Friedrich-Alexander-Universität
Faculty of Sciences

Bachelor's thesis

Observing the Moon Shadow in SWGO Cosmic Ray Simulations

Lena Jäger

supervisor: Prof. Dr. Christopher van Eldik

submission date: 6th of February, 2026

Abstract

The moon shadow in SWGO cosmic ray simulation data is observed. Using a custom simulation based on the instrument response found in established simulations, we examine the dependency of the shadow visibility for different observation durations, in different energy bins and with and without consideration of the deflection of primaries in the geomagnetic field. We successfully determine the value of the angular resolution from shadow observations in accordance with the resolution input into the simulations. Furthermore, we find our results of the shadow deflection to be in accordance with the findings of HAWC.

Zusammenfassung

Es wird der Mondschatten der kosmischen Strahlung in SWGO Simulationen beobachtet. Unter Verwendung einer selbsterstellten Simulation, die auf der durch etablierte Simulationen gefundenen Detektorfunktion basiert, betrachten wir die Abhängigkeit der Sichtbarkeit des Schattens bei variabler Observationszeit und in verschiedenen Energiebereichen. Zudem untersuchen wir den Einfluss der Ablenkung der Primärteilchen im Erdmagnetfeld. Uns gelingt eine erfolgreiche Bestimmung der Winkelauflösung aus Schattenbeobachtungen, welche mit der zuvor in die Simulationen eingegebenen Winkelauflösung übereinstimmt. Des Weiteren decken sich unsere Ergebnisse der Verschiebung des Mondschattens mit dem Befund von HAWC.

Contents

1 Motivation & Introduction

The Southern Wide-field Gamma-ray Observatory (SWGO) is set to be a Water Cherenkov Observatory, which will be used primarily to detect γ-rays in the Southern Hemisphere. Since the detection of cosmic rays functions similarly to gamma rays due to their comparable shower behavior, we can also detect the nearly isotropic flux of cosmic rays with the same instrument and transfer detector characteristics from cosmic to gamma rays.
Since primary particle flux is blocked by the moon, a shadow can be observed in cosmic ray data. In this thesis, we are aiming to observe this moon shadow using SWGO proton simulation data. In doing this, we strive to generate reasonable expectations regarding the appearance of the shadow and the time necessary to observe it. Once the detector is in use, this can be re-examined to ideally verify our findings.
One of the challenging aspects of this analysis comes from the interaction of cosmic rays with magnetic fields within our solar system. The main deflection of the trajectories of the cosmic rays reaching us from the direction of the moon is caused by the Earth's magnetic field. To accurately predict the size and position of the observable moon shadow, the geomagnetic field modeled by the IGRF is utilized to correct the apparent direction of protons by calculating the estimated tracks of the particles through the magnetic field up to moon height.
Since we know the true size and position of the moon precisely, the measured appearance of the shadow can also be used to determine the direction bias and resolution of the detector. Furthermore, having this kind of external calibration method is a useful tool for gamma ray observation with water-Cherenkov observatories. Whereas Imaging Atmospheric Cherenkov of Telescopes (IACTs) observe the whole shower in the atmosphere, water-Cherenkov observatories only see a horizontal slice of the shower. The shower fluctuations and limited available information make the reconstruction of such showers challenging. Thus having this kind of external calibration method is extremely useful to test the accuracy of the reconstruction algorithms.
Although not relevant for this thesis, compelling research can also be conducted in the field of cosmic ray physics, for which we refer to the SWGO White Paper [11, p. 36 ff.].
In section 2 we prepare the thesis by providing the necessary scientific knowledge. In section 3, we first evaluate the simulation data provided by CORSIKA and HAWCSim, before employing a custom simulation to create the data needed to observe the moon shadow, which is done in section 4. Finally, in section 5, we take the deflection of primaries in the geomagnetic field into account and simulate the appearance of the deflected moon shadow, quantifying our results and comparing them to the findings of the HAWC collaboration.

2 Theoretical Background & Preparations

2.1 Gamma Ray Astronomy

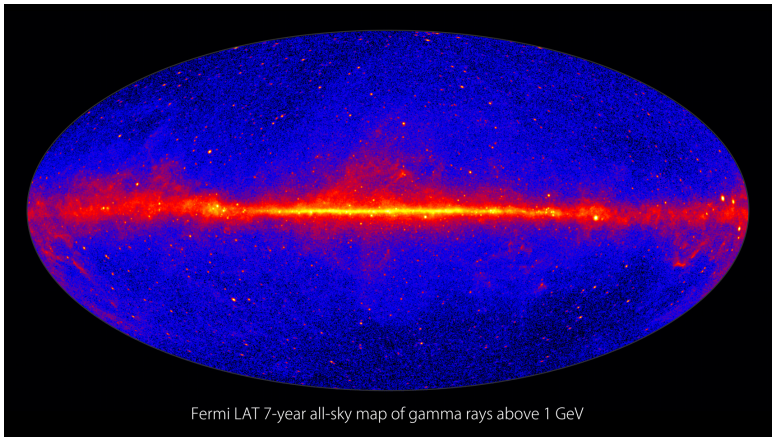

Figure 1: The all-sky map as seen in gamma rays above 1 GeV as observed with Fermi's Large Area Telescope (LAT). The graphic was taken from [10].

Gamma rays are photons with an energy of $E \geqslant 100\,\mathrm{keV}$, which are especially relevant in high energy astrophysics. Their distribution on a sky map can be seen in Figure 1. The main observed sources lie in the plane of the milky way, where emission consists of both gamma rays from discrete sources [2] and diffuse emission from π^0-decay in interstellar gas. Major Galactic sources are supernova remnants and pulsar wind nebulae, while extragalactic gamma rays mainly originate from active galactic nuclei and gamma ray bursts [27]. Their observation is especially useful, since photons are not charged and therefore not deflected by electromagnetic fields, which results in their arrival direction matching the direction of their origin. This makes it possible to locate and observe sources of high energy particles. [23, p. 25 ff.]
There are two main methods used to detect gamma rays, each entailing their own advantages and challenges: ground based and space based observation. When observing gamma rays via satellites, interaction with the atmosphere is avoided, but the measurable energy is limited, since the detector area becomes too small to measure with sufficient statistics for higher energies. At the ground, extended observatories with a large detection area can be built, but the direct detection is no longer possible as the atmosphere is opaque for gamma rays. Nonetheless, the atmosphere can be used as a calorimeter. The energy of primary particles is deposited into the atmosphere and generates a shower of secondary particles. These are then detected either by particle detector arrays, like HAWC, or IACTs, as is the case with H.E.S.S.. We will discuss the shower behavior in subsection 2.3. [23, p. 25 ff.]

2.2 Cosmic Ray Spectrum

When observing gamma rays at ground level, we also detect cosmic rays by default and need to differentiate between the two. Cosmic rays are atomic nuclei of different species that are emitted by astrophysical sources or produced in interactions with interstellar gas. Due to their charge, cosmic rays can be deflected and diffuse in the galactic magnetic

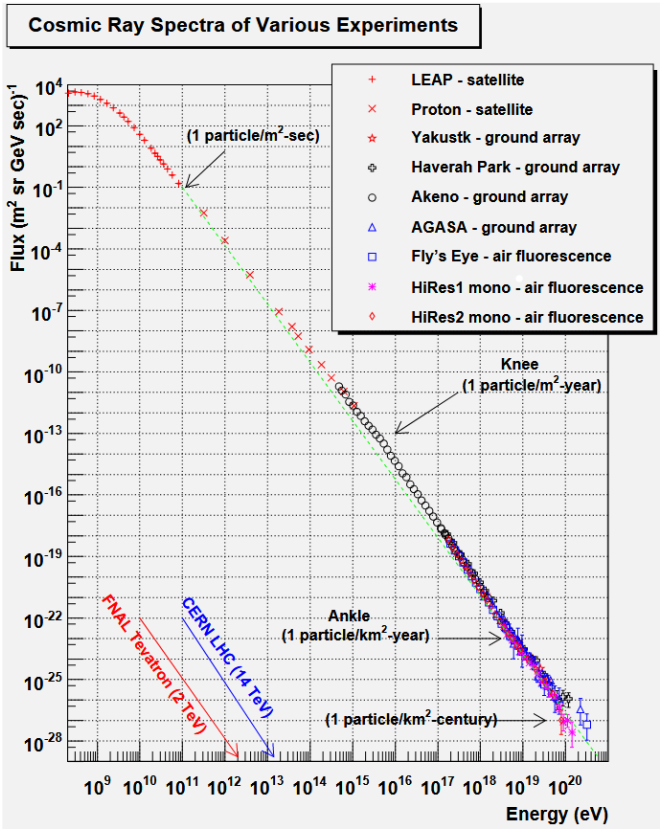

Figure 2: The flux of cosmic rays as observed by various experiments over eleven decades of energy. The first knee and ankle are shown, whereas the second knee is omitted. The graphic was taken from [17, p. 9].

field, making their flux on earth nearly isotropic [26], neglecting the anisotropy at the highest energies [3, p. 7, Fig. 3]. Their energy spans a wide range from about $1 \cdot 10^9$ eV to $1 \cdot 10^{20}$ eV, of which the latter are exceedingly rare [26]. The spectrum of cosmic rays, depicted by their flux Φ, can be well described by a power law of the following form:

$$\Phi(E)\,\mathrm{d}E = \Phi_0 \left(\frac{E}{E_0}\right)^{-\gamma} \mathrm{d}E \tag{1}$$

The formula follows [23, p. 30, Eq. 1.2]. In the range of about 100 GeV – 100 TeV, which will be relevant for the employed simulations, the value of γ is assumed to be constant at $\gamma = 2.7$ for the differential spectrum [23, p. 493 ff.]. At higher energies between 1 PeV – 10 PeV, steepening in the spectrum is observed, which is commonly referred to as the "knee" of the spectrum, followed by a second knee at 100 PeV and the "ankle" at ≈ 32 EeV where the exponent of the of the underlying power law seems to differ [26]. The described spectrum can be viewed in Figure 2. This change of the power law will not be relevant for our further discussions, since particles in that energy range cannot be detected by SWGO due to its limited detector area.
There are also differences in the observed spectrum for the multiple cosmic ray species that occur (s. Figure 3), of which hydrogen nuclei make up about 87 % [23, p. 493 ff.]. We only discuss proton flux in this thesis, because they are the only species simulated

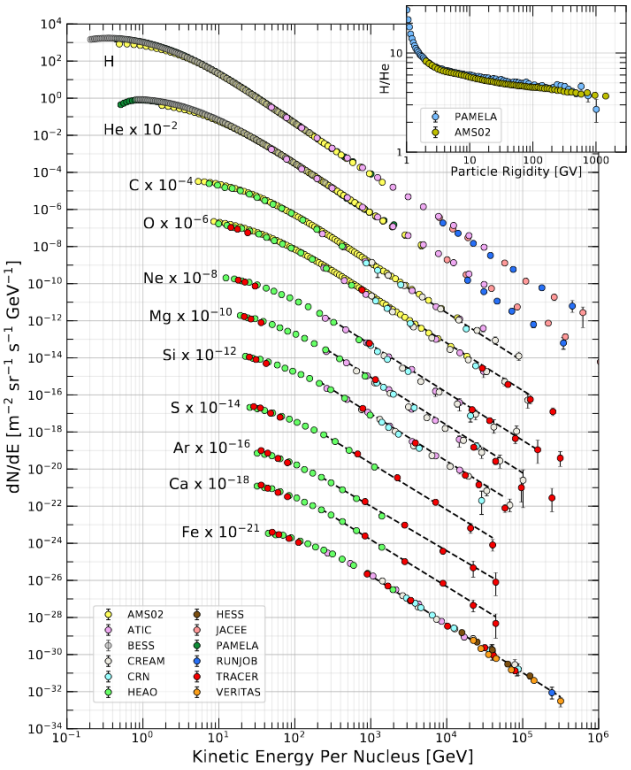

Figure 3: The flux of nuclei of cosmic rays divided into their respective species. The flux of hydrogen nuclei makes up by far the biggest part of cosmic ray flux. The graphic was taken from [26, p. 3].

in our evaluated data. Later, we will compute the expected rate of these protons on the detector area, for which it is necessary to know the value of Φ_0 in Equation 1. For hydrogen at an energy of $E_0 = 10.6\,\mathrm{GeV}$, a flux of

$$\Phi_0 = 18.095\,\frac{\text{nucleon}}{\mathrm{m^2\,s\,sr\,GeV}} \tag{2}$$

can be determined [26]. To measure this flux at ground level, we need to take into account how these events will reach the detector. Since both gamma and cosmic rays produce showers, it is possible to observe them with the same instrument. For this reason, we want to familiarize ourselves with a useful model to describe these showers.

2.3 Heitler Shower Model

If we want to observe gamma or cosmic rays at ground level, we need to account for the behavior of those particles when interacting with Earth's atmosphere. Both photons and protons initiate cascades of secondary particles cascades, called showers, upon entering the atmosphere [23, p. 26]. We use the Heitler model to describe these showers, since it is useful although being fairly simple in terms of computational needs. In this model, shower behavior can be described by a few simple parameters and performs well when compared to a more complicated shower simulation.
In the following, we are referring to the paper in [24]. The main idea the model is based on, is that in electromagnetic showers induced by a gamma ray photon, a two-body

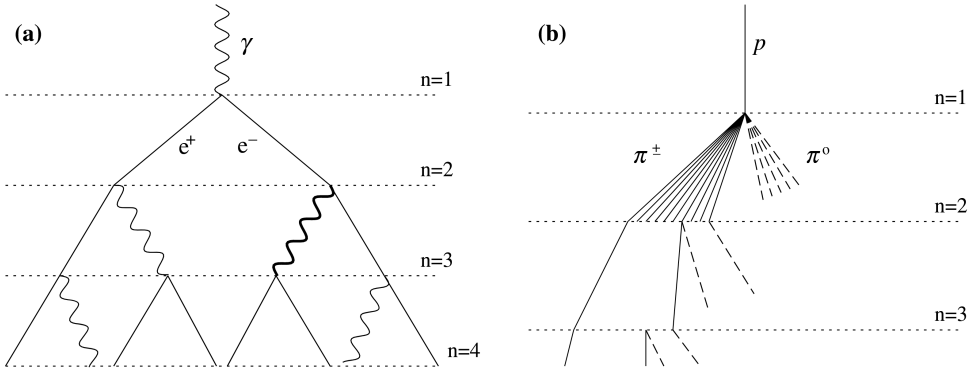

Figure 4: Schematic view of an electromagnetic (a) and hadronic (b) shower cascade. Not all pion lines are shown after the $n = 2$ level. The graphic was taken from [24].

splitting in the form of one-photon bremsstrahlung or e^+e^- pair production occurs at each interaction. The distance between interactions is dependent on the radiation length of the resulting particles. The total number of particles in the shower at each stage can then be easily described by the number of splittings n as 2^n. A schematic graphic of such an electromagnetic shower can be seen on the left side of Figure 4. The described multiplication process is stopped when the individual energies of the produced electrons and positrons drop below a critical energy ξ_c^e, where the energy by collision becomes dominant compared to radiative losses. This simplified model does not account for all details of shower behavior, but some relevant quantities can be simply described by their proportionality to the initial energy E_0, namely the final total number of particles

$$N_{\max} \propto E_0 \tag{3}$$

and the depth, where the shower reaches its maximum number of particles

$$X_{\max} \propto \log(E_0) \tag{4}$$

The behavior of hadronic showers is slightly more difficult than that of electromagnetic showers, but can still be approximately described in a similar fashion. In the case of a hadronic shower, each interaction is not limited to two-body-splittings, e.g. one proton might produce multiple neutral (π^0) and charged ($\pi^\pm$) pions, of which the former almost immediately decay to photons, producing electromagnetic subshowers as described before. During π^0 decay, at each step a third of the energy is deposited from the hadronic to the electromagnetic shower component. Charged pions continue to interact and produce a new generation of pions. Analogous to electromagnetic showers, multiplication is not stopped until particle energies fall below a critical energy ξ_c^π. Then, interaction is no longer favored compared to pion decay into muons, which can then be observed at ground level. An example of such a hadronic shower can be seen on the right side of Figure 4. In the case of hadronic showers, the atmospheric depth of the first interaction can be approximated by the primary proton energy:

$$X_0 = \lambda_I \ln 2 \tag{5}$$

$$= \left(61 \frac{\mathrm{g}}{\mathrm{cm}^2}\right) \left(1.0 - 0.1 \ln\left(\frac{E_0}{\mathrm{PeV}}\right)\right) \ln 2 \tag{6}$$

with the interaction length of the primary proton λ_I. This however only gives an estimation of typical values for X_0, since the point of first interaction is a highly statistical and probabilistic quantity and can therefore not be easily computed from the primary energy. The model also includes an approximation of $X_{\max}$:

$$X_{\max} = X_0 + \lambda_r \ln\left(\frac{E_0}{3N_{\text{ch}}\xi_c^e}\right) \tag{7}$$

with the radiation length λ_r, the multiplicity of charged particles N_{ch} and the critical energy ξ_c^e found in electromagnetic showers. [24]

2.4 Shower Detection

In ground-level detection, the ultra-relativistic particles produced in the shower are observed via Cherenkov radiation. Cherenkov radiation is blue light that is emitted when a charged particle moves through a medium at a velocity $v > c_n$ greater than the speed of light in that medium. Therefore only particles with an absolute velocity greater than $\frac{c}{n}$, where n is the refractive index of the medium, are detected. Since the particles produced in the cascades are fast enough to emit Cherenkov radiation, it is possible to detect them via the Cherenkov light they produce in air, as is the case at H.E.S.S. (s. Figure 5). Another option is the detection of Cherenkov light produced in water, as done at HAWC (s. Figure 6) [23, p. 264 f.]. When secondary particles pass through large water tanks, their emitted Cherenkov radiation is detected via photomultipliers (PMTs). From the number, location and time when PMTs on the array light up, one can reconstruct the direction of the shower core, which corresponds to the arrival direction of the primary photon or proton using a Gaussian fitting procedure. In this case, a plane with a second-order curvature model of a shower front is fitted to arrival times, making a very accurate event timing indispensable [4, p. 6 f.]. To differentiate between photons and protons, a neural network, which was trained to differentiate between gamma-ray showers and showers of hadronic origin, is often used [15].
Later we show, how by monitoring the shadow cast by the moon in cosmic ray flux, the resolution and pointing accuracy of the detector's angular reconstruction can be examined and transferred to gamma rays for the case of SWGO [14, p. 1].

Figure 5: The full H.E.S.S. array in Namibia. The figure was taken from [9].

Figure 6: The HAWC Observatory in Mexico. The figure was taken from [8].

