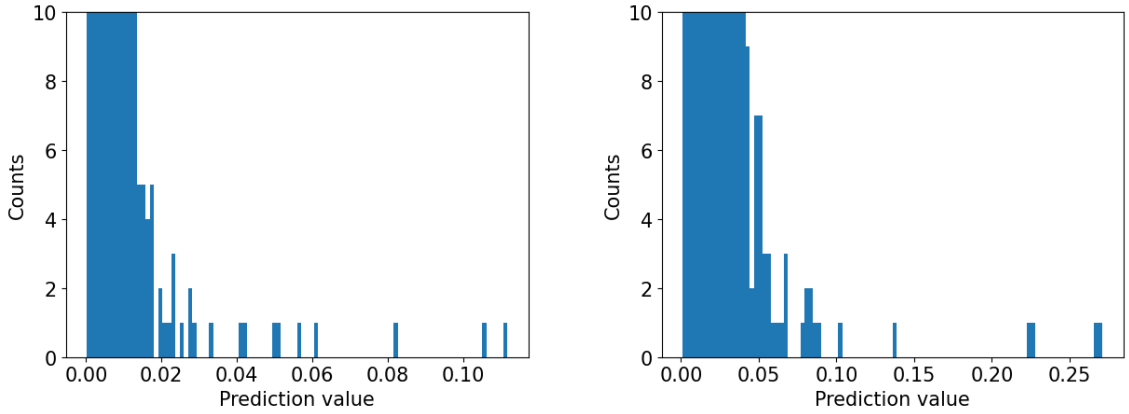


Figure 6.13.: Arrival directions of 2635 UHECRs above 32 EeV observed by Auger. Taken from [105].

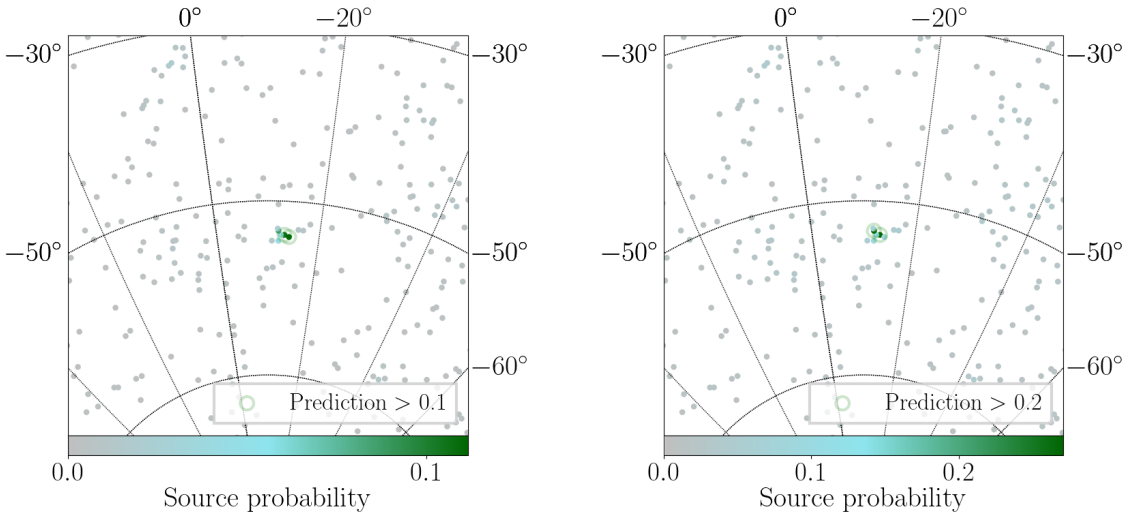
When applying the different neural networks to the Auger data an optimized threshold can obviously not be found anymore. Due to the less reliable performance of our neural network for heavier nuclei and few test particles, we rather focus on those UHECRs that receive the highest prediction values. Shown in Figure 6.14 are the output nodes of the proton and Nitrogen trained neural networks applied to the Auger data. Both distributions show a peak close to zero, indicating a high level of background. Only a few single UHECRs receive a noticeably higher source probability and could possibly be attributed to a source, however, their values remain questionable. For all trained models we find

the highest source probabilities for two small clusters of UHECRs shown in Figure 6.16. The pure-proton NN shows a mean prediction value of 0.004 for the entire Auger data set and gives 2 UHECRs a value above 0.1 while the Nitrogen model outputs a mean prediction value of 0.015 with 4 UHECRs above a threshold of 0.2.



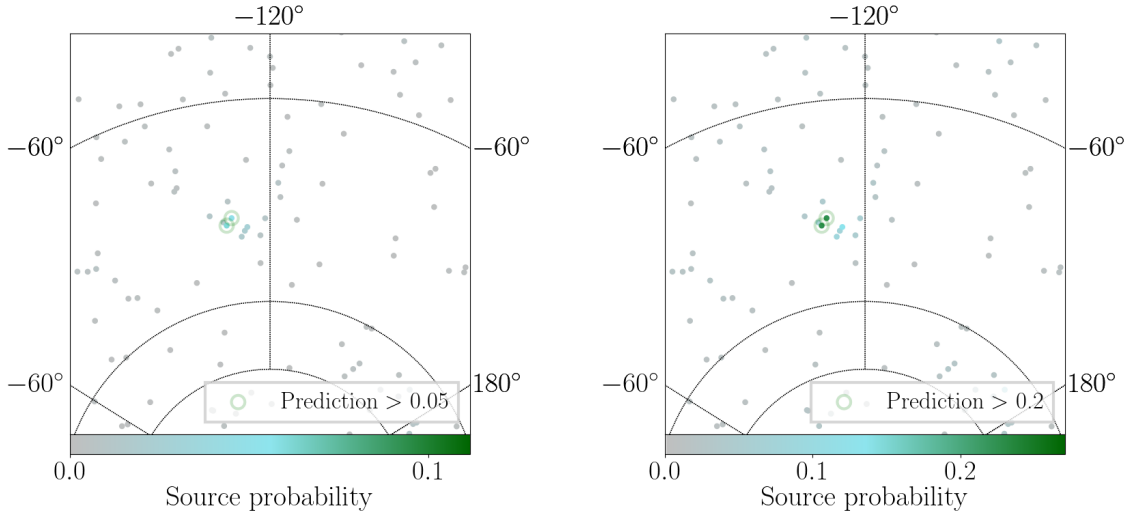
(a) Proton neural network **(b)** Nitrogen neural network

Figure 6.14.: Ouput node of the (a) proton and (b) Nitrogen trained neural network applied to the Auger data.