2.5 SWGO's Layout & Detector Design

The main predecessors of SWGO are HAWC in Puebla, Mexico [8] and LHAASO in Sichuan, China [28]. When completed, SWGO is expected to enable the observation

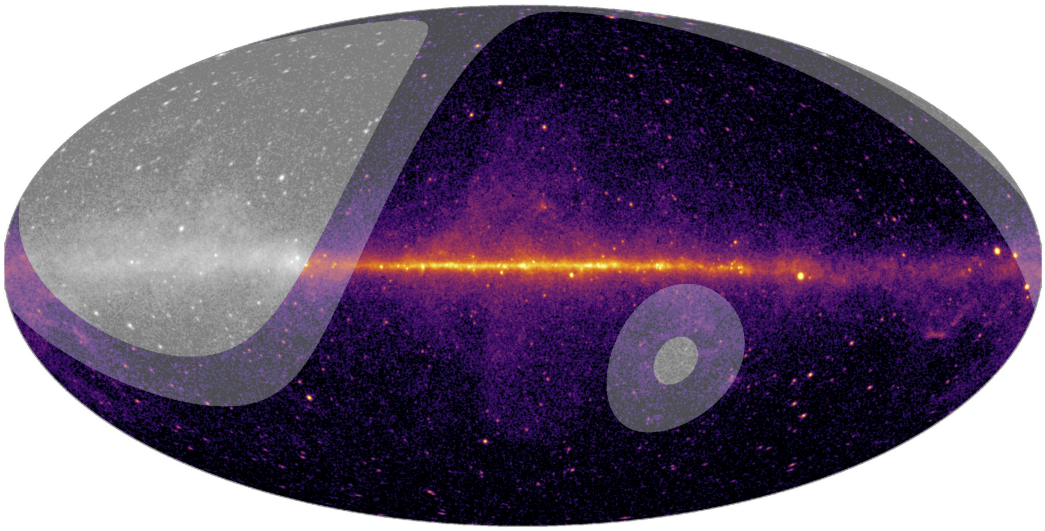

Figure 7: Visibility map showing the sky observable to SWGO above 45° (full color) and 60° (color shaded) together with Fermi-LAT all-sky image. Regions invisible to SWGO are shown in grayscale. The graphic was taken from [11, p. 5, Fig. 1].

in a field of view previously inaccessible to previous detectors with unprecedented precision. The site the observatory is set to be constructed at is located near San Pedro de Atacama, Chile, at a longitude of $-67.7°$ and a latitude of $-22.9°$. The total radius of the observatory is set out to be 560 m, which is split into three detection zones with increasing radius and decreasing fill factor of water Cherenkov detectors (WCDs). An overview of the planned array can be seen in Figure 8. [11, p. 4]

The design of these WCDs differs from previous approaches, since SWGO would be the first detector to use a two layer tank design, consisting of a cylindrical steel tank containing a bladder filled with water. This bladder is in turn subdivided into two volumes of different dimensions. Each of the two volumes includes a PMT with the lower one directed downwards and the upper one directed upwards, which will be able to detect the Cherenkov photons emitted by secondary particles passing the WCD. This dual-layer approach is employed to further improve the cosmic ray rejection of the detector. A simulation of a vertical muon passing the WCD with added dimensions of the tank is shown in Figure 9. [11, p. 4 ff.]

Beside practical criteria, like connection to an international road and proximity to the nearest city, this location is especially compelling, since it lies in the Southern Hemisphere. This makes it an interesting starting point for mapping of large scale emission and observing the full sky, since no instrument with such a wide field of view exists in the Southern Hemisphere. The visibility of SWGO of the sky can be seen in Figure 7. It will also provide access to observations of the Galactic Center. Furthermore, it is located at a height of 4770 m above sea level. A high altitude is preferable for this type of detector, since secondary shower particles need to be measured directly and therefore, a position enabling the observation of a large amount of particles is needed, especially for lower energy gamma rays. As shown by Equation 4, this makes a higher position in the atmosphere necessary. The observable energy of gamma rays expected to

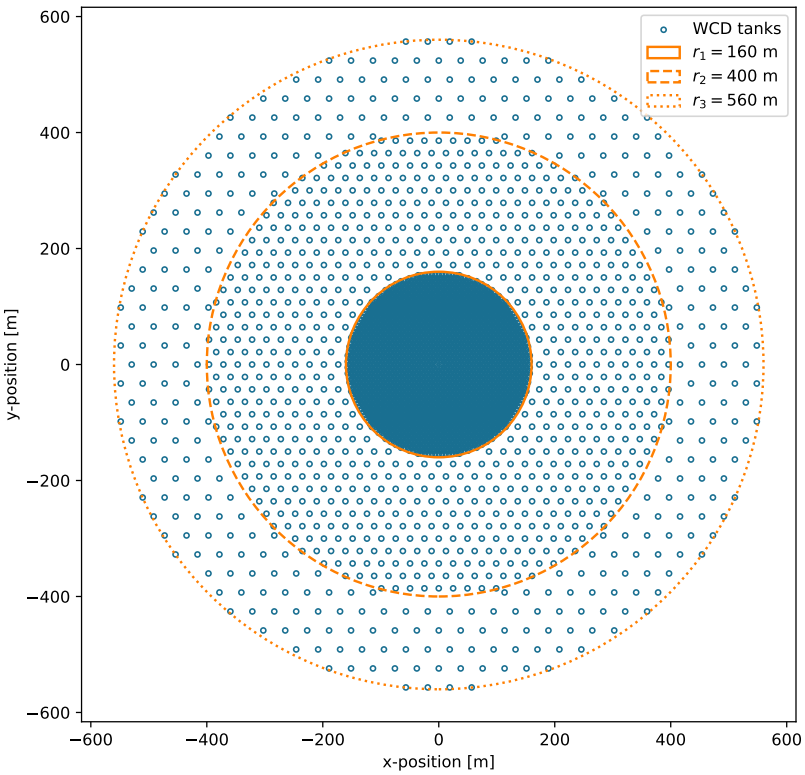

Figure 8: Planned layout of WCD tanks on the detector array. The tanks on the innermost array are positioned directly next to each other.

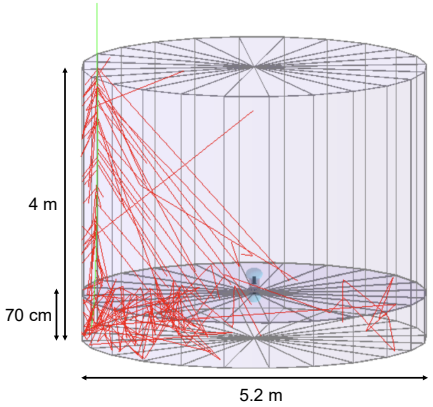

Figure 9: Simulation of a vertical muon producing Cherenkov photons, which are shown in red and green respectively. Note that the photons are reflected in the lower chamber. The two photo-multipliers are shown in light blue. The graphic was taken from [11, p. 6, Fig. 3]. The shown dimensions were added retrospectively.

be achieved lies in a range of 100 GeV − 1 PeV in orders of magnitude. [11, p. 3 f.]

2.6 Coordinate Systems

For the later examinations of the moon shadow, we need to first define suitable coordinate systems for each use case. It is necessary to describe the arrival direction of particles relative to the detector and their trajectory in space, in addition to the geomagnetic field that they pass on their way to our detector. The basis for each of these coordinate systems is a spherical coordinate system, which we want to explain first.

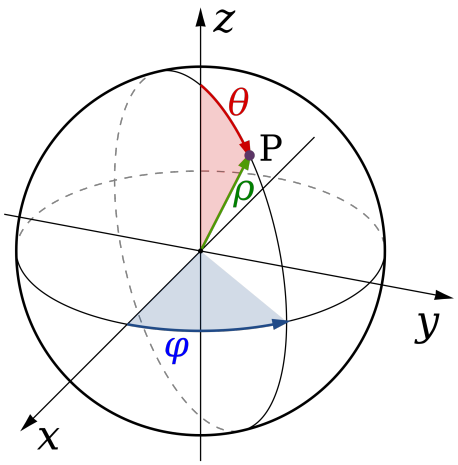

Figure 10: Definition of angles in a spherical coordinate system relative to cartesian coordinate axes. The figure was taken from [19].

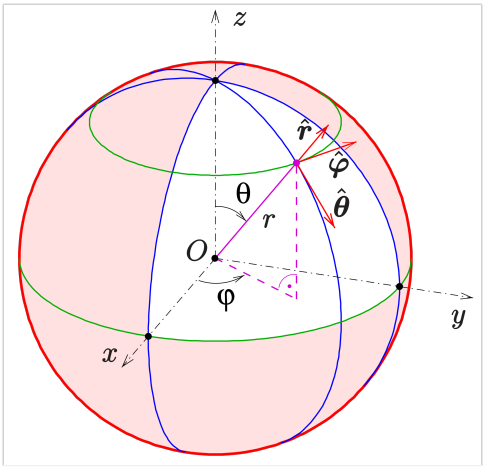

Figure 11: Unit vectors in spherical coordinates. The figure was taken from [5].

Spherical Coordinates Right handed spherical coordinates are the base for most astronomical coordinate systems. In this system, a point is specified in three-dimensional space using a distance r, sometimes also referred to as ρ, and two angles θ and φ, which are defined as shown in Figure 10. The transformation to cartesian coordinates is defined as follows:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \sin\theta \cos\varphi \\ r \sin\theta \sin\varphi \\ r \cos\theta \end{pmatrix} \tag{8}$$

In this definition, the cartesian axes point in the following directions in spherical coordinates:

$$\begin{aligned} x &\to \varphi = 0°, \theta = 90° \\ y &\to \varphi = 90°, \theta = 90° \\ z &\to \theta = 0° \end{aligned}$$

The value of φ for $\theta = 0$ is arbitrary and therefore defined as $\varphi = 0°$. To describe a vector field or a velocity in spherical coordinates, we are in need of the corresponding unit vectors $\hat{e}_r, \hat{e}_\theta, \hat{e}_\varphi$, which are shown in Figure 11, where they are referred to as $\hat{r}, \hat{\theta}, \hat{\varphi}$.

They can be derived as follows [29, p. 59 f.]:

$$\hat{e}_r = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}, \quad \hat{e}_\theta = \begin{pmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{pmatrix}, \quad \hat{e}_\varphi = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix} \tag{9}$$

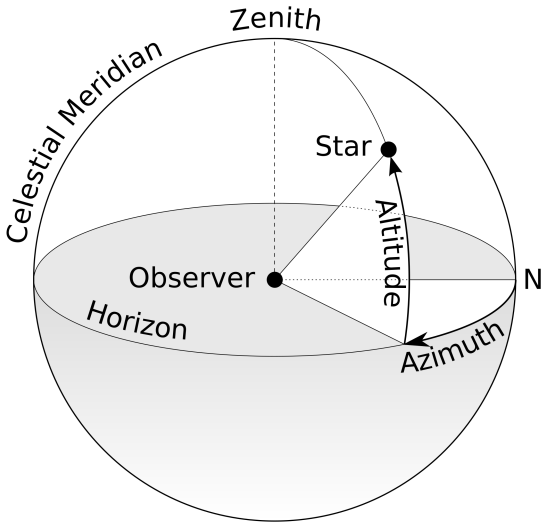

Figure 12: A schematic view of the horizontal coordinate system. In this picture, a cartesian x-axis would point in the direction of "N", the y-axis in positive Azimuth = 90° direction on the same plane and the z-axis in the direction of "Zenith". The graphic was taken from [30].

Horizontal Coordinate System The coordinate system that is used to describe the arrival direction of particles relative to the center of the detector array is called the horizontal coordinate system, since it uses the observer's horizon as a reference plane. The horizon is defined as the tangential plane to the observer's location. By defining the altitude and azimuth angle relative to this horizon as seen in Figure 12, the position of celestial objects or particles on the observer's sky can be described. An equivalent definition uses a zenith angle instead of the altitude, which is then the angle between the object and the pole of the upper hemisphere, also named zenith. [6, p. 16 f.]
It is important to note here, that, unlike the usual spherical coordinates, the horizontal coordinate system is a left-handed system, since the z-axis points outwards, perpendicular to Earth's surface, the x-axis points in the direction of North and the y-axis in the direction of East, as seen from the point of observation. A schematic view of the horizontal coordinate system is shown in Figure 12. The transformation of different coordinate systems will be explained further in subsection 5.3, when we compute the deflection of cosmic rays in the geomagnetic field.

Equatorial Coordinates In equatorial coordinates, the Earth's rotational axis and equatorial plane are used as a reference, since their positions remain constant. The coordinates describing this system are the angles right ascension, commonly referred to as α, (RA) and declination δ (DEC). As seen in Figure 13, the right ascension measures the

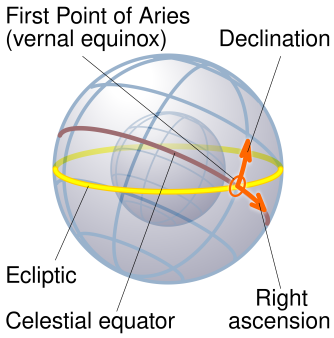

Figure 13: A schematic view of the equatorial coordinate system relative to the earth. The direction of right ascension and declination are shown. The graphic was taken from [21]

angular distance of the observed object to the March equinox along the celestial equator, whereas the declination measures the distance perpendicular to the celestial equator in positive and negative direction, meaning the North celestial pole has a declination of $+90°$. Its advantage lies in its universality, being independent of the observer's position. This makes it useful to compare results across observatories located at different sites.

Geocentric Coordinates Furthermore, we need a coordinate system to describe quantities relative to a position on or above the Earth's surface. In this context, we define spherical geocentric coordinates, using the distance from the center of the Earth r, the colatitude θ and the longitude φ as coordinates. The definitions of the longitude and latitude are shown in Figure 14, being an equivalent description of the coordinates. The difference lies in the range of the defined angle, since the colatitude varies from $0°$ at the North pole to $180°$ at the South pole instead of the range of the latitude shown in the figure. So our geocentric coordinate system is simply a right-handed spherical coordinate system as described before, with the center of the Earth as the coordinate origin and rotating with it. A field in this system can then be described by its components in the direction of $\hat{e}_r, \hat{e}_\theta$ and $\hat{e}_\varphi$. [20]

Sky Offset Frame Lastly, to observe the moon shadow and produce counts and sky maps, we need to define our particle's coordinates relative to the position of the moon at the time of observation. In order to achieve this, the `SkyOffsetFrame` frame class from the `astropy` library is used. The frame is defined with an origin, which we set to the moon position. The orientation of the frame will then match the orientation of the original coordinate system, with the new coordinate angles being named longitude and latitude. In our case, the longitude will point in the direction of azimuth, and latitude will point in the zenith direction, since we define the moon position in a horizontal coordinate system. Later we will rotate this frame to depict equatorial coordinates, meaning the longitude and latitude will then correspond to right ascension and declination, respectively. [13]

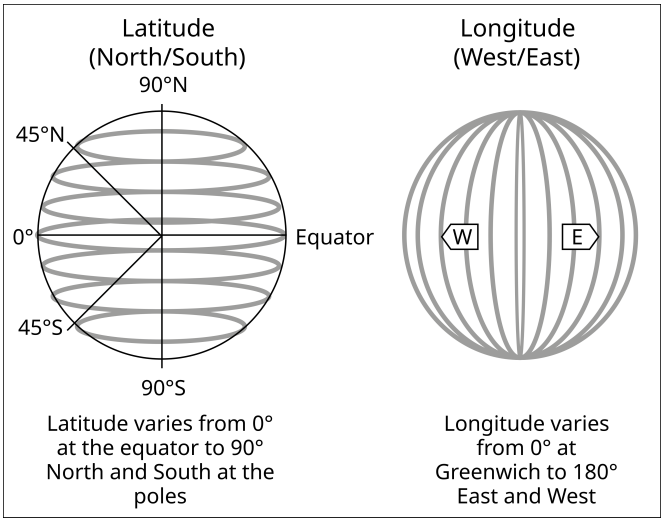

Figure 14: The definitions of longitude and latitude in a geocentric coordinate system. The graphic was taken from [1].

3 Simulation Analysis & Data Set Creation

To interpret the real data that will be taken by the future observatory SWGO, we need to understand the physics of the air showers and the functionality of the detector. Usually this is done by using Monte Carlo simulations. For the air shower, our simulation is based on the CORSIKA simulations [18] and the detector's response is simulated based on a software called HAWCSim [11, p. 8]. Later we will find that we can't directly use these simulations for moon shadow observations, leading us to develop a custom, simpler simulation based on the Instrument Response Functions (IRFs) that we inferred from the original simulation data.

3.1 Shower & Detector Simulations

3.1.1 Primary Data

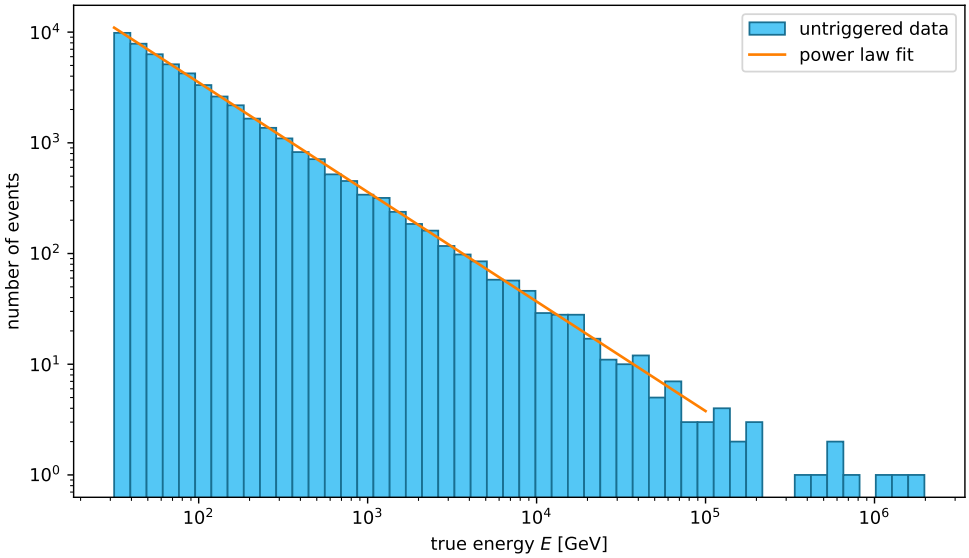

Figure 15: The energy spectrum as simulated by CORSIKA. A power law was fitted to the histogram to examine the function the distribution follows.

Firstly, we will take a look at the primary data. These are events as they are being simulated in the atmosphere above the observatory before being thrown onto the detector, therefore including untriggered events. In our case, only protons are simulated with an energy spectrum as shown in Equation 1 with $\gamma = 2$ and an isotropic distribution over the detection array. The utilized spectrum is therefore less steep than physically true, since this is useful to get a sufficient amount of events in the higher energy range, making the statistics more meaningful and enabling us to also evaluate our algorithms at these higher energies. To verify this expected energy spectrum and the event isotropy, we now discuss the distributions found in the primary data. However, the results of this evaluation are not used for further applications, making it sufficient to investigate only a small data set of 50000 events.
For the energy spectrum, we obtain a histogram that seems to be distributed according to the expected power law (s. Figure 15). To verify this and determine its exponent

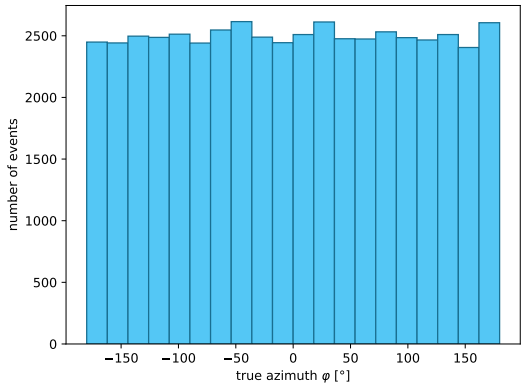

Figure 16: Distribution of the azimuth angle as it is being simulated by CORSIKA.

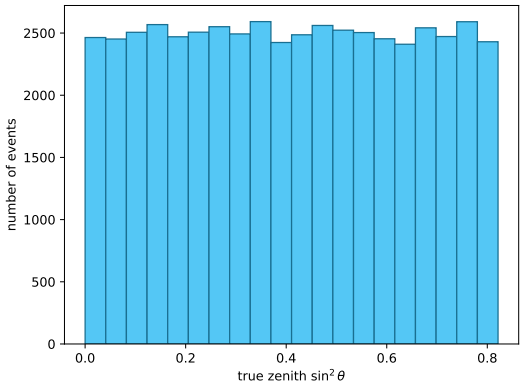

Figure 17: Distribution of the zenith angle as it is being simulated by CORSIKA.