(a) Proton neural network **(b)** Nitrogen neural network

Figure 6.15.: Sky segment in Galactic coordinates of the 2635 UHECRs observed by Auger. Shown is the prediction value of a NN trained on a pure Nitrogen composition indicated by the colour. Taken from [105].



(a) Proton neural network

(b) Nitrogen neural network

Figure 6.16.: Sky segment in Galactic coordinates of the 2635 UHECRs observed by Auger. Shown is the prediction value of a NN trained on a pure Nitrogen composition indicated by the colour. Auger data taken from reference [105].

We note that although the source probability of neural networks is disputable, these two found regions could be of interest for further analysis. We could lower the cut on the threshold even further but arguably their prediction would be even less trustworthy.

6.5. Comparison to other methods

Finally, we still want to shortly place this work within similar conducted research. In the already mentioned studies [91] searches for multiplets of proton and Helium nuclei were performed. By deriving the correlation coefficients and thrust ratios within a preselected search region for certain source candidates they studied the number of events needed to surpass a 3σ significance. Our network performs quite well with protons and we expect it to show similar behaviour still for Helium, which is motivated by Figure 6.9. We gain comparable values for the number of proton source events needed to assign the majority of particles to the source class correctly. However, for a better comparison, our analysis would have to be adjusted so that the neural network would be trained for only singular sources emitting the same nuclei fractions for stepwise decreasing particle counts N_s and averaged over several repetitions. We however do not rule out that a closer combination than presented in Section 5.2 of both methods could show beneficial results. Also, when they applied their method to Auger data they did not find any significant enough correlations with any candidate sources which is consistent with our results.

Another method using deep neural networks to study multiplets coming from a few source candidates was performed in [107]. Here, however, convolutional deep neural networks (CNNs) were used to reject the null hypothesis as a function of the fraction of source events N_s, meaning deriving the probability that the arrival-direction distribution can not result from an isotropic pattern for a certain fraction of source events. Herefore, the arrival directions were not viewed on an event-to-event basis, but the entire distribution over the observable sky as an image. Thereby they can not determine if some selected UHECRs come from any given source or not.

We kept our machine learning approach very general by adding as few assumptions as possible e.g. training on several random source positions and no charge information. As we were quite pessimistic about the information we included, our results show to be limited towards more realistic scenarios. In [74] the effect of the random turbulent magnetic field component on iron nuclei above 60 EeV was extensively studied. It was shown that the random turbulent field could cause a large distortion effect on the arrival direction of an iron UHECR for just a relatively small change in its trajectory. As such a deflection in the GMF model is random by design, detecting a pattern is extremely difficult as the acquirable information on differences between isotropic background and source events might just be too insignificant. However, with increased data size we have shown that a deep-learning approach could pave the way to possibly detect such patterns in the future. Furthermore, with more limitations and better modelling of the GMF, NNs could be trained more specifically and better features constructed. Although we used a fully connected linear neural network, we suggest that a graph neural network (GNN) [108] might be better suited for the detection of heavier nuclei.

CHAPTER 7

CONCLUSION AND OUTLOOK

The origin of UHECRs is one of the biggest ongoing mysteries in astroparticle physics. Since charged particles are deflected in the Galactic magnetic field (GMF), searching for magnetically-induced patterns in their arrival directions might help to finally discover their sources.

Within this thesis, deep neural networks were used to find magnetically-induced patterns of UHECRs arriving from a few sources and distinguish them from a larger isotropic background on an event-to-event basis. For this, simulations of the propagation of UHECRs through the Galactic magnetic field were created to reproduce an arrival-direction distribution as could be observed by the Pierre Auger Observatory. The neural networks were trained and tested for different mass composition scenarios, on both randomly distributed and a few selected candidate sources. Different plausible feature extensions were tested to stress certain attributes in the data and improve the performance of the network. It was shown, that by training and validating on different variations of the random and striated magnetic field components, the network could be made more robust towards uncertainties in these characteristic properties. We found that for protons, the neural network is able to detect source events in an all-sky search with a high level of accuracy. To predict more than 80% of the source events correctly, the network needed 10 events from M87, 15–20 events from Cen A and Fornax A, but required a significantly larger multiplicity for Cyg A due to its location in the Galactic plane of also lower geometrical exposure. For pure Nitrogen simulations, the network still maintained a considerable detection ability when selecting the UHECRs with the highest prediction values, yet, it showed to be unsuccessful in an iron-dominated mixed composition scenario. From an input variable analysis, we found that the angular distance feature extension was the most influential and the energy correlation became increasingly more important for the nitrogen deflections. However, we also discovered that the placement of the sources for the simulation of training data is crucial and should be implemented in the form of a grid in future works. When applying the differently trained networks to data taken by Auger

we found two regions of interest that could possibly be attributed to a source origin. Nonetheless, the prediction output of the neural network suggests that the clear majority arises from an isotropic background within our analysis. As our performance for heavier nuclei was insufficient on simulated data, this prediction remains however questionable. We concluded that the currently available data and the knowledge of the Galactic magnetic field are not enough to accurately detect heavier-nuclei-dominated source events from an isotropic background using deep neural networks with such a general approach as we have presented.

However, extending the simulations to also consider the observations done by the Telescope Array [109] could show a crucial improvement. Additionally, by individually analysing different candidate sources and their power-law index, their energy spectrum can be adjusted accordingly. When trained on this property, the neural network could capture such an underlying structure in the observational data. Also, with larger computational resources we suggest that a graph neural network approach, where each pixel of the observed sky represents a node in the input layer, might be well suited for this problem [108]. With the success that convolutional neural networks have shown in recent years and the increasingly available computational resources, this method could show more promising results. Furthermore, with more limitations and knowledge on the deflection of the Galactic magnetic field, more accurate features could be constructed to better capture correlations between more largely deflected events from a common source.

BIBLIOGRAPHY

[1] V. F. Hess, "The Discovery of Cosmic Radiation", *Thought: Fordham University Quarterly* **15** (1940), no. 2, 225–236, doi:10.5840/thought1940152136.

[2] R. A. Batista, A. Dundovic, M. Erdmann et al., "CRPropa 3—a public astrophysical simulation framework for propagating extraterrestrial ultra-high energy particles", *Journal of Cosmology and Astroparticle Physics* **2016** (may, 2016) 038–038, doi:10.1088/1475-7516/2016/05/038.

[3] R. Jansson and G. R. Farrar, "The Galactic Magnetic Field", *The Astrophysical Journal Letters* **761** (nov, 2012) L11, doi:10.1088/2041-8205/761/1/L11.

[4] P. Collaboration, O. Adriani, G. Barbarino et al., "Ten years of PAMELA in space", *La Rivista del Nuovo Cimento* **40** (2017) 473–522, doi:10.1393/ncr/i2017-10140-x.

[5] T. K. Gaisser, R. Engel, and E. Resconi, "Cosmic Rays and Particle Physics". Cambridge University Press, 2 edition, 2016.

[6] T. P. A. collaboration, "Reconstruction of inclined air showers detected with the Pierre Auger Observatory", *Journal of Cosmology and Astroparticle Physics* **2014** (aug, 2014) 019–019, doi:10.1088/1475-7516/2014/08/019.

[7] Particle Data Group Collaboration, "Review of Particle Physics", *Phys. Rev. D* **98** (Aug, 2018) 030001, doi:10.1103/PhysRevD.98.030001.

[8] J. R. Hörandel, "Models of the knee in the energy spectrum of cosmic rays", *Astroparticle Physics* **21** (jun, 2004) 241–265, doi:10.1016/j.astropartphys.2004.01.004.

[9] T. Abu-Zayyad, D. Ivanov, C. C. H. Jui et al., "The Knee and the Second Knee of the Cosmic-Ray Energy Spectrum", 2018. doi:https://doi.org/10.48550/arXiv.1803.07052.

[10] The Pierre Auger Collaboration Collaboration, "Measurement of the cosmic-ray energy spectrum above 2.5×10^{18} eV using the Pierre Auger Observatory", *Phys. Rev. D* **102** (Sep, 2020) 062005, doi:10.1103/PhysRevD.102.062005.

[11] J. Heinze, D. Boncioli, M. Bustamante et al., "COSMOGENIC NEUTRINOS CHALLENGE THE COSMIC-RAY PROTON DIP MODEL", *The Astrophysical Journal* **825** (jul, 2016) 122, doi:10.3847/0004-637x/825/2/122.

[12] J. Glombitza, "Deep-learning based measurement of the mass composition of ultra-high energy cosmic rays using the surface detector of the Pierre Auger Observatory". PhD thesis, 2021. doi:10.18154/RWTH-2022-00759.

[13] M. Unger, "Highlights from the Pierre Auger Observatory (ICRC17)", doi:10.48550/ARXIV.1710.09478.

[14] B. Peters, "Primary cosmic radiation and extensive air showers", *Il Nuovo Cimento* **22** (November, 1961) 800–819, doi:10.1007/BF02783106.

[15] P. A. M. A. J. A. I. A. R. de Almeida, The Pierre Auger Collaboration, "Large-scale and multipolar anisotropies of cosmic rays detected at the Pierre Auger Observatory with energies above 4 EeV", *Proceedings of Science* **395** (October, 2021) doi:https://doi.org/10.22323/1.395.0335.

[16] A. C. Crook, J. P. Huchra, N. Martimbeau et al., "Groups of Galaxies in the Two Micron All Sky Redshift Survey", *The Astrophysical Journal* **655** (2007), no. 2, 790, doi:10.1086/510201.

[17] and A. Aab, P. Abreu, M. Aglietta et al., "Observation of a large-scale anisotropy in the arrival directions of cosmic rays above 8 × 10 sup18/sup eV", *Science* **357** (sep, 2017) 1266–1270, doi:10.1126/science.aan4338.

[18] A. H. Compton and I. A. Getting, "An Apparent Effect of Galactic Rotation on the Intensity of Cosmic Rays", *Phys. Rev.* **47** (Jun, 1935) 817–821, doi:10.1103/PhysRev.47.817.

[19] D. Harari, S. Mollerach, and E. Roulet, "Anisotropies of ultrahigh energy cosmic rays diffusing from extragalactic sources", *Phys. Rev. D* **89** (Jun, 2014) 123001, doi:10.1103/PhysRevD.89.123001.

[20] D. Harari, S. Mollerach, and E. Roulet, "Anisotropies of ultrahigh energy cosmic ray nuclei diffusing from extragalactic sources", *Physical Review D* **92** (2015), no. 6, 063014.

[21] V. Berezinskii, S. Grigor'eva, and V. Dogiel, "Predicted spectrum and anisotropy of the ultra-high energy cosmic rays in a single-source model", *Astronomy and Astrophysics* **232** (1990) 582–588.

[22] A. Aab, P. Abreu, M. Aglietta et al., "An Indication of Anisotropy in Arrival Directions of Ultra-high-energy Cosmic Rays through Comparison to the Flux Pattern of Extragalactic Gamma-Ray Sources", *The Astrophysical Journal* **853** (feb, 2018) L29, doi:10.3847/2041-8213/aaa66d.

[23] R. de Almeida for the Pierre Auger Collaboration., "Large-scale and multipolar anisotropies of cosmic rays detected at the Pierre Auger Observatory with energies above 4 EeV", *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021).* **358** (2019) 408, doi:10.22323/1.358.0408.