γ, we perform a fit using the `curve_fit` function from the `scipy` library [12]. We only include values in the energy range up to 100 TeV, since statistics are generally not sufficient at the highest energies. We use the following function

$$f(E) = a \cdot \left(\frac{E}{E_0}\right)^{-\gamma}, \quad E_0 = 1\,\text{GeV} \tag{10}$$

where E is the true energy and a and γ are the parameters being fitted. We obtain the following optimal parameters:

$$\gamma_{\text{opt}} = 0.9839 \pm 0.0042, \quad a_{\text{opt}} = (334 \pm 6) \cdot 10^3\,\text{nucleons} \tag{11}$$

The standard deviations were taken from the resulting covariance matrix determined in the fitting process. This result of $\gamma_{\text{opt}} \approx 1$ is consistent with our expectation, because binning the energy spectrum effectively shows us an integrated spectrum, for which the exponent should in turn be

$$\gamma_{\text{opt}} \approx \gamma - 1 = 1 \tag{12}$$

To examine the isotropy, we plot distributions of the zenith angle θ and the azimuth angle φ. The azimuth is, as expected, uniformly distributed in a range of $[-180°, 180°)$ (s. Figure 16), since an equal amount of particles should arrive from every direction. The zenith angle distribution is not as trivial. The area of a circle of latitude decreases with the zenith angle with a proportionality of $\propto \sin\theta$, as visible in Figure 14. Additionally, the area of the detector as seen from the arrival direction of the particle is dependent on the zenith angle with $A_\theta = A_{\text{det}} \cdot \cos\theta$. The zenith angle should therefore be distributed with

$$\theta \sim \sin\theta \cos\theta \tag{13}$$

[16, p. 42] making the distribution uniform when plotting values of $\sin^2\theta$. This expectation is again verified in Figure 17, where a uniform distribution is observed, with zenith angles simulated in a range of $[0°, 65°]$.

3.1.2 Detection Data & Event Selection

Now we want to take a look at how the detector is expected to respond to these simulated hadronic showers, creating the data we ultimately expect to observe with the detector.

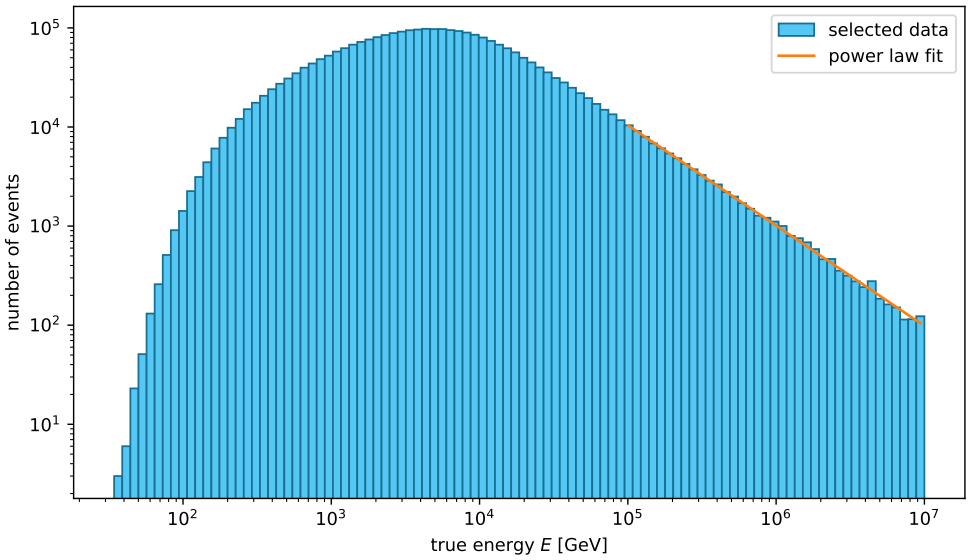

Figure 18: The energy spectrum after event selection as simulated by HAWCSim. A power law was fitted to the histogram to examine the underlying function.

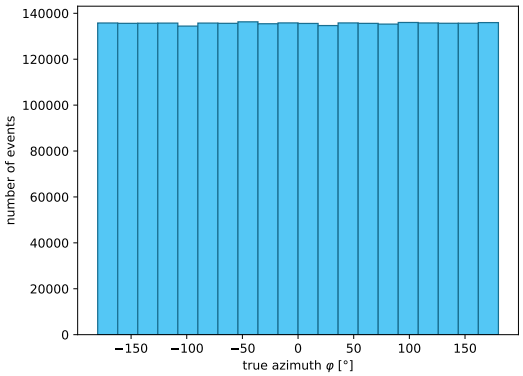

Figure 19: Distribution of the azimuth angle in HAWCSim simulations after event selection.

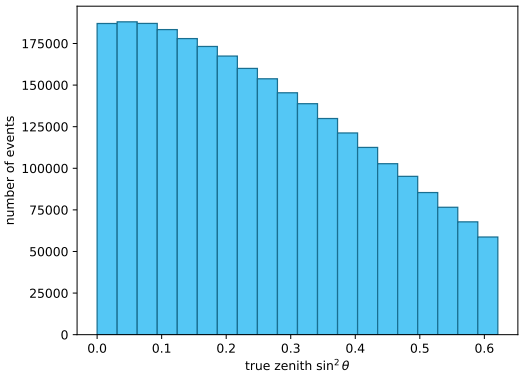

Figure 20: Distribution of the zenith angle in HAWCSim simulations after event selection.

We will not be able to reconstruct all cosmic rays that land on or near the detector with a sufficient accuracy, and therefore are in need of sensible criteria to select meaningful events. These criteria are:

- The number of triggered PMTs should have a threshold: $N_{\text{Hit}} \geqslant 65$. This is in accordance with the convention used by the collaboration [11, p. 8][1].
- The zenith angle should have a maximum value: $\theta \leqslant 52°$. At large zenith angles, particles have a long way to pass through the atmosphere, making reconstruction challenging. This is why we exclude events arriving at very large zenith angles.

[1]The collaboration wrongfully mentions a threshold in the number of stations instead of PMTs. This is a typo in the preprint on arXiv and to be corrected in next versions.

- The shower core should lie on the detector array: $\sqrt{x^2 + y^2} \leqslant r_{\text{det}} = 560\,\text{m}$
- The plane fit used for the event reconstruction should have converged. This criterion is not rooted in a physical event property, but a technical limit to the events observable by the detector.

After analyzing a total of $49659 \cdot 50000 \approx 2.5$ billion thrown events, about 2.72 million events remain after applying these criteria. We can now look at the distributions of energy, azimuth and zenith and compare them to the primary data to get an idea of the detector's behavior. The resulting energy spectrum can be seen in Figure 18. Here it is important to note, that the simulation used to create the data examined in this thesis does not yet include an energy reconstruction, which is why we regard the true energy of primaries E_{true} in our evaluations. For higher energies, we can still observe the spectrum behaving according to the expected power law, which we again verify via a fit, only including values above 100 TeV. We receive the following optimal parameters:

$$\gamma_{\text{opt}} = 1.101 \pm 0.006, \quad a_{\text{opt}} = (122 \pm 8) \cdot 10^7 \tag{14}$$

Here, the value of the exponent still follows the expectation $\gamma \approx 1$, and the value of a is larger simply because we consider a much larger amount of data than in the primary data discussion. For lower energies, there is a strong deviation from the original spectrum. Because of the small shower area at those energies and the smaller amount of particles (cf. Equation 4), not a lot of PMTs on the array are triggered, resulting in a worse resolution or rejection by the event selection, ultimately leading to less events being observed in that section of the energy range. The new spectrum can be described as a convolution of the energy spectrum with the energy dependent effective area of the detector

$$\int \mathrm{d}E\ \Phi(E) A(E) \tag{15}$$

which we will determine later. As seen in Figure 19, the azimuth angle still follows a uniform distribution, since showers behave identically in all directions. When plotting the zenith angle as $\sin^2\theta$, the distribution is not uniform anymore (s. Figure 20), as less events are detected at larger zenith angles. The reason for this lies in the physical properties of the shower: For a large zenith angle, a bigger amount of atmosphere has to be traversed by the shower, making it reach its maximum further away from the detector than for a small zenith angle. Therefore less secondary particles can be observed by the detector, resulting in a smaller count of events. We will further examine this phenomenon later.

3.1.3 Effective Detector Area for Protons

As we have already mentioned when comparing Figure 18 to Figure 15, not all events are detected by the array and the probability of an event being observed seems to depend on the energy of the primary particle, resulting in a different energy spectrum. To quantify this effect, we define a detection efficiency $p(E)$ that describes how likely a shower is to be successfully detected given its primary energy. This quantity is directly proportional to the often used effective detector area:

$$A(E) = p(E) \cdot A_{\text{sim}} \tag{16}$$

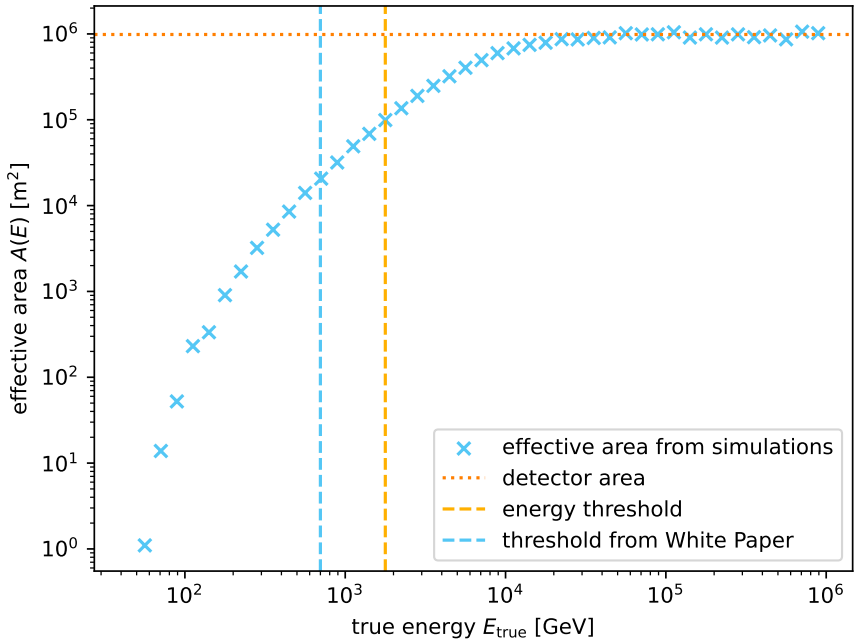

Figure 21: The effective area $A(E)$ as determined in HAWCSim simulations. The marker lies on the middle point of each energy bin on a logarithmic scale. The detector area is drawn to illustrate the saturated effective area equaling the area of the detector. The energy threshold is the energy, at which the effective area is 10 % of the saturated value, in our case $E_{\text{thr}} \approx 1800\,\text{GeV}$. The energy threshold found by the collaboration was $E_{\text{thr}} \approx 700\,\text{GeV}$, since different selection criteria were applied [11, p. 37, Fig. 27].

Where A_{sim} is the area on which the primary particles are simulated, in our case $A_{\text{sim}} = (1500\,\text{m})^2\pi$. We use a numerical approach to determine the detection efficiency and effective area. In this method, we simply divide the bin counts determined in the selected events by their respective counts in the same bins in the primary data. We receive a discrete function for $p(E)$ and accordingly $A(E)$, shown in Figure 21. The effective area increases until reaching a saturation level at the value of the detector area A_{det}. This is because after a certain saturation energy, almost all showers landing on the array are detected, and due to the cut applied on the core positions in the event selection, the effective area cannot exceed the area of the detector. This is consistent with the energy dependent effective area determined by the SWGO Collaboration [11, p. 37, Fig. 27], although the energy threshold E_{thr}, at which the effective area is 10 % of the saturated value, found by us was about one order of magnitude larger. This could be caused by a slight difference in the event selection criteria, since with stricter criteria less events are detected at lower energies, making the threshold energy larger.

3.1.4 Angular Resolution of the Event Reconstruction

Previously, we were only looking at the true directions from which the primary particles arrived, defined by their azimuth and zenith angles. We now want to determine how accurately the events are reconstructed by the plane fit. For this we compare the true and predicted directions for each proton and derive the angular resolution of the detector

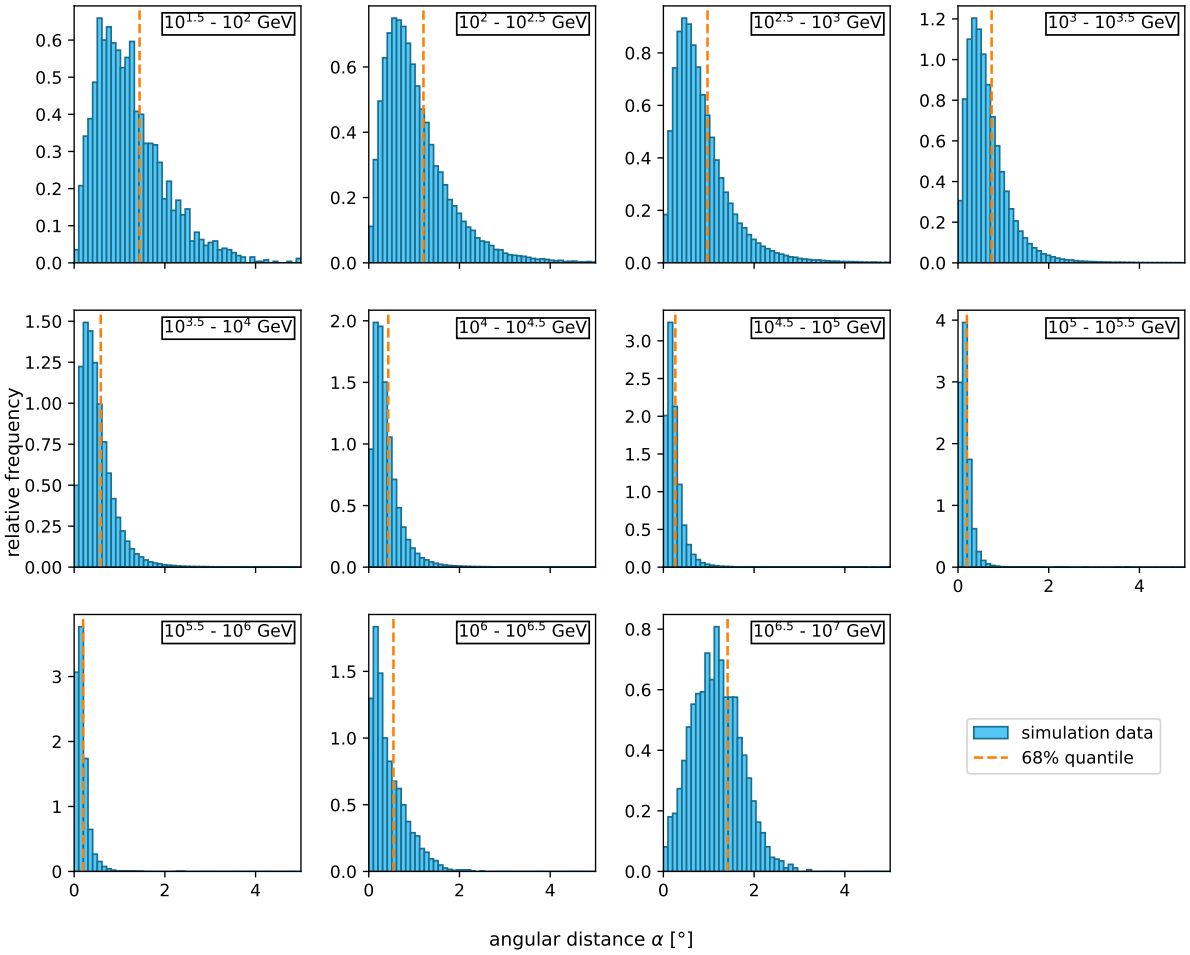

Figure 22: Distribution of the angular distance for each energy bin. Only a range of $\alpha \in [0°, 5°]$, because the amount of events with a greater angular distance is relatively small. Note that the y-axes show a relative frequency of events, the amount of events in each energy bin is determined by the energy spectrum of detected events.

from that data. We firstly define the angular distance of a singular event as the angle α between the true and reconstructed arrival direction, $\vec{r}_{\text{true}}$ and $\vec{r}_{\text{rec}}$, which can be described by the following formula:

$$\cos\alpha = \frac{\vec{r}_{\text{true}} \cdot \vec{r}_{\text{rec}}}{||\vec{r}_{\text{true}}|| \cdot ||\vec{r}_{\text{rec}}||} \tag{17}$$

After computing this angular distance, we split all events into energy bins (2 bins per decade), following the convention used in the collaboration. The distributions in all these energy bins can be found in Figure 22. Although events with $\alpha > 5°$ are not shown in these distributions, all showers are of course taken into consideration in the further analysis. In general, distributions become more narrow with higher energies, except at the highest energies, where they get broader again. It is also interesting to note that the distributions do not peak at $\alpha = 0°$, but are shifted to the right. This behavior can be traced back to a geometrical effect. The number of possible reconstructed directions at

a given angular distance increases with the solid angle element $\mathrm{d}\Omega = 2\pi \sin\alpha \, \mathrm{d}\alpha$, which vanishes for $\alpha \to 0°$. Therefore the probability density tends towards zero for very small α.

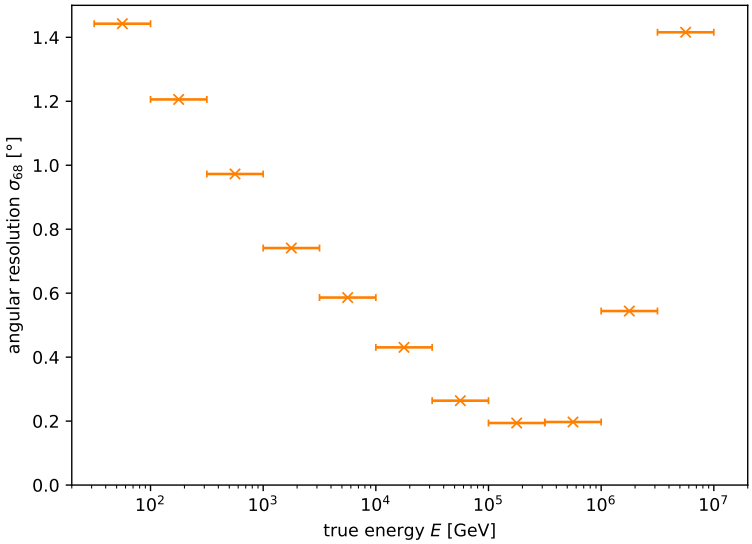

Figure 23: The angular resolution as a function of energy. The marker is placed in the center of the corresponding energy bin, while the error bars show its size.

The angular resolution as a function of energy is defined as the 68 % quantile σ_{68} of the angular distance for each energy, the energy dependency for all events can be seen in Figure 23. It is mainly important to note that, as expected, the resolution generally gets better for higher energies, since the event reconstruction becomes more accurate due to the larger amount of secondary particles being produced in the shower and being able to hit the detector. This is true up to a point at the highest energies where the statistics become insufficient and the event reconstruction becomes harder, since the whole detector array might light up during a particularly large air shower. This is also consistent with the findings of the collaboration [11]. Next, we want to consider the possible influence the zenith angle and shower core position on the array might have on the angular resolution. For this, we differentiate between three zenith bands

$$0° - 30°, \quad 30° - 45°, \quad 45° - 52° \tag{18}$$

and take into account whether the shower core is on the inner or the outer part of the detector array, divided by the radius of the smallest detection zone $r_1 = 156\,\mathrm{m}$. As one might expect from the higher fill factor of this innermost section, the angular resolution tends to be better for events landing on the inner array across all zenith bands (s. Figure 24). Furthermore, it behaves similarly for all zenith bands, although a slight tendency for a better resolution, especially in the outer array, at smaller zenith angles can be seen.

3.1.5 Atmospheric Depth $X_{\max}$ & Interaction Height X_0

Finally, we want to get an idea of at what height primary particles begin to interact with the atmosphere and initiate shower cascades. For this, we first look at the

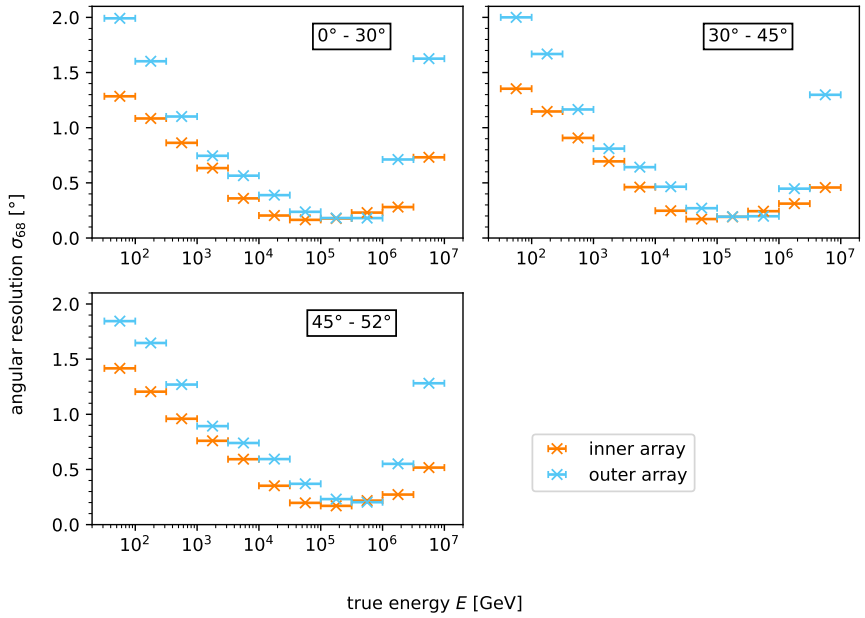

Figure 24: The angular resolution for different zenith bands and split into events landing on the inner and the outer array.

atmospheric depth at which the shower reaches its maximum number of shower particles $X_{\max}$, to which we have access because the CORSIKA simulations store this value. In Figure 25 we see that the values for $X_{\max}$ seem to be mainly located in a range between $2 \cdot 10^2 \frac{\text{g}}{\text{cm}^2} - 1 \cdot 10^3 \frac{\text{g}}{\text{cm}^2}$ with their maximum at $\approx 400 \frac{\text{g}}{\text{cm}^2}$. For comparison, the atmospheric depth of the entire atmosphere is $\approx 1000 \frac{\text{g}}{\text{cm}^2}$. Figure 26 shows the same distribution but with counts on a logarithmic scale to show that the tail of the distribution on the right of the maximum is larger than the left, since values on the left are restricted. This is because the atmospheric depth of a height where there is no atmosphere is defined as $X = 0 \frac{\text{g}}{\text{cm}^2}$ and primary particles can't initiate cascades before reaching the atmosphere, limiting low values of $X_{\max}$. On the right-hand side, the zenith angle of the arrival direction can make the value of $X_{\max}$ larger than the total atmosphere we would expect when moving straight upwards. Therefore, there is a higher amount of events in that section, although it is not entirely assured the distribution can be explained by only this effect, which is why this is an aspect to be examined further. We will take the mentioned impact of the particle slant into account later.

From this point, we want to find the height above sea level of the first interaction point of the particle. The atmospheric depth associated with this height is X_0, which is unfortunately not easily determined, since it is a highly statistical quantity. Equation 5 gives us an estimation of typical values of X_0 in dependency of the respective primary particle energies E_0. The overall values of X_0 are expected to be distributed according to an exponential function, with counts decreasing for higher values of X_0 [31, p. 2, Eq. 1], which can be attributed to the probability at which a primary particle will interact with the atmosphere at a certain atmospheric depth. When using Equation 5 for our energy spectrum in Figure 27, we therefore only find a value corresponding to a mean value of the distribution in each energy bin. Here, the values of X_0 are mainly distributed in the

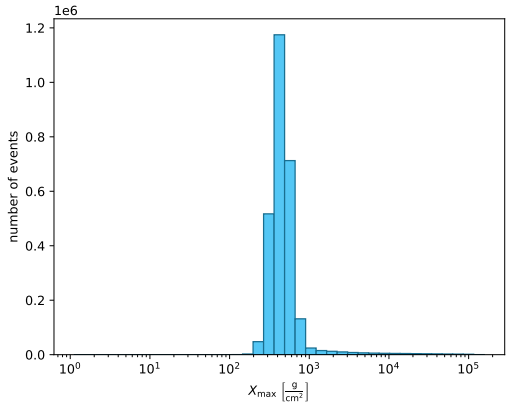

Figure 25: The distribution of $X_{\max}$ of CORSIKA showers. Singular events where $X_{\max} < 0\,\frac{g}{cm^2}$ were removed since they are not physically sensible.

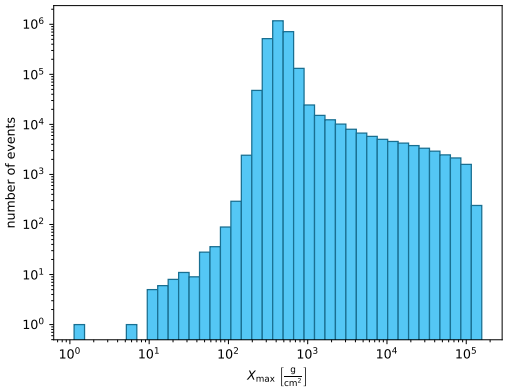

Figure 26: The distribution of $X_{\max}$ with event number on a logarithmic scale.

range between $50\,\frac{g}{cm^2} - 80\,\frac{g}{cm^2}$, showing a maximum at about $65\,\frac{g}{cm^2}$, with a larger tail on the left-hand side of the maximum. When plotting the computed distribution again with the event count on a logarithmic scale, we can see that the spectrum look very similar to the detected energy spectrum, the difference being that the x-axis is inverted and not on a logarithmic scale, as we would expect when regarding the energy spectrum. This is because lower values of E_0 correspond to higher values of X_0, resulting in an inverted distribution, and because the relation between X_0 and E_0 is logarithmic, which is why we observe a similar curve even though one axis is not logarithmic. Alternatively, we can compute an estimation for X_0 from Equation 7 to take the distribution $X_{\max}$ into account and to receive a distribution of values for particles of the same primary energy. Rearranging of this equation yields

$$X_0 = X_{\max} - \lambda_r \ln\left(\frac{E_0}{3N_{\text{ch}}\xi_c^e}\right) \tag{19}$$

For the other parameters, we use values determined in [24]:

$$N_{\text{ch}} = 41.2\left(\frac{E_0}{1\,\text{PeV}}\right)^{\frac{1}{5}}, \quad \xi_c^e = 85\,\text{MeV}, \quad \lambda_r = 37\,\frac{\text{g}}{\text{cm}^2} \tag{20}$$

Using this formula we receive the distribution of X_0 seen in Figure 29. Here we can see that the maximum is higher than before, being located at about $150\,\frac{g}{cm^2}$. When looking on the same distribution on a logarithmic scale (s. Figure 30), we can again see an exponential curve left of the maximum. This does not agree with our expectation, since we assume the values of X_0 to be distributed according to a declining exponential curve, which is opposite of what is seen here. It is possible, that the formula given in [24] is not as accurate as we originally assumed for the determination of X_0 from $X_{\max}$, and that the distribution is still dominated by the primary energy spectrum instead of the interaction probability. If we were to accurately simulate X_0, we would need to utilize the values found in Figure 27 as a mean in each energy bin and pick the values of X_0 from an exponential distribution of this mean. Nevertheless, we decide to continue

with the values found from Equation 7, since $X_{\max}$ is a quantity that will be available to observe once the detector is built. This decision is justified by the realization that the height of the particle we want to examine here has only a minuscule effect on the overall deflection in the geomagnetic field, which we will see later in subsection 5.4.

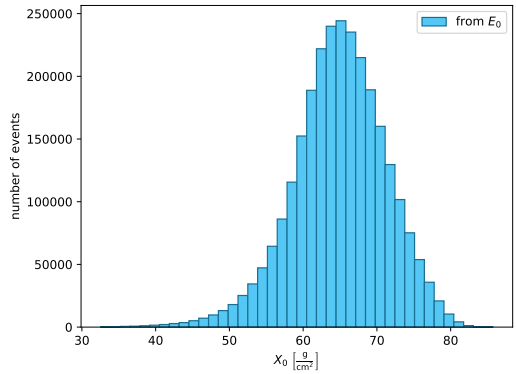

Figure 27: The distribution of X_0 computed via Equation 5 using only the primary energy E_0.

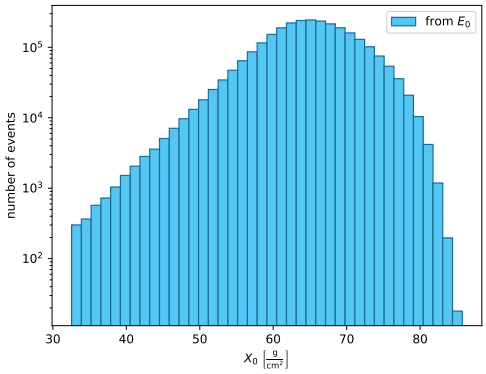

Figure 28: The distribution of X_0 shown in Figure 27 with event number on a logarithmic scale.

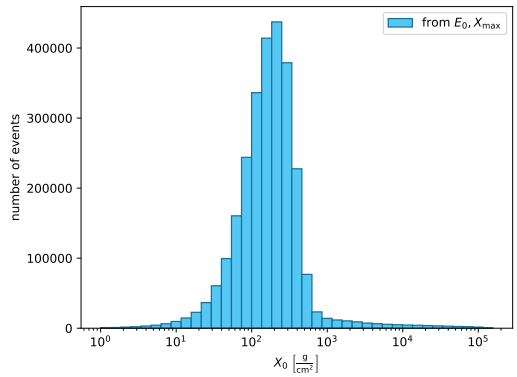

Figure 29: The distribution of X_0 computed via Equation 7 using both $X_{\max}$ and E_0.

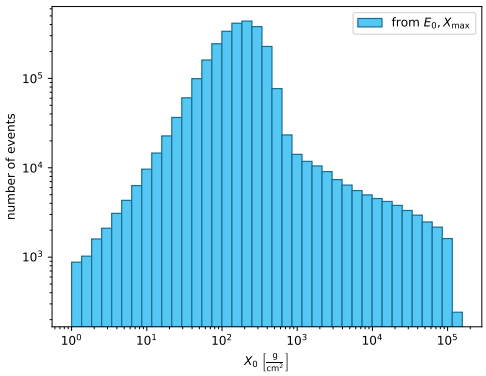

Figure 30: The distribution of X_0 shown in Figure 29 with event number on a logarithmic scale.

Lastly, we want to compute the height above sea level h this atmospheric depth X_0 corresponds to. For this, we assume an isothermic atmospheric model:

$$X(h) = X_{\text{sea}} \cdot \mathrm{e}^{-\frac{h}{H}} \tag{21}$$

We can rearrange this equation to:

$$h(X) = H \cdot \ln\left(\frac{X_{\text{sea}}}{X}\right) \tag{22}$$

We assume a scale factor of $H = 7.99\,\text{km}$ and an atmospheric depth at sea level of $X_{\text{sea}} = 1000\,\frac{\text{g}}{\text{cm}^2}$. Again, this is not an exact conversion but rather an estimation.

Additionally, we need to account for the different zenith angles of the events by unslanting the values of $X_{\max}$:

$$X_{\text{unslanted}} = X \cdot \cos\theta \tag{23}$$

It is important to note, that this formula is only valid if we can assume the Earth and atmosphere to be approximately flat in the domain of interest. This is however justified for our restricted range of zenith angels. We receive the distribution that can be seen in Figure 31 and find most values at about 20 km. On a logarithmic scale (s. Figure 32), we again find an exponential distribution with slope opposite to the expectation, which is now located at the right side of the maximum, since a higher value of h corresponds to a lower value of X as seen in Equation 22. For further usage of this data, we want to examine whether and how the zenith angle and energy of a particle have an effect on the height of first interaction. We therefore part our data set into the zenith bands and energy bins established in subsubsection 3.1.4. As seen in Figure 33, the zenith angle does not have a significant effect on the interaction height, which is why in following we assume and independency of the zenith angle, especially because the distribution is dominated by the first zenith band anyway and we only look at a moon position in this band in this thesis. The energy on the other hand has a significant influence on the distribution, as seen in Figure 34. Primaries with a lower energy seem to interact much later in the atmosphere, while higher energies yield a wider distribution of heights located at higher values. For further usage, we will therefore need to take the primary particle energy E_0 into account.

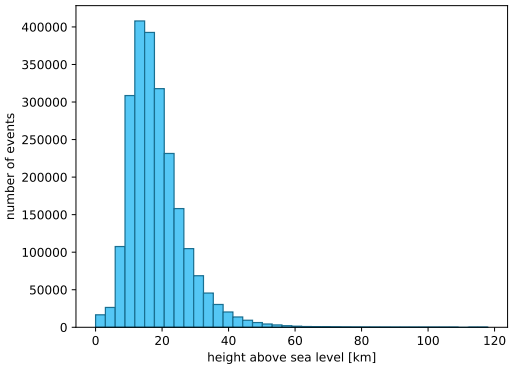

Figure 31: The height of first interaction h computed via Equation 22.

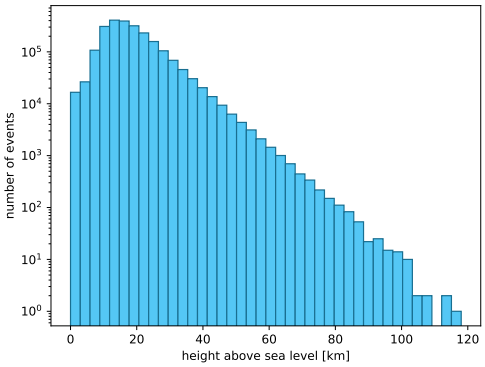

Figure 32: The distribution of h shown in Figure 31 with event number on a logarithmic scale.

3.1.6 Limitations of using the CORSIKA and HAWCSim Simulations

After obtaining all relevant IRFs, we now need to determine whether we can use these events to observe the moon shadow. For this, we can first calculate the observation time that would be equivalent to this amount of data. We can derive the rate by integrating

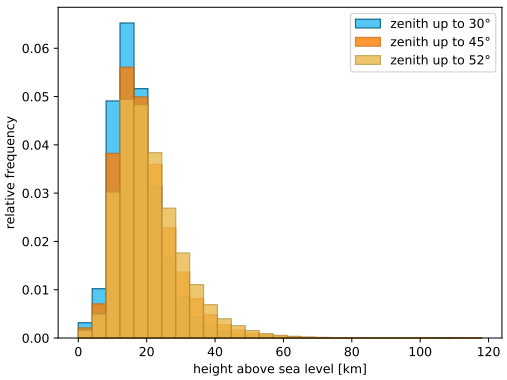

Figure 33: The distribution of h split into three zenith bands. A relative frequency is shown to enable better comparability between distributions.

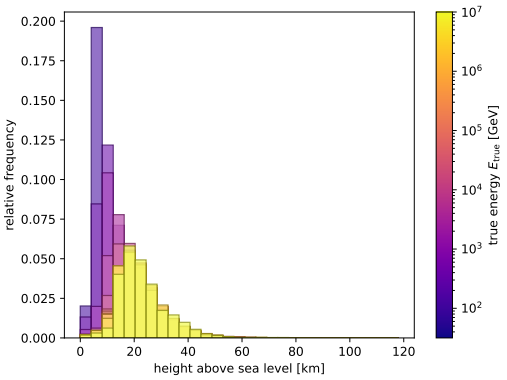

Figure 34: The distribution of h split into energy bins. A relative frequency is shown to enable better comparability between distributions. The color scale shows the energy of the upper bin edge for each bin.

the flux of cosmic rays:

$$f = \int_{E_{\min}}^{E_{\max}} \mathrm{d}E \int_0^{\Omega_{\det}} \mathrm{d}\Omega \int_0^{A_{\mathrm{sim}}} \mathrm{d}A \, \Phi_0 \left(\frac{E}{E_0}\right)^{-\gamma} \tag{24}$$

$$= \Phi_0 \cdot \frac{\Omega_{\det}}{-\gamma+1} \cdot A_{\mathrm{sim}} \cdot \left(\left(\frac{E_{\max}}{E_0}\right)^{-\gamma+1} - \left(\frac{E_{\min}}{E_0}\right)^{-\gamma+1} \right) \tag{25}$$

For this estimation, we assume the flux of the $\gamma = 2$ spectrum to have the same value Φ_0 as the true $\gamma = 2.7$ spectrum at an energy $E = 1\,\mathrm{TeV}$. This is just a rough assumption only employed to give an approximate value of the observation time. With the other values used in the simulation of $\gamma = 2$, $E_{\min} = 31.6\,\mathrm{GeV}$, $E_{\max} = 1.97\,\mathrm{PeV}$, $\Omega_{\det} = 2\pi(1 - \cos 65°)$ and $A_{\mathrm{sim}} = (1500\,\mathrm{m})^2\pi$ we find an event rate on the simulation area of

$$f_{\mathrm{sim}} = 16.5\,\mathrm{GHz} \tag{26}$$

making the observation time for our 2.5 billion events

$$t = \frac{n_{\mathrm{tot}}}{f} = \frac{2.5\,\mathrm{billion}}{54.3\,\mathrm{GHz}} = 0.15\,\mathrm{s} \tag{27}$$

Since we generally want to simulate a moon shadow observation for at least a few hours, this is obviously much too little. This is especially problematic since we will only need events in close proximity of the moon at a given observation time, making our sample data even smaller. To avoid oversampling from these distributions, we instead employ a custom simulation based on the determined IRFs from the shower and detector simulations, which will be further explained in the following.

3.2 Modular Simulation of Event Properties

To create our needed data, we design a modular simulation mechanism which can be expanded upon based on which event properties are needed. For this we first determine

how many events we need for a given observation time and then gradually add each property to the data set.

3.2.1 Observation Time & Detection Rate

To compute the number of events we expect to observe in a given time, we first need to know the rate of events on the simulated area. The simulated area will be a square window with dimension $n \times n$ degrees around the moon position on a `SkyOffsetFrame` [13] at a given observation time.