[24] D. Harari, S. Mollerach, and E. Roulet, "Anisotropies of ultrahigh energy cosmic ray nuclei diffusing from extragalactic sources", *Phys. Rev. D* **92** (Sep, 2015) 063014, doi:10.1103/PhysRevD.92.063014.

[25] A. Aab, P. Abreu, M. Aglietta et al., "Cosmic-Ray Anisotropies in Right Ascension Measured by the Pierre Auger Observatory ", *The Astrophysical Journal* **891** (March, 2020) 142, doi:10.3847/1538-4357/ab7236.

[26] P. Abreu, M. Aglietta, J. M. Albury et al., "Arrival Directions of Cosmic Rays above 32 EeV from Phase One of the Pierre Auger Observatory", *The Astrophysical Journal* **935** (aug, 2022) 170, doi:10.3847/1538-4357/ac7d4e.

[27] M. Kachelrieß, S. Ostapchenko, and R. Tomà s, "High energy radiation from Centaurus A", *New Journal of Physics* **11** (jun, 2009) 065017, doi:10.1088/1367-2630/11/6/065017.

[28] J. Biteau, R. Alves Batista, F. Canfora et al., "The ultra-high-energy cosmic-ray sky above 32 EeV viewed from the Pierre Auger Observatory",.

[29] A. Aab, P. Abreu, M. Aglietta et al. *The Astrophysical Journal* **804** (apr, 2015) 15, doi:10.1088/0004-637X/804/1/15.

[30] R. Protheroe and P. Biermann, "A new estimate of the extragalactic radio background and implications for ultra-high-energy γ-ray propagation", *Astroparticle Physics* **6** (dec, 1996) 45–54, doi:10.1016/s0927-6505(96)00041-2.

[31] A. Domínguez, J. R. Primack, D. J. Rosario et al., "Extragalactic background light inferred from AEGIS galaxy-SED-type fractions", **410** (February, 2011) 2556–2578, doi:10.1111/j.1365-2966.2010.17631.x, arXiv:1007.1459.

[32] D. S. Walz, "Constraining Models of the Extragalactic Cosmic-Ray Origin with the Pierre Auger Observatory". PhD thesis, 2016.

[33] A. A. Penzias and R. W. Wilson, "A Measurement of Excess Antenna Temperature at 4080 Mc/s.", **142** (July, 1965) 419–421, doi:10.1086/148307.

[34] D. J. Fixsen, "The Temperature of the Cosmic Microwave Background", *The Astrophysical Journal* **707** (nov, 2009) 916, doi:10.1088/0004-637X/707/2/916.

[35] K. Greisen, "End to the Cosmic-Ray Spectrum?", *Phys. Rev. Lett.* **16** (Apr, 1966) 748–750, doi:10.1103/PhysRevLett.16.748.

[36] A. Cooray, "Extragalactic Background Light: Measurements and Applications", doi:https://doi.org/10.48550/arXiv.1602.03512, arXiv:1602.03512.

[37] R. Alves Batista et al., "Open Questions in Cosmic-Ray Research at Ultrahigh Energies", *Front. Astron. Space Sci.* **6** (2019) 23, doi:10.3389/fspas.2019.00023, arXiv:1903.06714.

[38] D. Kümpel, "Multivariate Search for a Directional Excess of EeV Photons with the Pierre Auger Observatory". PhD thesis, 2011.

[39] J. Abraham, P. Abreu, M. Aglietta et al., "Upper limit on the cosmic-ray photon flux above 1019eV using the surface detector of the Pierre Auger Observatory", *Astroparticle Physics* **29** (may, 2008) 243–256, doi:10.1016/j.astropartphys.2008.01.003.

[40] K. Fang, K. Kotera, and A. V. Olinto, "Newly-born pulsars as sources of ultrahigh energy cosmic rays", *The Astrophysical Journal* **750** (apr, 2012) 118, doi:10.1088/0004-637X/750/2/118.

[41] E. Fermi, "On the Origin of the Cosmic Radiation", *Phys. Rev.* **75** (Apr, 1949) 1169–1174, doi:10.1103/PhysRev.75.1169.

[42] J. G. Kirk and P. Duffy, "Particle acceleration and relativistic shocks", *Journal of Physics G: Nuclear and Particle Physics* **25** (jul, 1999) R163–R194, doi:10.1088/0954-3899/25/8/201.

[43] A. R. Bell, "The acceleration of cosmic rays in shock fronts - I.", **182** (January, 1978) 147–156, doi:10.1093/mnras/182.2.147.

[44] C. D. Dermer, "Maximum Particle Energies by Fermi Acceleration and the Origin of Cosmic Rays above the Knee", 2001. doi:https://doi.org/10.48550/arXiv.astro-ph/0012490.

[45] A. Heavens and L. Drury, "Relativistic shocks and particle acceleration", *Monthly Notices of the Royal Astronomical Society* **235** (1988) 997–1009.

[46] A. Achterberg, Y. A. Gallant, J. G. Kirk et al., "Particle acceleration by ultrarelativistic shocks: theory and simulations", *Monthly Notices of the Royal Astronomical Society* **328** (2001), no. 2, 393–408.

[47] A. M. Hillas, "The Origin of Ultra-High-Energy Cosmic Rays", **22** (January, 1984) 425–444, doi:10.1146/annurev.aa.22.090184.002233.

[48] K. Kotera and A. V. Olinto, "The Astrophysics of Ultrahigh-Energy Cosmic Rays",

Annual Review of Astronomy and Astrophysics **49** (sep, 2011) 119–153, doi:10.1146/annurev-astro-081710-102620.

[49] J. H. Matthews, A. R. Bell, K. M. Blundell et al., "Fornax A, Centaurus A, and other radio galaxies as sources of ultrahigh energy cosmic rays", *Monthly Notices of the Royal Astronomical Society: Letters* **479** (jun, 2018) L76–L80, doi:10.1093/mnrasl/sly099.

[50] J. H. Matthews, A. R. Bell, K. M. Blundell et al., "Ultrahigh energy cosmic rays from shocks in the lobes of powerful radio galaxies", **482** (February, 2019) 4303–4321, doi:10.1093/mnras/sty2936, arXiv:1810.12350.