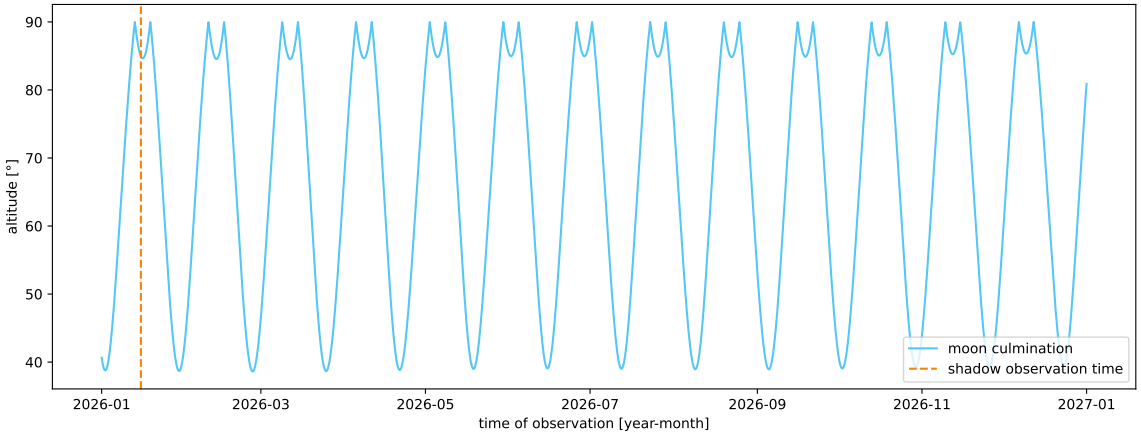

Figure 35: The altitude of the upper culmination of the moon as seen from the SWGO site over the course of one year. The dashed line marks the time, where the moon is in the position we have chosen to observe the shadow.

In Figure 35, the culmination of the moon in the year 2026 is shown, meaning the highest point the moon reaches in a given course over the observable hemisphere. The altitude angles of this culmination range from $\approx 40°$ to a maximum of $90°$, since that is the highest value the altitude can take on. However, the altitude of the actual moon position varies from $0°$ to $90°$, meaning the moon trajectory runs between the respective culmination and the horizon. The apparent indent in the culmination curve at high altitudes is caused by the latitude angle the SWGO site is located at: The culmination of the moon $h_{\max}$ can be calculated using the latitude angle φ and the declination angle of the moon δ:

$$h_{\max} = 90° - |\delta - \varphi| \tag{28}$$

For the latitude angle of the SWGO site, $\varphi = -22.9°$, there exist times where the declination is larger than the latitude, causing the culmination to dip during these periods of time.
In this thesis, we choose an observation time where the moon is positioned at an azimuth of $\varphi = 113.8°$ and a zenith of $\theta = 14.5°$ in order to profit from the better detection conditions generally found in the first zenith band. The chosen time is one of the first instances, where the moon reaches such an altitude in 2026.
The rate of events in that window is then again analogous to Equation 24:

$$f = \Phi_0 \cdot \frac{\Omega_{n\times n}}{-\gamma + 1} \cdot A_{\text{sim}} \cdot \left(E_{\max}^{-\gamma+1} - E_{\min}^{-\gamma+1} \right) \tag{29}$$

with Φ_0 as established in Equation 2, $\gamma = 2.7$, $E_{\text{min}} = 100\,\text{GeV}$, $E_{\text{max}} = 1\,\text{PeV}$, $\Omega_{\text{det}} = \frac{\pi}{(180°)^2} \cdot n^2$ and $A_{\text{sim}} = (1500\,\text{m})^2\pi$. For a 10 hour observation of a $3° \times 3°$ window around the moon center we find an event rate and count of

$$f = 15.33\,\text{kHz}, \quad n_{\text{tot}} = 0.55\,\text{billion} \tag{30}$$

This is however just the expected total number of proton cosmic ray events we expect in the window. We will not have to simulate all of these events, since most of them are not detected by our instrument with sufficient accuracy. The number of events we need to simulate will be determined in the following.

3.2.2 Detected Energy Spectrum Simulation

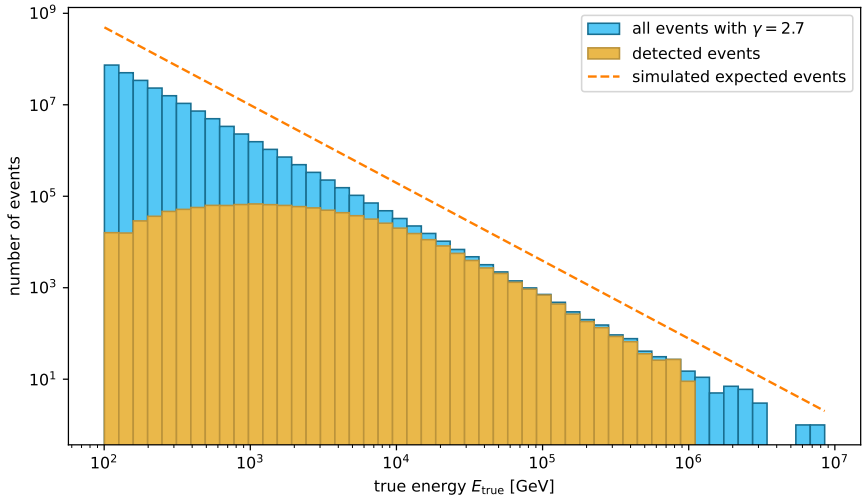

Figure 36: The energy spectrum simulated by our custom simulation. The orange line shows the total expected cosmic ray spectrum. The total spectrum of simulated events according to a power law are shown in blue, whereas the yellow histogram shows the energy spectrum after taking the effective area of the detector into account.

Another advantage of creating a custom simulations lies in being able to simulate an energy spectrum using $\gamma = 2.7$ instead of $\gamma = 2$ in the CORSIKA simulation. This provides us with a more accurate measure of moon shadow visibility, since the results will not include disproportionately more events in the higher energy range, where we have seen a better angular resolution. As seen in subsubsection 3.1.3, at higher energies with the best detection efficiency there is a constant offset between the detected and theoretical spectrum of about $\frac{(560\,\text{m})^2}{(1500\,\text{m})^2} \approx 0.14$ due to the size of the detector. Therefore we do not need to simulate the number of events corresponding to the entire expected energy flux, saving computation time by only simulating in the lower power law (s. Figure 36), at first assuming the best detection efficiency for all events. We can simulate this power law using the inverse transformation method: When picking values R from a uniform distribution $R \sim u(0, 1)$ and inserting the in the inverse definite integral of the energy function,

$$E = \left(R \cdot \frac{1-\gamma}{\Phi_0} \right)^{\frac{1}{1-\gamma}} \tag{31}$$

we receive energies E distributed according to our desired power law. We simulate a total of $n_{\text{tot}} \cdot p_{\max}$ of these energies (s. Figure 36) and then throw away values according to the detection efficiency curve in Figure 21. For this, we give each event the detection probability that corresponds to its energy bin used in determining the effective area and then throw out events based on this probability. After this, we receive energies distributed according to a power law spectrum that has been folded with the effective detection area, as seen in Figure 36, and a total number of events that varies slightly due to the detection efficiency. In further steps, we will only need to simulate properties for this number of events n_{det}, which is why it is useful to simulate the energy first to determine the event number.

3.2.3 True Arrival Direction

To simulate the true position of cosmic ray primaries, we first define a `SkyOffsetFrame` (s. subsection 2.6) pointed in the direction of the moon at the time of observation in a horizontal system. We then simply pick values for longitude and latitude in this frame distributed uniformly in a $n \times n$ square window around the moon position. It is important to note that this simulation window should be larger than the desired window we want to show to remove edge effects. The largest window we simulate has an extent of $10° \times 10°$, although we will later see, that an even larger window would be beneficial when taking the geomagnetic field into account. We have verified that this method produces isotropic events as seen by the observer. For this thesis, we choose a fixed time of observation on January 15, 2026 at 13:00 to get a moon position with a relatively high zenith angle as mentioned before in subsubsection 3.2.1. The moon position at this time was determined via the function `get_body` from the `astropy` library (s. Figure 35).

3.2.4 Angular Resolution & Detected Position

Finally, to determine how we will detect these simulated events on the instrument, we need to account for the angular resolution we determined before, by randomly adding noise based on the before determined angular resolution. We do this by assigning each event an angular distance, for which we get the data from HAWCSim of the corresponding zenith band and split our events into the same energy bins we used before. We randomly choose a value of α from the respective distribution for each of these events. Due to the larger amount of events needed for our custom simulation, we will find identical values of α multiple times. However, this should not be problematic for this application, since the new detected position is sufficiently randomized and the exact distribution of α is not needed later. In Figure 37 we find the distribution of α for all energies and again determine 68 % quantile of this all-energy distribution for a later application. After giving each event an angular distance, we simply determine the new, detected position $\vec{r}_{\text{det}}$ by randomly picking a position uniformly distributed on a circle around the true position $\vec{r}_{\text{true}}$ with radius α (s. Figure 38). The aforementioned procedure only accounts for statistical fluctuations, whereas systematic effects, like the deflection of the primary particle due to the geomagnetic field, are not taken into consideration. This is an information determined by CORSIKA, which we lose by employing our custom simulation. We will further discuss this when determining the deflection of the particles in the geomagnetic field.

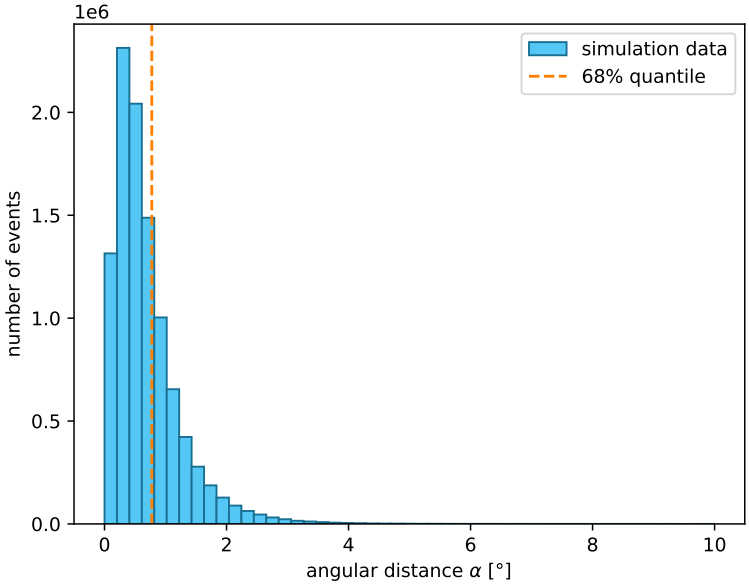

Figure 37: The angular distance of our simulated events. There are again singular values larger than seen here, but the shown data is limited to $\alpha \leqslant 10°$. The 68 % quantile is marked for a later application.

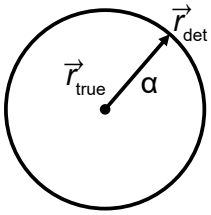

Figure 38: Visualization of the simulation of the detected direction $\vec{r}_{\text{det}}$. The position $\vec{r}_{\text{det}}$ is randomly chosen on a circle of radius α around the true position $\vec{r}_{\text{true}}$.

Since we have now determined all important characteristics of our cosmic ray events we need so far, we can finally take a first look at the moon shadow we expect to observe with SWGO.

4 First Look at the Moon Shadow

Before continuing with the more complicated consideration of the influences caused by the geomagnetic field, we want to evaluate the observable moon shadow in the data as simulated so far. We will produce counts, sky and subtracted maps to determine the significance at which the shadow can be observed and to quantify the angular resolution we can determine from evaluating the shadow.

4.1 Observation of Particles in Counts Maps

To observe the moon shadow in this undeflected data set, we first need to remove all events that would be covered by a disc with 0.5° in diameter centered at the moon position in the true positions of events:

$$r = \sqrt{\text{lon}^2 + \text{lat}^2} \leqslant r_{\text{moon}} = 0.25^\circ \tag{32}$$

since the angular diameter of the moon varies between $0.49^\circ - 0.57^\circ$. The result of removing these events can be seen in Figure 39.

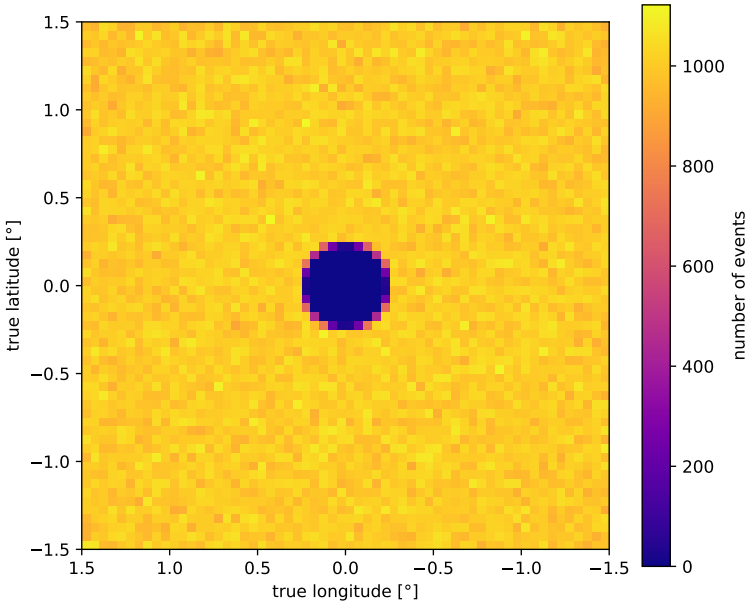

Figure 39: The all-energy moon shadow in a counts map of true positions. The moon can be seen as a sharp, circular shadow, since there is no blurring by the difference between true and detected position determined from the angular resolution.

We remove the events covered by the moon from the entire data set to find the detected positions of the events in the counts map in Figure 39. As a result, we plot our first counts map of the detected moon shadow. This plot is shown in Figure 40 for a simulation corresponding to an observation time of 100 hours. We use a standard binning of 20 bins/°, which corresponds to a bin size of 0.05°, which is a bit smaller than the best determined angular resolution. A slight moon shadow in the center of the plot can be seen. To further improve the visibility and determine the significance with which we

observe the moon shadow, we now want to take a look at the moon shadow on a sky map.

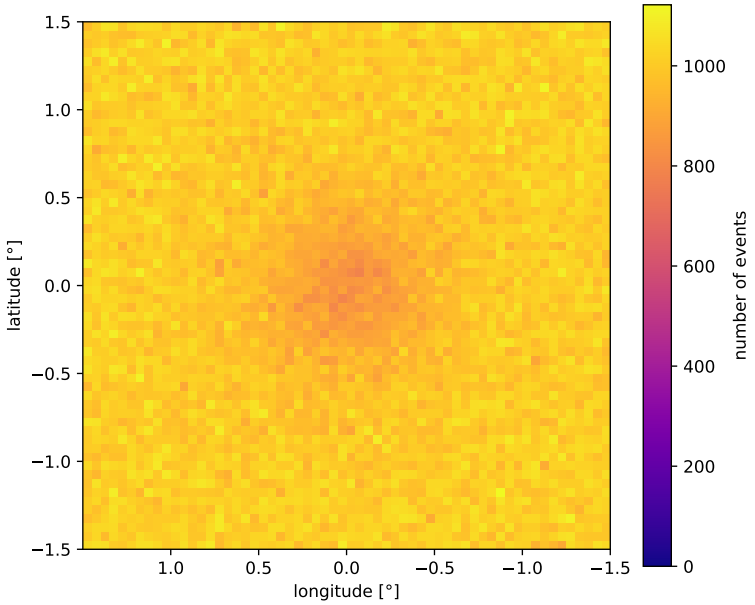

Figure 40: An all-energy counts map of the moon shadow for a 100 h observation at a fixed zenith angle of 14.5°. The color scale ranges from zero to the maximum number of detected events in one bin.

4.2 Evaluation of the Significance via Sky Maps

To produce a significance map of the moon shadow, we use the formula established by Li and Ma [22, p. 320, Eq. 17] with N_{on} corresponding to our shadow data and N_{off} being a set of background data. In this formula, we add the sign of the difference $N_{\text{on}} - \alpha N_{\text{off}}$ as a coefficient, since we want to associate a deficit in events with a negative significance. For the computation of the significance, we simply create another independent data set equivalent to the same observation time, this time without removing the moon shadow. After applying the Li and Ma Formula, we employ a circular kernelling with a diameter of $\frac{1}{2}\sigma_{68} = 0.39°$, using the 68 % quantile for all energies determined in subsubsection 3.2.4. The resulting sky map for a 100 h observation can be found in Figure 41. Now there is a prominent visible shadow centered on the moon coordinates $(0, 0)$.
To determine the dependency of the observable shadow on the observation time, we split this data set into 10 hour increments. For each increment, we cumulate the data-subsets up to that observation time and plot sky map for each cumulated subset. The result can be seen in Figure 42. As one might expect, the moon shadow becomes more prominent for longer observation times, resulting in a larger value of σ. To further examine this behavior, we determine the maximum significance of each sky map and plot the dependency of this significance on the observation time, expecting a relation of $\sim \sqrt{t}$. In Figure 43, we therefore employ a fit of the function

$$f(t) = a \cdot t^c + b \tag{33}$$

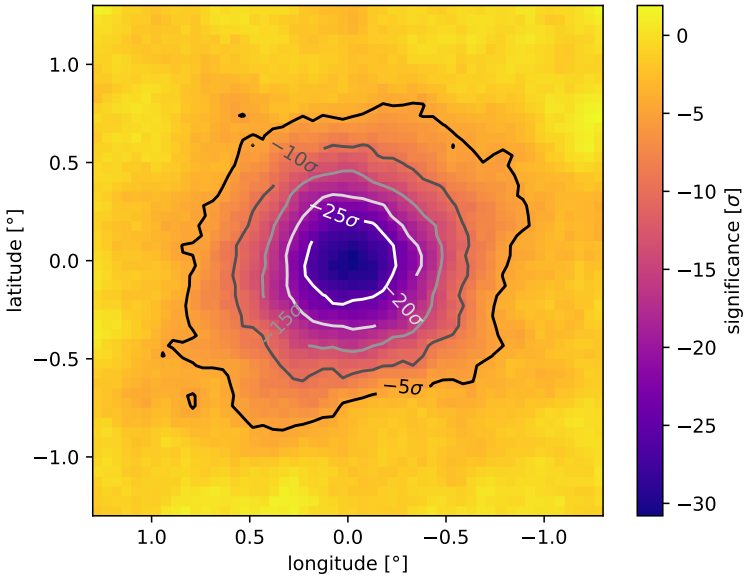

Figure 41: An all-energy sky map of the moon shadow for a 100 h observation. The contour lines are meant to visualize areas of the same significance level.

to verify this relation, expecting $c = \frac{1}{2}$. For the exponent, we find an optimal value of

$$c_{\text{opt}} = 0.420 \pm 0.019 \tag{34}$$

where the error is determined from the covariance matrix. This does not entirely match our expectation, but is still close to $c = 0.5$.
Lastly, we want to examine the dependency of the moon shadow on the energy of primary particles. We employ the established energy binning and receive the resulting significance maps in Figure 44. Generally, a higher energy is connected to a better visibility, due to the better angular resolution. However, this is limited by the smaller amount of events in the higher energy bins, since the steep $E^{-2.7}$ spectrum is dominated by lower energies.
To further quantify our observations for later comparison with the deflected data, we now want to take a look at the subtracted maps available by our simulation.

4.3 Quantifying the Shadow with Subtracted Maps

A subtracted map is a count map that shows only the events we would normally remove in order to observe the moon shadow. It is very important to note that these subtracted maps do not correspond to a physical observation possible by an instrument and is therefore only available to us in simulated data. Later on, one would have to subtract a separate observed background measurement and perform a binned maximum likelihood fit to determine the position and size of the shadow, taking into account fluctuations in the background data. This is a difficult process, which might not lead to the desired measurement outcome, especially for energy bins with small amounts of data. The advantage of evaluating these subtracted maps in simulation data is that we can easily quantify the position and size of the moon shadow without the need for a fitting routine,

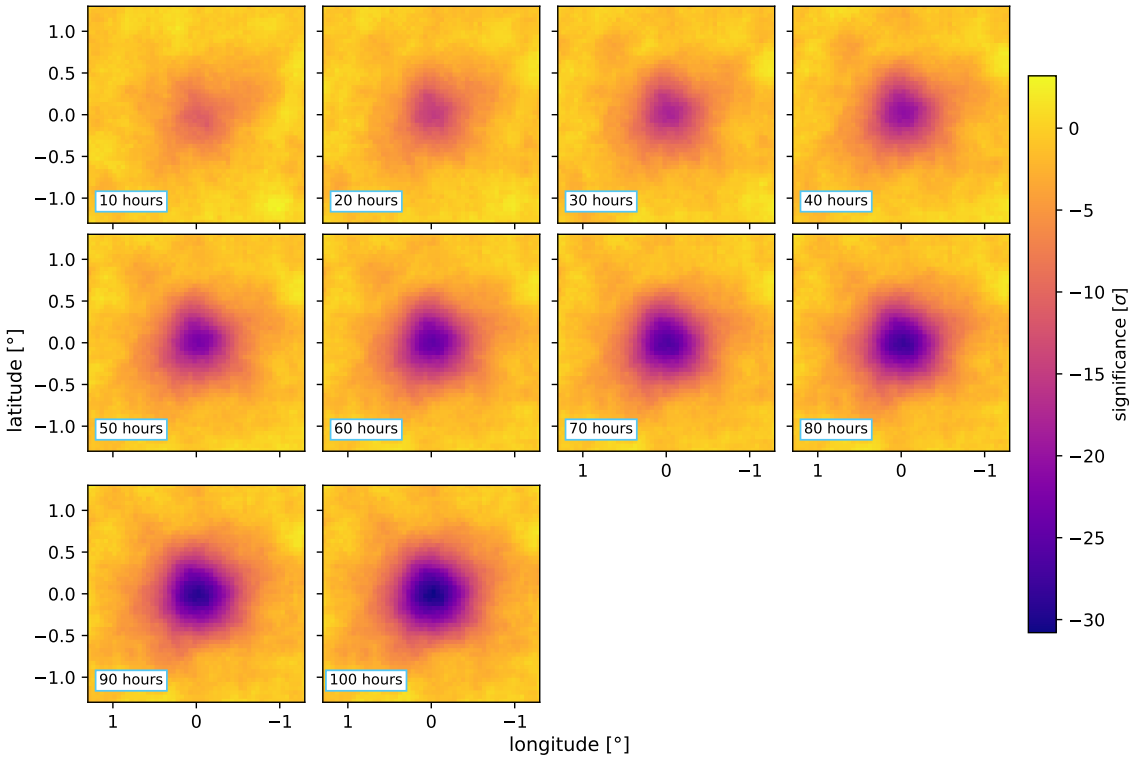

Figure 42: The all-energy sky maps for 10 h increments. The same color scale was applied to all maps.

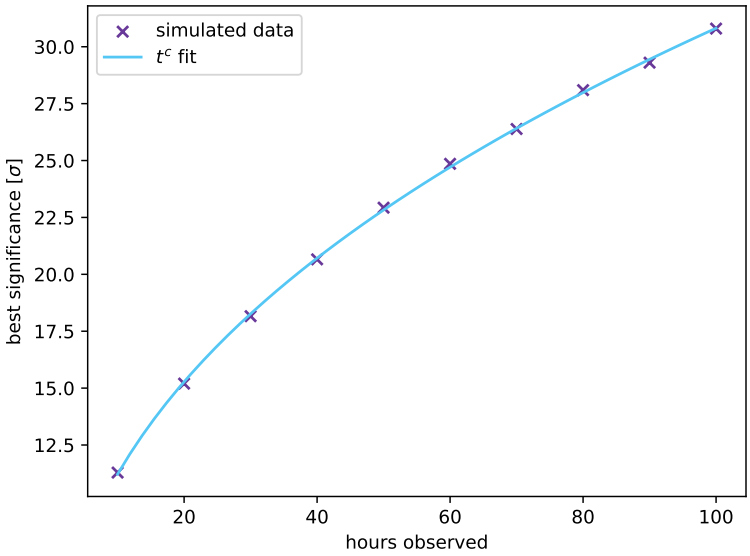

Figure 43: The time dependency of the maximum significance of each time increment from Figure 42. A t^c fit was performed to show the underlying dependency.

which is especially useful for the deflected data we will simulate later, when the observed shadow is not symmetric anymore.

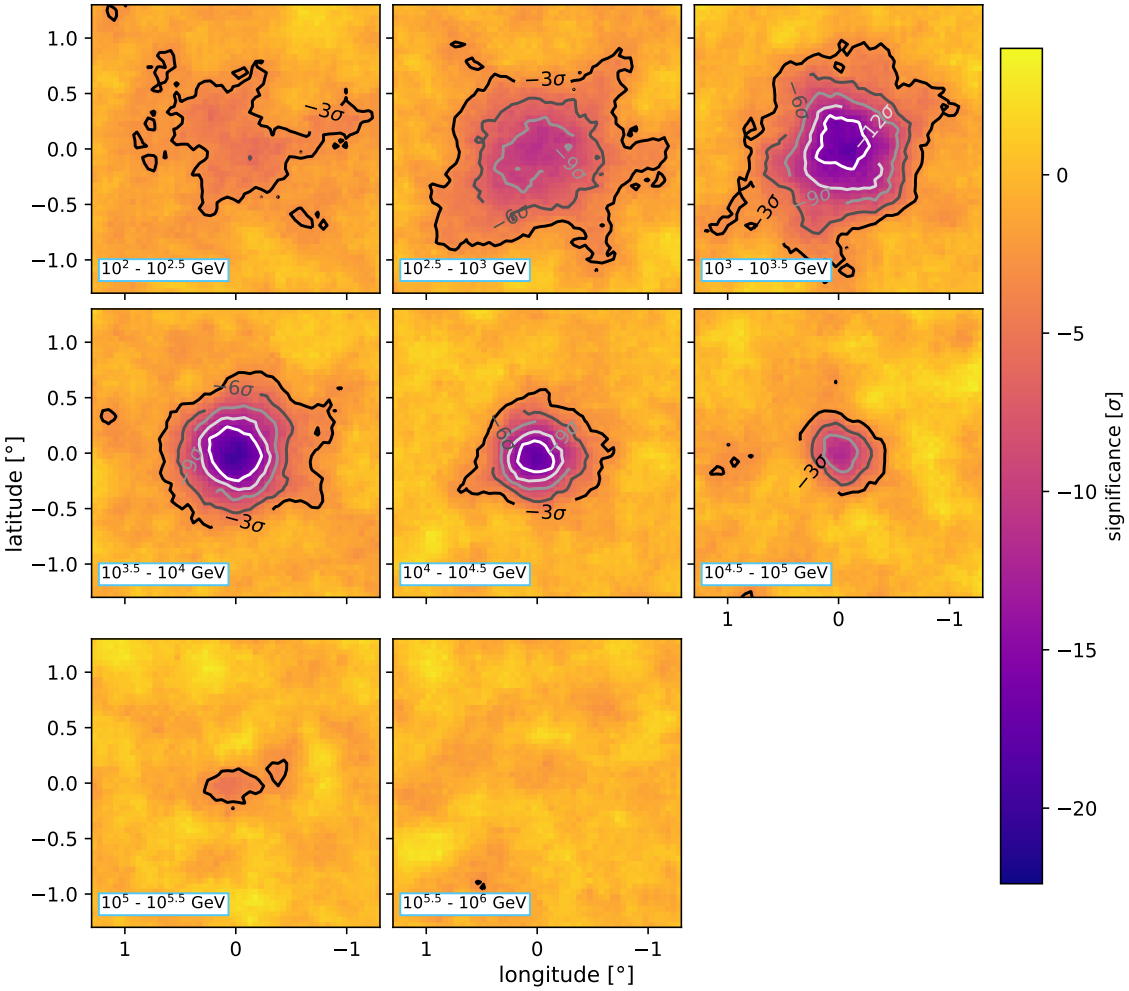

Figure 44: Sky maps for a 100 h observation, binned in primary energy. The same color scale was applied to all maps. The contour lines are meant to visualize areas of the same significance level.

When plotting only the events covered by the moon, we receive the subtracted map in Figure 45. To determine the center of gravity (CoG) of this map, we compute the mean of the longitude and latitude values of the subtracted events, making the considered window as large as our data permits, to avoid edge effects. For the all-energy map we find

$$\Delta\text{lon} = -0.002^\circ \pm 0.003^\circ \quad \text{and} \quad \Delta\text{lat} = -0.003^\circ \pm 0.003^\circ \tag{35}$$

With the error yielding from the standard error of the arithmetic mean estimated by the sample standard deviation:

$$\Delta(\Delta x) = \frac{\sigma_x}{\sqrt{n}} \tag{36}$$

where n is the number of events in the data set the mean was determined from. The result is in accordance to our expectation of the CoG being at $\approx (0^\circ, 0^\circ)$, since the moon is centered on the shown window and we have not yet introduced the deflection of the primary particles. Similarly, we compute the Root Mean Square (RMS) of

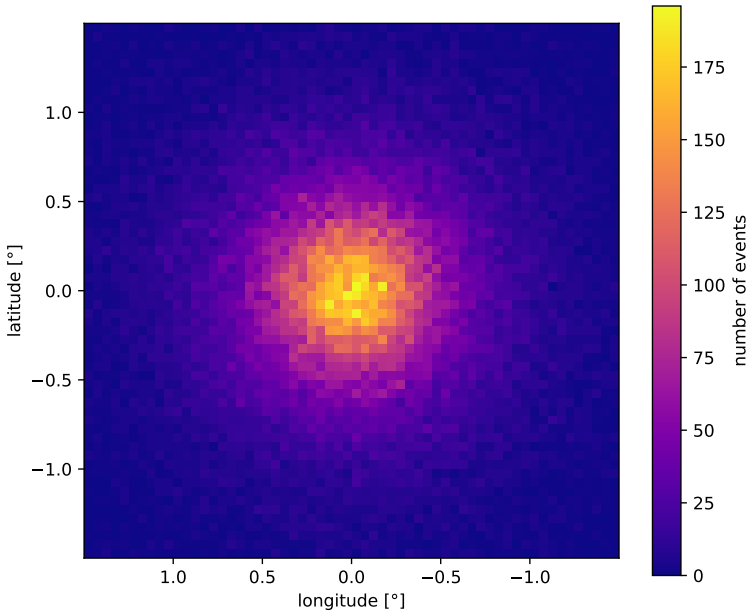

Figure 45: An all-energy subtracted map of the moon shadow for a 100 h observation.

the distribution by computing the square root of the corrected sample variance S^2 of longitude and latitude values. We receive

$$\sigma_{\text{lon}} = 0.6173^\circ \pm 0.0016^\circ \quad \text{and} \quad \sigma_{\text{lat}} = 0.6173^\circ \pm 0.0016^\circ \tag{37}$$

The errors are estimated by (s. derivation in appendix A)

$$\Delta\sigma = \frac{\sigma}{\sqrt{2(n-1)}} \tag{38}$$

As expected, we receive the same value of σ for longitude and latitude, since the undeflected moon shadow is expected to be symmetric. The total RMS for both directions is then

$$\sigma_{\text{tot}} = \sqrt{\sigma_{\text{lon}}^2 + \sigma_{\text{lat}}^2} = 0.8730^\circ \pm 0.0016^\circ \tag{39}$$

with the error determined from error propagation:

$$\Delta\sigma_{\text{tot}} = \frac{1}{\sigma_{\text{tot}}}\sqrt{\sigma_{\text{lon}}^2\Delta\sigma_{\text{lon}}^2 + \sigma_{\text{lat}}^2\Delta\sigma_{\text{lat}}^2} \tag{40}$$

We can use this total RMS as a cross check of our simulations so far, by splitting it up into its components:

$$\sigma_{\text{tot}}^2 = \sigma_{\text{moon}}^2 + \sigma_{\text{angres}}^2 \tag{41}$$

where σ_{moon} is the contribution of the moon disc and σ_{angres} is the contribution by our determined angular resolution. For a uniform distribution on a circular disc, we find an RMS of (s. derivation in appendix A):

$$\sigma_{\text{moon}} = \frac{r}{\sqrt{2}} = \frac{0.25^\circ}{\sqrt{2}} \approx 0.1768^\circ \tag{42}$$

After rearranging Equation 41 and inserting σ_{moon} and σ_{tot}, we receive

$$\sigma_{\text{angres}} = \sqrt{\sigma_{\text{tot}}^2 - \sigma_{\text{moon}}^2} = 0.8550^\circ \pm 0.0016^\circ \tag{43}$$

where the propagated error of σ_{angres} is:

$$\Delta\sigma_{\text{angres}} = \frac{\sigma_{\text{tot}}}{\sigma_{\text{angres}}} \Delta\sigma_{\text{tot}} \tag{44}$$

The value of σ_{angres} is similar to the 68 % quantile of the angular distance we input into the shadow simulation, $\sigma_{68} = 0.7739^\circ$. We have therefore successfully utilized our simulated observation of the moon shadow to estimate the angular resolution we have input into the simulation.

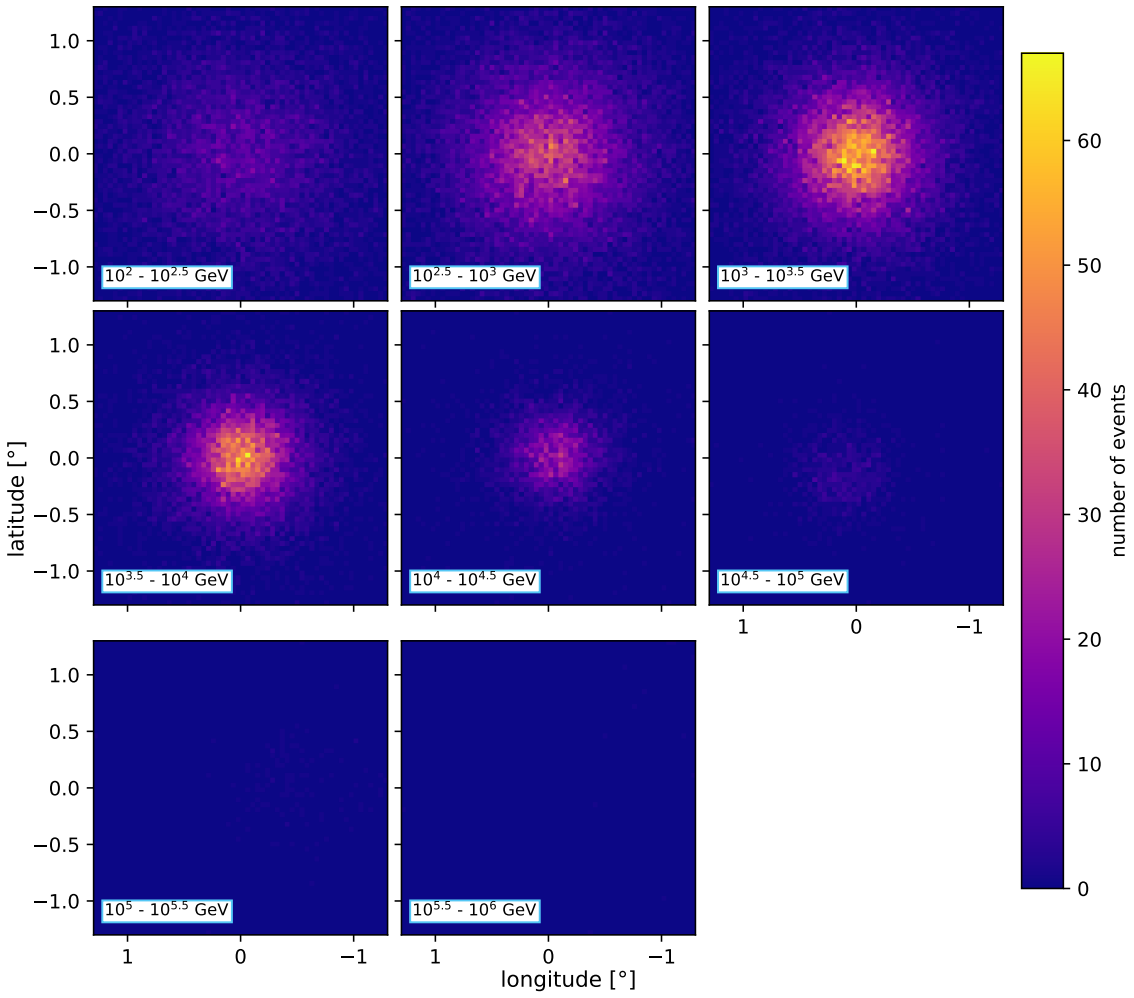

Figure 46: Subtracted maps for a 100 h observation, binned in primary energy. The same color scale was applied to all maps.

We also want to examine the energy dependent appearance of these Subtracted maps and apply the same binning in primary energy as before. The result can be seen in Figure 46. As before, the visibility improves with higher energies, being restricted by

the low event number in higher energy bins. We apply the same method to determine the Center of Gravity and the Root Mean Square in longitude and latitude for each of these bins for later comparison with the deflected shadow. The values can be found in Table 2.

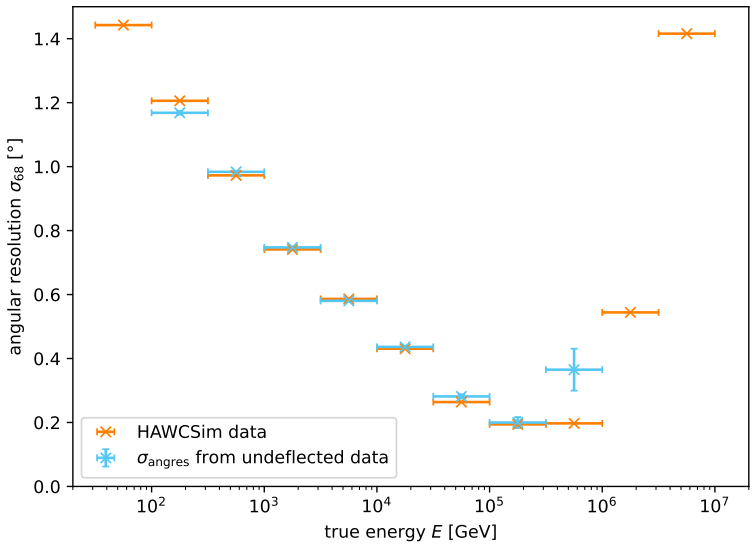

Figure 47: The values of σ_{angres} as determined from the subtracted moon shadow in comparison with the angular resolution determined from HAWCSim. The number of data points for the shadow data is smaller than that of HAWCSim, because the energy range we have chosen for our custom simulation is smaller.

We again compute σ_{angres} for each of these energies (s. Table 5 in the appendix) and compare them to the angular resolution determined from HAWCSim data. As seen in Figure 47, the result is in accordance with the expected resolution. The value of the first energy bin is slightly smaller than expected. This likely yields from the size of the plot window not being able to capture all events that belong to the shadow, since the angular uncertainties in this energy bin are rather large. The last measured point also has a strong deviation, which yields from the very small amount of events in that map ($n = 13$), indicated by the large error bar.
It is important to mention that the undeflected shadow we have simulated in this section is obviously not observable in real data, since all particles will be deflected before reaching the detector. The simulation of an undeflected shadow merely serves as a possibility to take a first look at the moon shadow and to compare to the shadow we will observe in deflected events in the following.