[51] S. S. Kimura, K. Murase, and B. T. Zhang, "Ultrahigh-energy cosmic-ray nuclei from black hole jets: Recycling galactic cosmic rays through shear acceleration", *Physical Review D* **97** (jan, 2018) doi:10.1103/physrevd.97.023026.

[52] B. Eichmann, J. Rachen, L. Merten et al., "Ultra-high-energy cosmic rays from radio galaxies", *Journal of Cosmology and Astroparticle Physics* **2018** (feb, 2018) 036–036, doi:10.1088/1475-7516/2018/02/036.

[53] P. Mészáros, D. B. Fox, C. Hanna et al., "Multi-messenger astrophysics", *Nature Reviews Physics* (2019) doi:10.1038/s42254-019-0101-z.

[54] T. Piran, "Gamma-ray bursts and the fireball model", *Physics Reports* **314** (jun, 1999) 575–667, doi:10.1016/s0370-1573(98)00127-6.

[55] P. Baerwald, M. Bustamante, and W. Winter, "Are gamma-ray bursts the sources of ultra-high energy cosmic rays?", *Astroparticle Physics* **62** (mar, 2015) 66–91, doi:10.1016/j.astropartphys.2014.07.007.

[56] V. M. Kaspi and A. M. Beloborodov, "Magnetars", *Annual Review of Astronomy and Astrophysics* **55** (aug, 2017) 261–301, doi:10.1146/annurev-astro-081915-023329.

[57] J. Pruet, S. Guiles, and G. Fuller, "Light Element Synthesis in High Entropy Relativistic Flows Associated with Gamma Ray Bursts", *The Astrophysical Journal* **580** (05, 2002) doi:10.1086/342838.

[58] B. A. Groves, T. M. Heckman, and G. Kauffmann, "Emission-line diagnostics of low-metallicity active galactic nuclei", *Monthly Notices of the Royal Astronomical Society* **371** (08, 2006) 1559–1569, doi:10.1111/j.1365-2966.2006.10812.x.

[59] K. Fang, K. Kotera, and A. V. Olinto, "Newly born pulsars as sources of ultrahigh energy cosmic rays", *The Astrophysical Journal* doi:10.1088/0004-637X/750/2/118.

[60] G. Golup, D. Harari, S. Mollerach et al., "Source position reconstruction and constraints on the galactic magnetic field from ultra-high energy cosmic rays", *Astroparticle Physics* **32** (dec, 2009) 269–277, doi:10.1016/j.astropartphys.2009.09.003.

[61] K. Beuermann, G. Kanbach, and E. M. Berkhuijsen, "Radio structure of the Galaxy: thick disk and thin disk at 408 MHz.", **153** (December, 1985) 17–34.

[62] Beck, R., "Magnetism in the spiral galaxy NGC 6946: magnetic arms, depolarization rings, dynamo modes, and helical fields", *A&A* **470** (2007), no. 2, 539–556, doi:10.1051/0004-6361:20066988.

[63] J. M. Stil, A. R. Taylor, and C. Sunstrum, "Sturcture in the rotation measure sky", *The Astrophysical Journal* **726** (dec, 2010) 4, doi:10.1088/0004-637x/726/1/4.

[64] A. R. Taylor, J. M. Stil, and C. Sunstrum, "A rotation measure image of the sky", *The Astrophysical Journal* **702** (aug, 2009) 1230, doi:10.1088/0004-637X/702/2/1230.

[65] A. W. Strong, E. Orlando, and T. R. Jaffe, "The interstellar cosmic-ray electron spectrum from synchrotron radiation and direct measurements", *Astronomy &amp Astrophysics* **534** (sep, 2011) A54, doi:10.1051/0004-6361/201116828.

[66] X. Sun, W. Reich, A. Waelkens et al., "Radio observational constraints on Galactic 3D-emission models", *Astronomy & Astrophysics* **477** (2008), no. 2, 573–592.

[67] M. S. Pshirkov, P. G. Tinyakov, P. P. Kronberg et al., "Deriving global structure of the Galactic Magnetic Field from Faraday Rotation Measures of extragalactic sources", *The Astrophysical Journal* **738** (aug, 2011) 192, doi:10.1088/0004-637X/738/2/192.

[68] G. R. Farrar, "The Galactic magnetic field and ultrahigh-energy cosmic ray deflections", *Comptes Rendus Physique* **15** (apr, 2014) 339–348, doi:10.1016/j.crhy.2014.04.002.

[69] J. Kleimann, T. Schorlepp, L. Merten et al., "Solenoidal Improvements for the JF12 Galactic Magnetic Field Model", *The Astrophysical Journal* **877** (may, 2019) 76, doi:10.3847/1538-4357/ab1913.

[70] B. Gold, N. Odegard, J. L. Weiland et al., "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Galactic Foreground Emission", *The Astrophysical Journal Supplement Series* **192** (jan, 2011) 15, doi:10.1088/0067-0049/192/2/15.

[71] C. L. Van Eck, J. C. Brown, J. M. Stil et al., "Modeling the Magnetic Field in the Galactic Disk Using New Rotation Measure Observations from the Very Large

Array", **728** (February, 2011) 97, doi:10.1088/0004-637X/728/2/97, arXiv:1012.2938.

[72] T.Bister, "A Combined Fit of Energy Spectrum, Shower Depth Distributions and Arrival Directions of Ultra-High-Energy Cosmic Rays measured at the Pierre Auger Observatory". PhD thesis, 2021.

[73] R. Beck, "Galactic dynamos and galactic winds", *Astrophysics and Space Science* **320** (2009), no. 1, 77–84, doi:10.1007/s10509-008-9825-4.

[74] G. Giacinti, M. Kachelrieß, D. Semikoz et al., "Ultrahigh energy nuclei in the turbulent Galactic magnetic field", *Astroparticle Physics* **35** (nov, 2011) 192–200, doi:10.1016/j.astropartphys.2011.07.006.