5 Deflection in the Geomagnetic Field

5.1 Properties of the Geomagnetic Field

To determine the deflection of cosmic rays in the Earth's magnetic field, we first want to familiarize ourselves with the field's properties. The geomagnetic field can be approximated by a magnetic dipole field, with the direction of the field lines pointing from the geographic South to the geographic North pole, which is illustrated in Figure 48. A more exact estimate is provided by the International Geomagnetic Reference Field (IGRF) [20], which employs a series expansion with coefficients being updated in a five year cycle:

$$V(r,\theta,\varphi,t) = a\sum_{n=1}^{N}\sum_{m=0}^{n}\left(\frac{a}{r}\right)^{n+1}\left[g_n^m(t)\cos(m\varphi)+h_n^m(t)\sin(m\varphi)\right]P_n^m(\cos\theta) \tag{45}$$

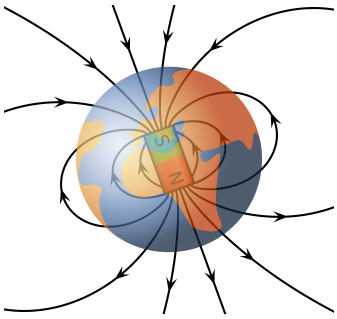

Figure 48: Illustration of the approximation of the geomagnetic field as a dipole. The graphic was taken from [25].

Using the Pure Python code provided for the IGRF, we receive the magnetic field vector of any location on and above earth in its components B_r, B_φ, B_θ in geocentric coordinates. We can then compute the field strength as

$$||\vec{B}|| = \sqrt{B_r^2 + B_\varphi^2 + B_\theta^2} \tag{46}$$

To get an idea of the strength of the magnetic field at the SWGO site, we show its height dependency in Figure 49. It is notable that the strength of the field remains fairly constant for the first $\sim 1000\,\text{km}$ before dropping off. In the further distance, the curve seems to follow a power law above a height of $\sim 10\,000\,\text{km}$, where the geomagnetic field seems to behave more like a dipole, with a dependency on distance of about $\sim \frac{1}{r^3}$.
Since we will later want to compare our simulated deflection of the moon shadow with literature from HAWC, we also want to compare the properties of the geomagnetic field at both sites. The magnetic field strength at the ground level of both sides is

$$B_{\text{SWGO}} = 22.3\,\mu\text{T} \tag{47}$$

$$B_{\text{HAWC}} = 39.3\,\mu\text{T} \tag{48}$$

This makes the geomagnetic field about 1.8 times stronger at the HAWC location than at the prospective SWGO site. The declination δ, which is defined as the angle between

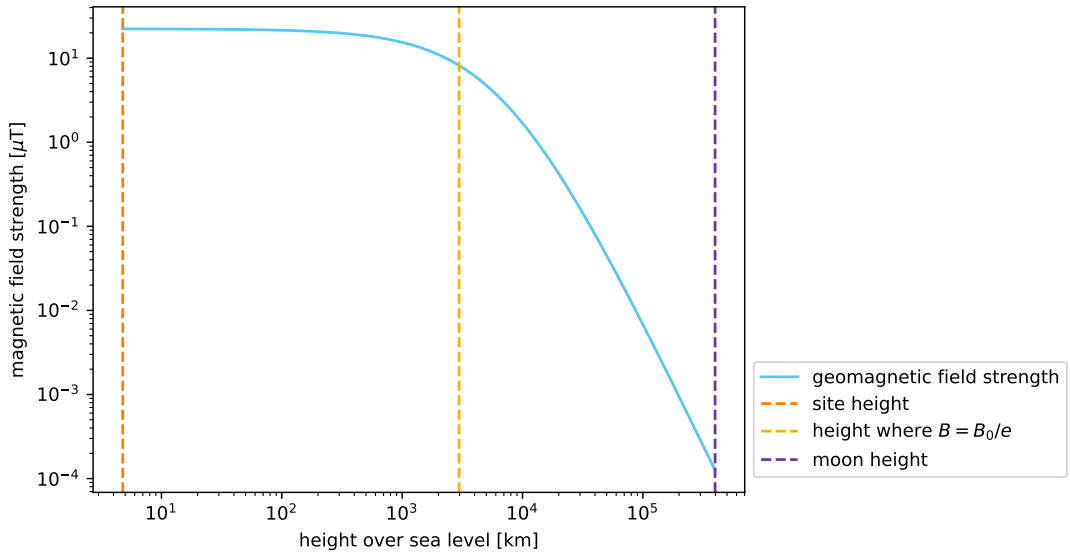

Figure 49: The magnetic field strength at the SWGO site for different heights. The values were determined via the IGRF coefficients [20]. The vertical lines represent relevant heights for illustration.

the magnetic field lines at a given location and the direction of the magnetic North pole, also differs at both observatories:

$$\delta_{\text{SWGO}} = -7.99^\circ \tag{49}$$

$$\delta_{\text{HAWC}} = 3.14^\circ \tag{50}$$

It is especially notable, that the magnetic field lines point in different directions, as shown by the different sign in their declination. This will be relevant later, when we examine the deflection of the moon shadow and see the consequences of this difference.

5.2 Event Height

To observe the moon shadow of deflected cosmic rays, we need to propagate the primaries back through the magnetic field from the point of observation. In order to do this, we need to decide when to start this propagation, effectively deciding on what is the starting height of each event. Two possible options include using the ground height or the height of the first interaction, as we have determined previously from X_{max} (s. subsubsection 3.1.5).
As mentioned before, our custom simulation does not take into account in which direction the detected position $\vec{r}_{\text{det}}$ is shifted compared to the true position $\vec{r}_{\text{true}}$. This also means, that even though CORSIKA takes the interaction of secondary and primary particles with the geomagnetic field into account, we do not use this information. If we were, we could simply start tracking the particles' trajectories above their creation height in the CORSIKA simulation, since below this height the deflection is already taken into account for. But since this is not possible with the way we have structured the custom simulation, we need to decide on a different starting point.
The main difference between the two mentioned options for the starting height can be described as follows: For the first option, we would assume that the shower as a

whole has approximately the same trajectory in the geomagnetic field as the primary particle would. Then we would need to compute the deflection starting at the height of the detector, since then the trajectory would be the same independent of whether the primary particle interacts with the atmosphere to trigger a cascade. For the other option of starting at the height of the first interaction, we would assume that the shower as a whole is not deflected by the geomagnetic field and therefore the shower core moves in a straight line as soon as the cascade is triggered. This could make sense, if we take into account that the shower consists of particles with different charges of different signs and the deflection of a single particle is therefore not very relevant to the average shower movement.
As far we could tell during research, this problem has unfortunately not been extensively discussed in literature, which is why we were not able to find a reliable answer on how the shower behavior can be best approximated in the magnetic field. Ultimately, we decide to use the height of first interaction [7, p. 18, Fig. 12], since we will see in subsection 5.4 that the influence of the starting height on the overall deflection is only minuscule. This means, that we only take into account the deflection in the geomagnetic field of primary particles, starting the propagation at the height of first interaction. Therefore, we again sample this height from the distribution shown in Figure 34 according to the primary event energy, similarly to our method for sampling the angular distance in subsubsection 3.2.4. It is important to note that no matter which of these described options we employ, they all describe only an approximation of the actual shower behavior, so we simply try to decide on the most sensible estimate for this use case. However, the behavior of a hadronic shower in the geomagnetic field remains an interesting field to look into for future considerations.

5.3 Coordinate Transformation

Since our events are being observed in horizontal coordinates and the geomagnetic field is given in geocentric coordinates, we need to find a transformation between the two to compute the trajectory of observed particles in the field. Since this transformation is non-trivial, we want to describe it in its consisting steps. A graphic describing the entire transformation from the horizontal to the geocentric system can be found in Figure 50.

1. The events in their spherical coordinate systems are transformed to left-handed a cartesian coordinate system, as described in subsection 2.6. Then the coordinate system is rotated with around the longitude angle φ around an axis that connects the North and South poles of the earth, which will later be the z-axis of our geocentric coordinate system. This results in a coordinate system located at $\text{lon} = 0°$.

2. This system is then rotated around an axis perpendicular to the pole axis and the vector connecting the current coordinate origin to the center of the earth. A rotation around the latitude angle θ yields a coordinate system located at the North pole, with the z- and y-axes already oriented as desired for the geocentric system.

3. To address the left- and right-handedness of the two system, the x-axis is now inverted, resulting in a coordinate system with all axes being correctly oriented.

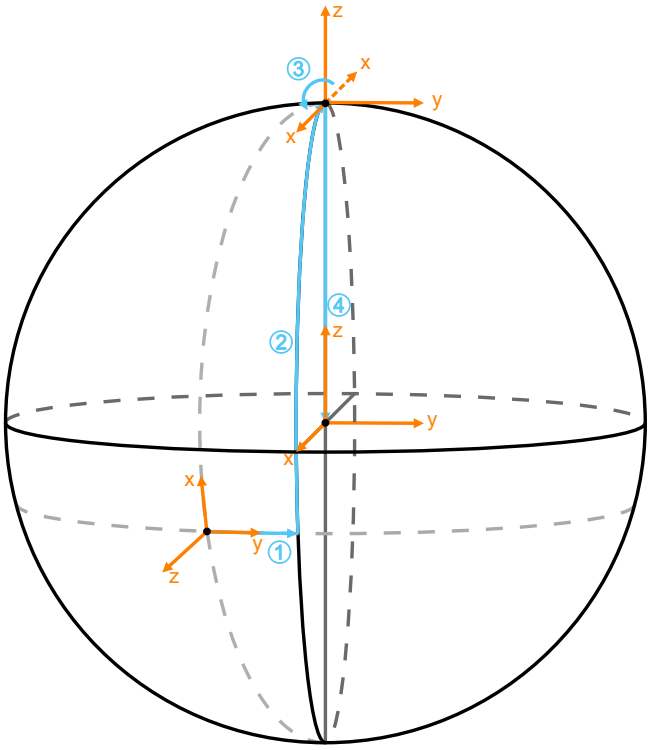

Figure 50: A schematic graphic describing the transformation from a horizontal to a geocentric coordinate system. The orange lines show the different coordinate systems, while the blue lines show the direction in which the systems are moved or rotated. The numbers mark the different steps of the transformation as they are being described in subsection 5.3.

4. Lastly, the offset of the coordinate system is removed, to transport the coordinate origin from the North pole to the center of the earth.

After this transformation, we receive events in a cartesian geocentric coordinate system, in which we can easily transform the spherical geomagnetic field vector by a standard transformation from spherical to cartesian coordinates.

5.4 Differential Equation & Trajectory

The cosmic ray primaries are deflected in the geomagnetic field via a Lorentz force

$$\vec{F}_L = q(\vec{v} \times \vec{B}) \tag{51}$$

with the particle's charge q, its velocity $\vec{v} = \dot{\vec{x}}$ and the magnetic field $\vec{B}$. Its trajectory is therefore described by the following differential equation:

$$m\ddot{\vec{x}} = q\left(\dot{\vec{x}} \times \vec{B}(\vec{x})\right) \tag{52}$$

This is a system of second order ordinary differential equations (ODEs) which can be rewritten as a system of first order differential equations. By defining

$$\vec{y}_1 = \vec{x} \quad \text{and} \quad \vec{y}_2 = \dot{\vec{x}} \tag{53}$$

we find

$$\dot{\vec{y}}_1 = \dot{\vec{x}} = \vec{y}_2 \tag{54}$$

$$\dot{\vec{y}}_2 = \ddot{\vec{x}} = \frac{q}{m}\left(\vec{y}_2 \times \vec{B}\right) \tag{55}$$

which can be easily solved by an ODE solver. We first employ a simplistic solving approach, using a self written Euler method to solve these ODEs, but find the accuracy of this first-order method to be insufficient. This is why we decide to utilize the function `solve_ivp` from the `scipy` library [12], which uses a fourth-order Runge-Kutta method, computing our trajectory with a sufficient accuracy. For the initial values, we give each event a starting location $\vec{x}_0$ given by its height and point of observation and a starting velocity $\vec{v}_0 = c \cdot \sqrt{1 - \frac{E_0}{E_0 + E_{\text{kin}}}}$ with the resting energy $E_0 = m_p c^2$ and the observed kinetic energy E_{kin}. Furthermore, the particles are given a charge $q = -e$. The sign of the charge and velocity are opposite of their true values, since we need to back-propagate the particle, making its trajectory identical to that of an anti-proton moving outward from the earth.

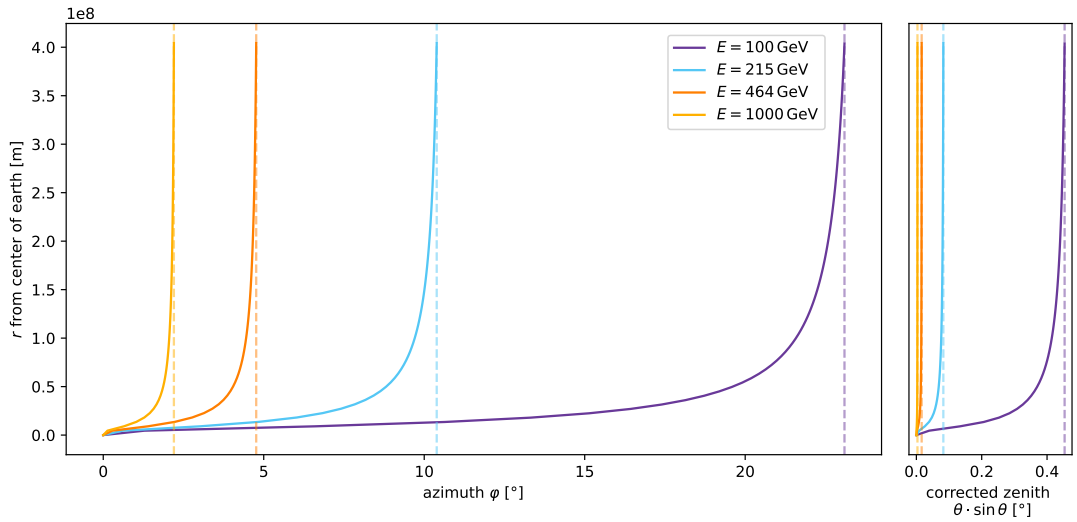

Figure 51: The trajectory of particles of different energies reaching in the geomagnetic field. The particles are all detected at a zenith angle of $\theta = 0°$ and back-propagated until reaching moon height. The dashed lines show the coordinates at which the particles would be detected without geomagnetic field deflection. The scale of the corrected zenith is magnified by ten in comparison to the azimuth, since the deflection in zenith is much smaller.

To familiarize ourselves with the deflection of these particles in the geomagnetic field at the SWGO site, we first want to take a look at the trajectory of a few single particles of different energies reaching our detector at a zenith angle of $\theta = 0°$, starting the computation at a $h = 0\,\text{km}$. As seen in Figure 51, the largest deflection can be found at the lowest energy. It is notable, that the deflection in azimuth angle is much larger than in the zenith angle, and the main deflection occurs closest to the ground, since this is where the magnetic field is strongest. We compute the total deflection as the angle between the arrival direction with an without the geomagnetic field and compare these

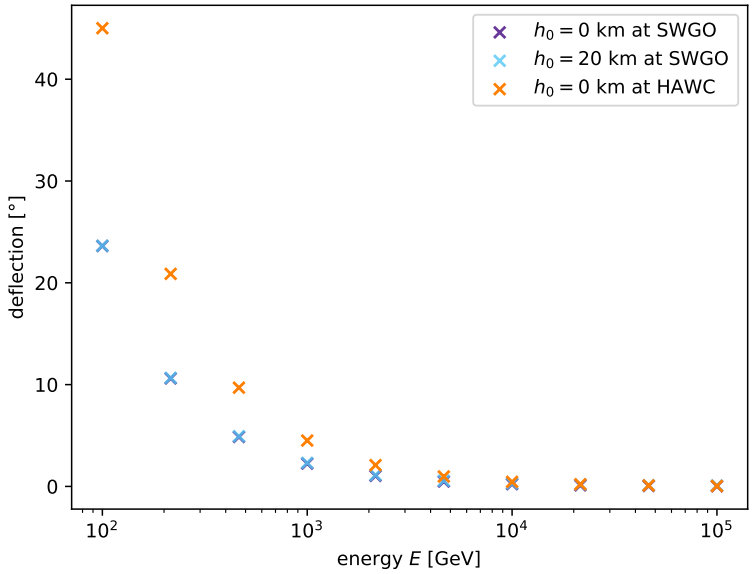

Figure 52: The total deflection at different energies for different starting heights at the SWGO site and for $h = 0\,\text{km}$ at the HAWC site.

values for a starting height of $h = 0\,\text{km}$ and a typical value of $h = 20\,\text{km}$. The comparison can be seen in Figure 52 and we find the difference to be very small, justifying our decision of using the height of the first interaction, since the influence on the deflection is only minuscule. We also compare these values to the deflection in the geomagnetic field found at the HAWC site, to estimate the comparability of the total deflection of the shadow at both sites. We find a large difference at low energies, which decreases for higher energy particles. We therefore should not see a large difference in deflection at higher energies.
We back-propagate all events in our observed window up to the height of the moon and save the longitude and latitude of the last evaluation point to determine the moon shadow as seen after deflection in the geomagnetic field.

5.5 The Shifted Moon Shadow

5.5.1 Sky Maps

To observe the moon shadow in the deflected data, a similar procedure as before is used. The events, which are blocked by a disc of 0.25° in radius around the moon position are removed, but now the deflected true positions are used, effectively shifting the observable shadow seen in the detected position data. We again extract a LiMa significance map from the data, this time using a `SkyOffsetFrame` with its origin as the moon position in equatorial coordinates, resulting in the longitude and latitude axes pointing in the direction of declination and right ascension respectively. For all energies we find the result in Figure 53. We see a deflection of the shadow in the negative direction for longitude and latitude, where the deflection in the direction corresponding to right ascension seems to be stronger than in declination. We will quantify this effect later, when regarding the subtracted maps (s. subsubsection 5.5.2). The shadow has an

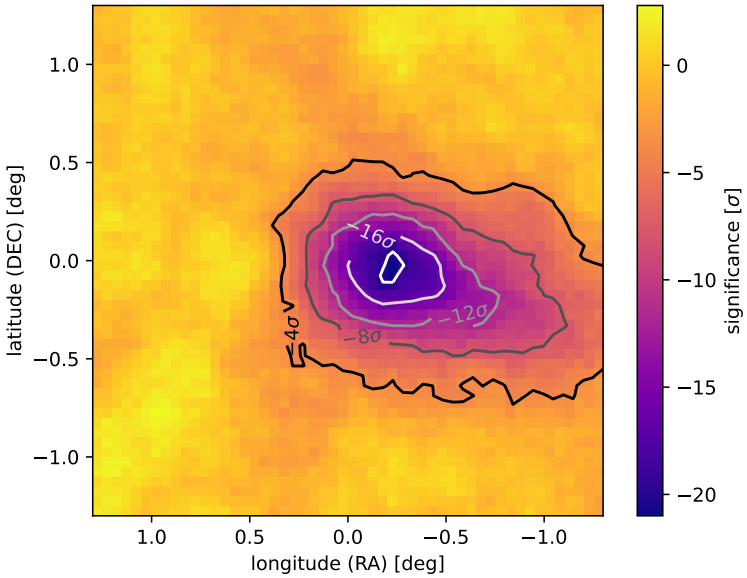

Figure 53: An all-energy sky map of the moon shadow for a 100 h observation of the moon, with particles being deflected in the geomagnetic field. The contour lines are meant to visualize areas of the same significance level.

asymmetric appearance, being drawn out to the right. This is because the portrayed energy range is very large, with deflection being strongly dependent on primary particle energy, as seen in subsection 5.4. Since lower energy particles are deflected much farther than higher energy particles, the all-energy sky map depicts the combined appearance of the shadow observed in the entire detected energy spectrum, resulting in an elongated shadow.
In Figure 54 we again split the simulation data into energy bins and look at the deflection in each bin. It is observable that the deflection is larger for bins of lower energy, as expected by the results in Figure 51. For higher energies, the position of the moon shadow moves toward $(0°, 0°)$, before visibility again fades due to the low number of events in the highest energy bins. To further quantify these findings, we again want to take a look a the subtracted maps. Interestingly, the elongation found in the all-energy sky map is no longer observable, resulting in a nearly symmetric shadow in the energy bins. This is because the energy range in each section is now small enough so that all particles are deflected approximately the same.