[75] and R. Adam, P. A. R. Ade, M. I. R. Alves et al., "iPlanck/iintermediate results", *Astronomy &amp Astrophysics* **596** (dec, 2016) A103, doi:10.1051/0004-6361/201528033.

[76] A. Keivani, G. R. Farrar, and M. Sutherland, "Magnetic deflections of ultra-high energy cosmic rays from Centaurus A", *Astroparticle Physics* **61** (feb, 2015) 47–55, doi:10.1016/j.astropartphys.2014.07.001.

[77] T. Sjöstrand, S. Ask, J. R. Christiansen et al., "An introduction to PYTHIA 8.2", *Computer Physics Communications* **191** (2015) 159–177, doi:https://doi.org/10.1016/j.cpc.2015.01.024.

[78] G. Aad, T. Abajyan, B. Abbott et al., "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Physics Letters B* **716** (2012), no. 1, 1–29.

[79] R. A. Batista, J. B. Tjus, and J. Dörner, "CRPropa3". https://crpropa.github.io/CRPropa3/, 2022. accessed:15/04/2023.

[80] J. R. Cash and A. H. Karp, "A Variable Order Runge-Kutta Method for Initial Value Problems with Rapidly Varying Right-Hand Sides", *ACM Trans. Math. Softw.* **16** (sep, 1990) 201–222, doi:10.1145/79505.79507.

[81] H.-P. Bretz, M. Erdmann, P. Schiffer et al., "PARSEC: A parametrized simulation engine for ultra-high energy cosmic ray protons", *Astroparticle Physics* **54** (feb, 2014) 110–117, doi:10.1016/j.astropartphys.2013.12.002.

[82] A. Zonca, L. Singer, D. Lenz et al., "healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python", *Journal of Open Source Software* **4** (March, 2019) 1298, doi:10.21105/joss.01298.

[83] K. M. Górski, E. Hivon, A. J. Banday et al., "HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere", **622** (April, 2005) 759–771, doi:10.1086/427976.

[84] P. Sommers, "Cosmic ray anisotropy analysis with a full-sky observatory", *Astroparticle Physics* **14** (jan, 2001) 271–286, doi:10.1016/s0927-6505(00)00130-4.

[85] D. Ehlert, F. Oikonomou, and M. Unger, "The Curious Case of Near-Identical Cosmic-Ray Accelerators", 2022. doi:https://doi.org/10.48550/arXiv.2207.10691.

[86] Statista, "Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025". https://www.statista.com/statistics/871513/worldwide-data-created/, june, 2021. accessed:15/04/2023.

[87] The CERN Data Centre , "Key Facts and Figures – CERN Data Centre", nov, 2021.

[88] A. F. Agarap, "Deep learning using rectified linear units (relu)", *arXiv preprint arXiv:1803.08375* (2018).

[89] Torch Contributors, "BinaryCrossEntropyLoss", 2023. accessed:15/04/2023.

[90] J. Sola and J. Sevilla, "Importance of input data normalization for the application of neural networks to complex industrial problems", *IEEE Transactions on Nuclear Science* **44** (1997), no. 3, 1464–1468, doi:10.1109/23.589532.

[91] A. Aab, P. Abreu, M. Aglietta et al., "Search for magnetically-induced signatures in the arrival directions of ultra-high-energy cosmic rays measured at the Pierre Auger Observatory", *Journal of Cosmology and Astroparticle Physics* **2020** (jun, 2020) 017–017, doi:10.1088/1475-7516/2020/06/017.

[92] W. Jin, X. Liu, Y. Ma et al., "Feature Overcorrelation in Deep Graph Neural Networks", in *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. ACM, aug, 2022. doi:10.1145/3534678.3539445.

[93] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift", 2015. doi:https://doi.org/10.48550/arXiv.1502.03167.

[94] G. Klambauer, T. Unterthiner, A. Mayr et al., "Self-Normalizing Neural Networks", 2017. doi:https://doi.org/10.48550/arXiv.1706.02515.

[95] N. Srivastava, G. Hinton, A. Krizhevsky et al., "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", *Journal of Machine Learning Research* **15** (2014), no. 56, 1929–1958.

[96] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization", *arXiv preprint arXiv:1711.05101* (2017) doi:https://doi.org/10.48550/arXiv.1711.05101.

[97] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms", *Pattern Recognition* **30** (1997), no. 7, 1145–1159, doi:https://doi.org/10.1016/S0031-3203(96)00142-2.

[98] C. Goutte and E. Gaussier, "A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation", in *Advances in Information Retrieval*, D. E. Losada and J. M. Fernández-Luna, eds., pp. 345–359. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[99] A. Paszke, S. Gross, F. Massa et al., "PyTorch: An Imperative Style, High-Performance Deep Learning Library", 2019.

[100] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization", 2017. doi:https://doi.org/10.48550/arXiv.1412.6980.

[101] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System", in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pp. 785–794. ACM, New York, NY, USA, 2016. doi:10.1145/2939672.2939785.

[102] G. Ke, Q. Meng, T. Finley et al., "Lightgbm: A highly efficient gradient boosting decision tree", *Advances in neural information processing systems* **30** (2017) 3146–3154.

[103] F. Pedregosa, G. Varoquaux, A. Gramfort et al., "Scikit-learn: Machine Learning in Python ", *Journal of Machine Learning Research* **12** (2011) 2825–2830.

[104] T. Akiba, S. Sano, T. Yanase et al., "Optuna: A Next-generation Hyperparameter Optimization Framework", in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2019. doi:https://doi.org/10.48550/arXiv.1907.10902.

[105] P. A. Collaboration, "Material for the analyses presented in "Arrival Directions of Cosmic Rays above 32 EeV from Phase One of the Pierre Auger Observatory"", June, 2022. doi:10.5281/zenodo.6759610.

[106] S. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions", 2017. doi:https://doi.org/10.48550/arXiv.1705.07874.

[107] O. Kalashev, M. Pshirkov, and M. Zotov, "Identifying nearby sources of ultra-high-energy cosmic rays with deep learning", *Journal of Cosmology and Astroparticle Physics* **2020** (nov, 2020) 005–005, doi:10.1088/1475-7516/2020/11/005.

[108] J. Zhou, G. Cui, S. Hu et al., "Graph neural networks: A review of methods and applications", *AI Open* **1** (2020) 57–81, doi:https://doi.org/10.1016/j.aiopen.2021.01.001.

[109] T. Abu-Zayyad, R. Aida, M. Allen et al., "The surface detector array of the Telescope Array experiment", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **689** (2012) 87–97, doi:https://doi.org/10.1016/j.nima.2012.05.079.

[110] R. A. Batista, J. B. Tjus, J. Dörner et al., "CRPropa 3.2 — an advanced framework for high-energy particle propagation in extragalactic and galactic spaces", *Journal of Cosmology and Astroparticle Physics* **2022** (sep, 2022) 035, doi:10.1088/1475-7516/2022/09/035.

[111] T. M. Kneiske, T. Bretz, K. Mannheim et al., "Implications of cosmological gamma-ray absorption", *Astronomy &amp Astrophysics* **413** (jan, 2004) 807–815, doi:10.1051/0004-6361:20031542.

APPENDIX A

EXTRAGALACTIC 1D SIMULATION

A.1. Extragalactic 1D simulation

Although this work studies the Galactic deflection of UHECRs, 1D extragalactic simulations were created to see how the energy spectrum changes with distance. Herefore, the interaction processes previously introduced in Section 3.1 are already implemented within the framework of CRPropa and were used as such. For more information on the different available modules see the documentation page [110]. First, protons originating from possible candidate sources Centaurus A (Cen A), Cygnus A (Cyg A), Messier 87 (M87) and Fornax A were propagated through 1D extragalactic space. The simulated interaction processes were based on the EBL model of reference [111]. The protons were injected with a power-law spectrum of $\gamma = -1$ at the corresponding distance of each source. The resulting spectrum at Earth is shown in Figure A.1. It can be seen that Cyg A experiences a lot softer cutoff at roughly $10^{20.15}$ eV than the other sources due to its far further distance of 232 Mpc. A clear pileup of post-interacted protons can be seen at energies of $10^{19.5} - 10^{19.9}$ eV and a smaller one for M87 and Fornax A at approximately 10^{20} eV.

These derived energy spectra could now be implemented into the simulations. However, it has to be ensured that the total energy spectra of both isotropic background and that from the source candidates in total follow the one observed at Auger. With this, the neural network would be able to better distinguish the two classes as the energy information will not anymore be the same. The protons from Cen A on the other hand arrive almost at the same power-law of $\gamma = -1$ as they were injected. Cyg falls off with an index $\gamma = -7.1$ from $10^{19.75}$ eV and both Fornax A and M87 with roughly $\gamma = -1.4$ from 10^{20} eV. In Figure A.2 the same simulation was run for Cyg A but for also nitrogen and

iron. Although the curve for nitrogen and iron can be barely distinguished as they are almost identical, the cutoff for those two heavier elements occurs at even lower energies.

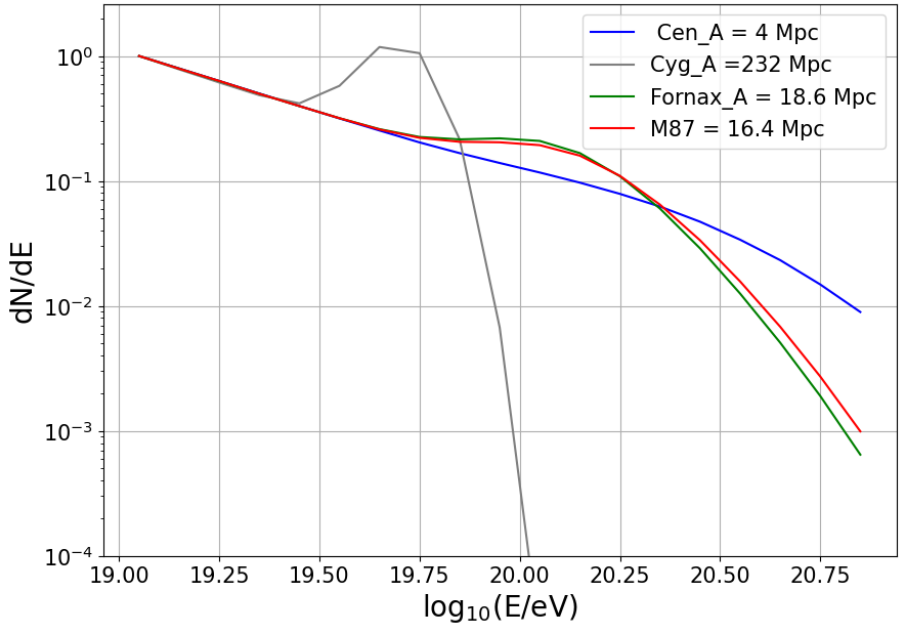


Figure A.1.: Extragalacic 1D propagation of UHECR protons from selected candidate sources as a function of distance only. The spectrum of Cyg A experiences a lot stronger cut-off and a far steeper power law. Cen A stays almost at the injection power-law index.

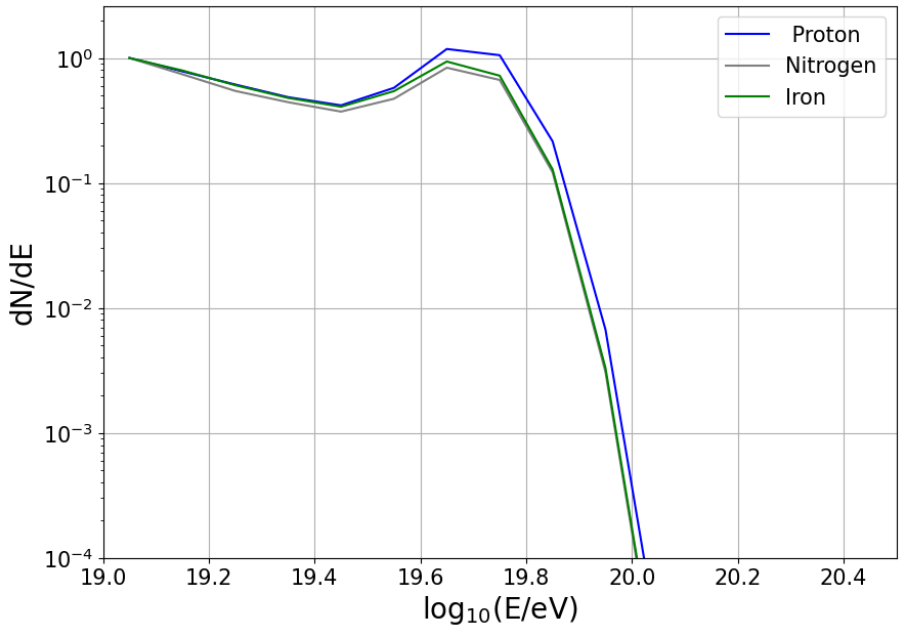
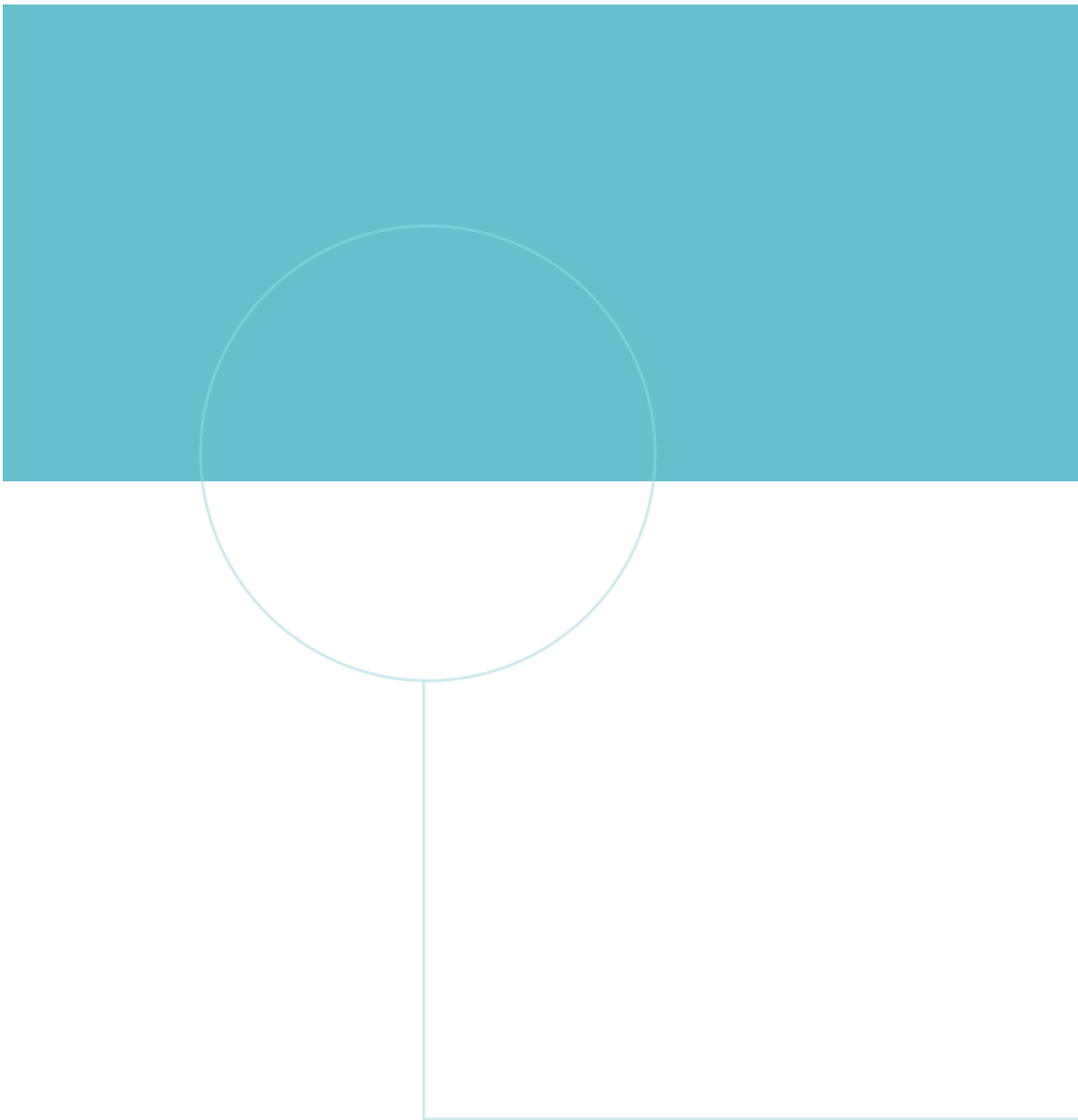


Figure A.2.: Comparison of protons (blue), nitrogen (grey) and iron (green) originating from Cyg A propagated extragalactic 1D space. Nitrogen and iron are almost indistinguishable and experience an earlier cut-off.

Acknowledgements

First off, I would like to thank Prof. Dr. Foteini Oikonomou for the opportunity to write such an exciting Master's thesis within the NTNU astroparticle physics group. Her supervision and help even outside the scope of this work, went far beyond what could have been expected. The time and care she invests in her students both on an academic and a personal level is truly unique. I would also like to thank Domenik Ehlert for helping me whenever it was needed and Christopher Matthies whose critical comments on my thesis helped make it what it is now. Also, a thank you goes out to my fellow students in C4 who not only always had an open ear but also made especially the final phase of the thesis a lot more enjoyable through our collective hardship and many jokes. I special thanks go out to Per Arne Selve, for the intense cooperation over the entire duration and long discussions of which many might have derailed into pointlessness.

Next, I would like to thank Dana King for always being there when I was frustrated about my model's performance and lifting my spirits when needed. Finally, I would like to especially thank my family for their constant support over the last two years, without them neither this work nor this amazing abroad experience would have been possible. Thank you!



NTNU

Norwegian University of
Science and Technology