5.5.2 Subtracted Maps

The depiction of only those events covered by the moon disc for all energies yields Figure 55. We determine the depicted errors with the same formulas shown in subsection 4.3. The difference between the CoG at

$$\Delta\text{lon} = -0.897° \pm 0.004° \quad \text{and} \quad \Delta\text{lon} = -0.236° \pm 0.003° \tag{56}$$

and the peak of the distribution at

$$\text{lon}_{\text{max}} = -0.226° \quad \text{and} \quad \text{lat}_{\text{max}} = -0.075° \tag{57}$$

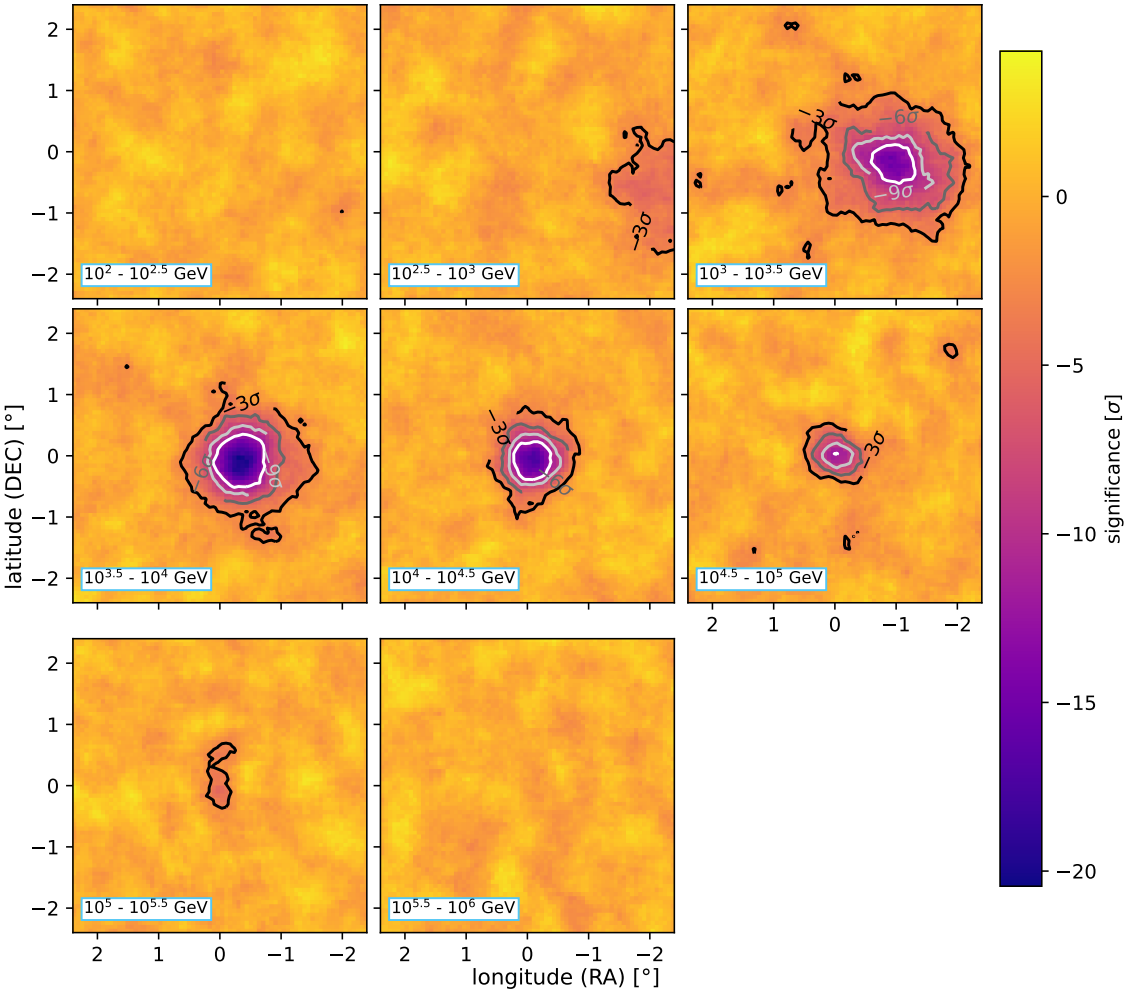

Figure 54: Sky maps for a 100 h observation of the moon, with particles being deflected in the geomagnetic field. The same color scale was applied to all maps. The contour lines are meant to visualize areas of the same significance level.

again verifies the asymmetry of this distribution. This is however only due to the large energy range depicted in this plot, since particles of different energies are deflected differently in the geomagnetic field. This is why, for an all-energy map, we can expand Equation 41 as follows

$$\sigma_{\text{tot}}^2 = \sigma_{\text{moon}}^2 + \sigma_{\text{angres}}^2 + \sigma_{\text{magnet}}^2 \tag{58}$$

since the geomagnetic field also contributes to the widening of the shadow distribution. If we assume the same σ_{angres} as found in the undeflected data and rearrange the equation as before, we find

$$\sigma_{\text{magnet}} = \sqrt{\sigma_{\text{tot}}^2 - \sigma_{\text{moon}}^2 - \sigma_{\text{angres}}^2} = 0.588^\circ \pm 0.005^\circ \tag{59}$$

where the error $\Delta\sigma_{\text{magnet}}$ is again determined by Gaussian error propagation:

$$\Delta\sigma_{\text{magnet}} = \sqrt{\left(\frac{\sigma_{\text{tot}}}{\sigma_{\text{magnet}}}\right)^2 \Delta\sigma_{\text{tot}}^2 + \left(-\frac{\sigma_{\text{angres}}}{\sigma_{\text{magnet}}}\right)^2 \Delta\sigma_{\text{angres}}^2} \tag{60}$$

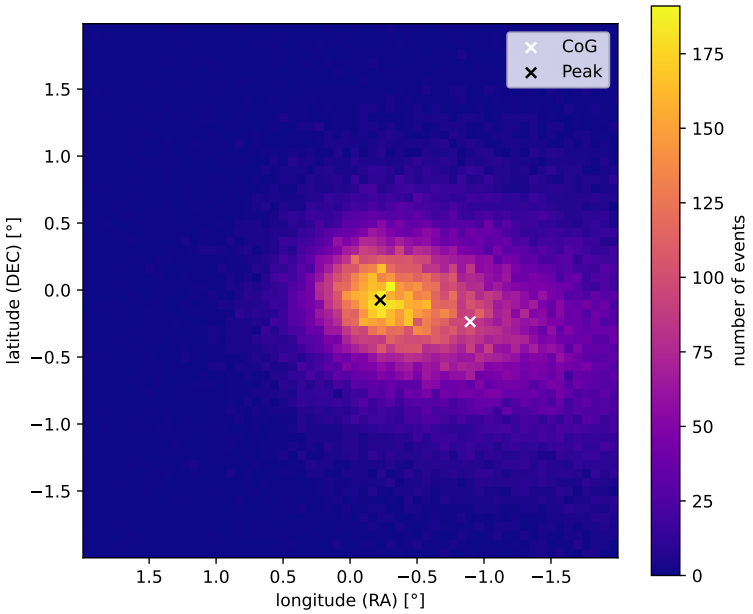

Figure 55: An all-energy subtracted map of the moon shadow for a 100 h observation, with particles being deflected in the geomagnetic field. The markers depict the CoG and the Peak of the distribution in longitude and latitude.

To compare the findings of the deflected and undeflected shadow evaluation, in Table 1 we have summarized the respective results in position and extent of the moon shadow.

Dataset	$\mathbf{\Delta lon}$ [°]	$\mathbf{\Delta lat}$ [°]	$\boldsymbol{\sigma_{lon}}$ [°]	$\boldsymbol{\sigma_{lat}}$ [°]
undeflected	-0.002±0.003	-0.003±0.003	0.6173±0.0016	0.6173±0.0016
deflected	-0.897±0.004	-0.236±0.003	0.875±0.003	0.5856±0.0017

Table 1: The overall result of the position and extent of the all-energy moon shadow for simulations with and without the geomagnetic field. In the deflected data, the longitude and latitude are equivalent to the right ascension and declination. The extent of the moon shadow is given by the RMS in the examined data set and direction.

For the energy-binned subtracted maps (s. Figure 56), we will not regard the contribution of σ_{magnet}, since the energy bins should be small enough to prevent a widening caused by the magnetic field, allowing us to use Equation 41 as before. For the CoG and the RMS we find the values in Table 3 in the appendix A.
We include the new values of σ_{angres} in the energy dependent angular resolution plot and find Figure 57. It can be seen that the angular resolution of both the undeflected and the deflected shadow are in accordance with the angular resolution from HAWCSim data, although σ_{angres} in the deflected data in the lower energy bins is slightly larger than in the undeflected. It is important to mention, that especially in the first two energy bins, the evaluated window was not big enough to portray the entire shadow, which can impact our result, as seen in the exceptionally large error bar in the first energy bin.
Lastly, we want to compare our findings regarding the deflection of the Center of Gravity

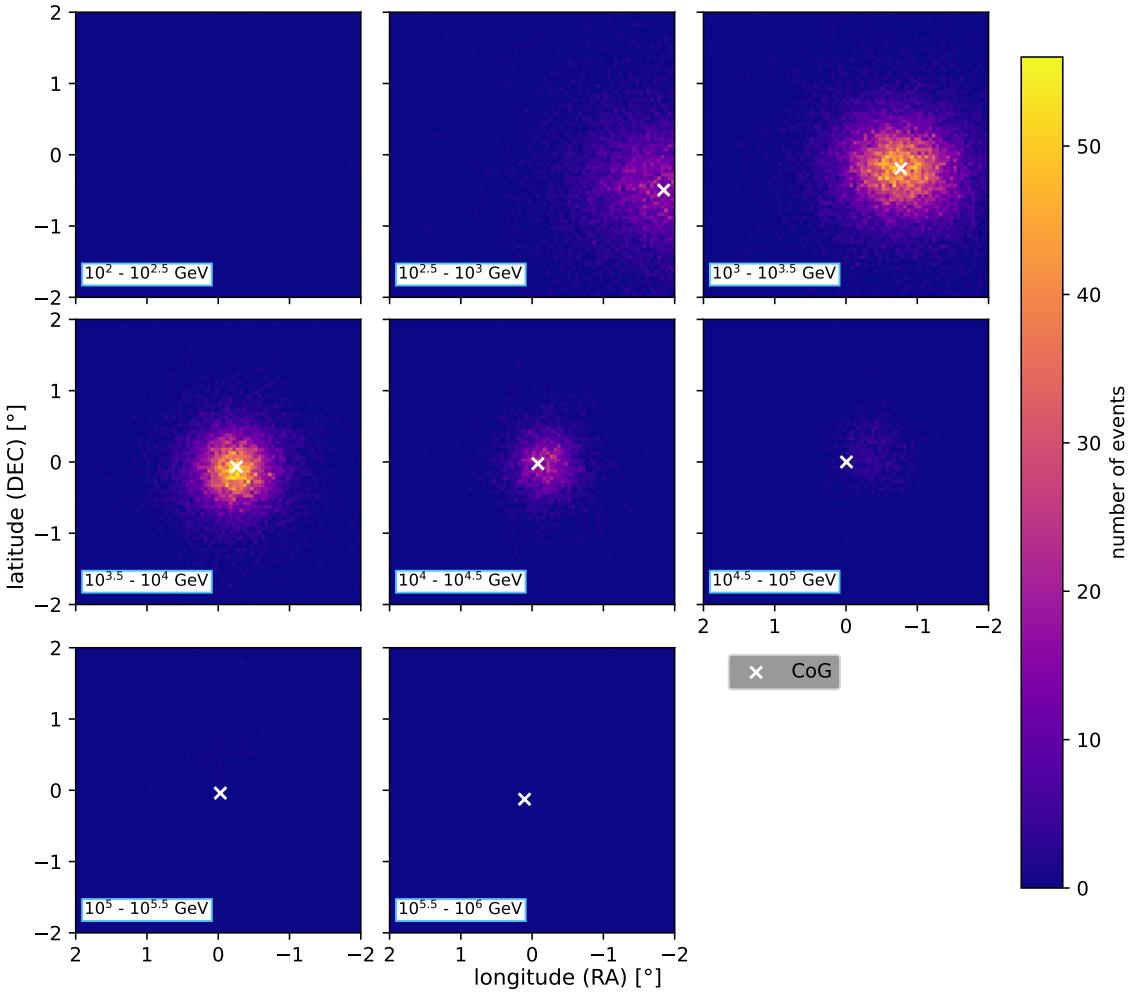

Figure 56: Subtracted maps for a 100 h observation, binned in primary energy, with particles being deflected in the geomagnetic field. The same color scale was applied to all maps. The CoG of the first energy bin exceeds the shown window, which is why it is not seen here.

of the shadow with the results found in HAWC data [14]. For this, we run our simulation with the same IRFs determined from HAWCSim data, only now located at the HAWC site, resulting in a stronger geomagnetic field as discussed in subsection 5.1. The resulting CoG for each energy bin in comparison to the data simulated at the SWGO site and the findings of the HAWC collaboration can be found in Figure 58 (s. Table 4 in the appendix A). It is notable that the deflection in right ascension at the HAWC site is only marginally larger, although the magnetic field strength at the ground is about 1.8 times as large as the value found at the SWGO site. This is why the data found after evaluating both simulations seems to show a similar deflection, while also being in accordance with the results from HAWC. The simulation at the HAWC site also serves as a cross check, verifying the simulation of the deflected moon shadow as sensible. The deflection in declination we found when simulating at the HAWC site also matches the findings in [14]. It is especially interesting to note that the cosmic ray primaries at

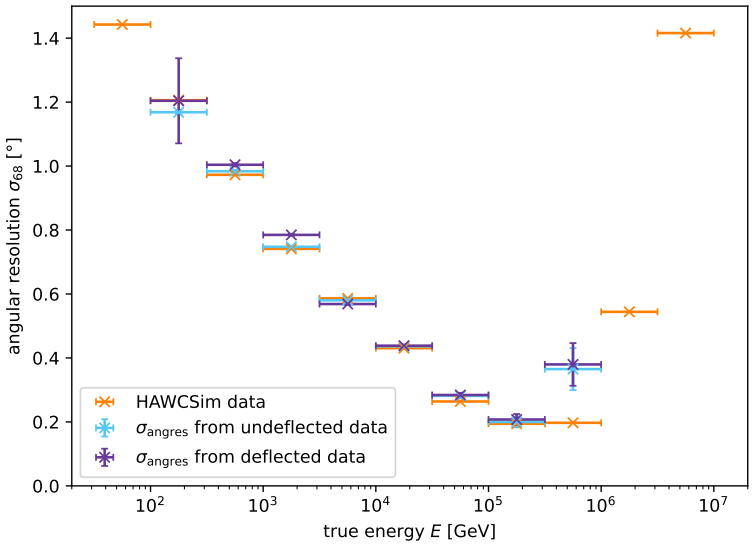

Figure 57: The values of σ_{angres} as determined from the subtracted moon shadow in comparison with the angular resolution determined from HAWCSim, now additionally with the values determined from the deflected shadow. The number of data points for the shadow data is smaller than that of HAWCSim, because the energy range we have chosen for our custom simulation is smaller.

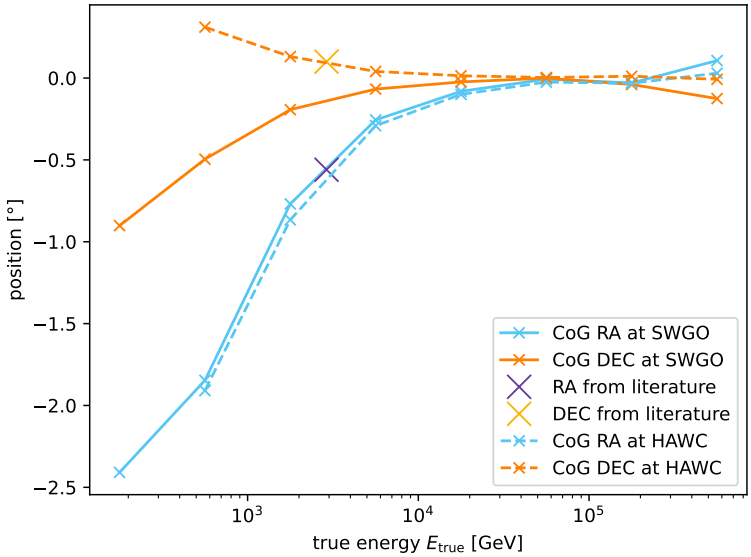

Figure 58: The Center of Gravity of the deflected moon shadow in each energy bin for a simulation at the SWGO and HAWC sites. The first energy bin is missing in the HAWC data. The markers represent the values found by the HAWC collaboration in [14]. The lines are shown as a guide for the eye.

the SWGO site are deflected in the opposite direction in declination. This is due to the different direction of the declination of the geomagnetic field at these locations, as determined in subsection 5.1. Figure 59 however shows the absolute values to again

approximately match in the simulations at the two sites, though now the deflection at the SWGO site is stronger due to the larger absolute declination of the geomagnetic field.

6 Conclusion

After this thorough analysis one can justifiably say that we have successfully observed the moon shadow in SWGO cosmic ray simulation data, using the IRFs of the detector we have independently determined from CORSIKA and HAWCSim data. By employing a custom modular simulation, we examined the visibility of the shadow in different energy bands for different observation times, computing detailed particle trajectories to account for the deflection of primary cosmic rays in the geomagnetic field. Particular emphasis lies on the scalability of this simulation pipeline, enabling an efficient simulation in extended windows and for long observation times. By utilizing statistically sensible evaluation procedures, we were able to determine values for the angular resolution σ_{angres} from the shadow that were in accordance with the original values employed in the simulation. Furthermore, we were able to determine the shift of the moon shadow due to the deflection of particles using an accurate model of the geomagnetic field and found values consistent with the findings of the HAWC collaboration.
However, there are many areas of interest to still be examined in further studies. Firstly, the simulations we have evaluated to determine the IRFs of our detector do not yet include an energy reconstruction, which is why only true energy values were used in this thesis. An energy reconstruction would be very interesting to get an idea of the energy resolution of the detector and to receive more realistic results of the moon shadow simulation.
Another aspect in need of further contemplation is the observability of the shadow for different positions of the moon, to determine the dependency of moon shadow visibility on the zenith angle and to be able to observe the moon shadow while the moon is moving over the hemisphere. This also includes the further study of the dependency of the energy spectrum and the angular resolution on the zenith angle of an observed event.
The most interesting prospect remains the outlook on the real data that will be recorded by SWGO, to apply the method described in this thesis and to hopefully verify our findings. The method can also be transferred to the observation of the sun shadow, which poses as another interesting application. Ideally, our results will serve as a template to determine the detector's angular resolution from moon shadow observation and thereby help enable the observation of gamma rays in the Southern Hemisphere or the detection of anti-protons in the cosmic ray spectrum.

A Appendix

Derivation of the error on the RMS

Suppose the detected event positions lon and lat of particles blocked by the moon are identically and independently distributed according to a normal distribution, which poses as a good approximation). Then the variance of this distribution can be approximated by the sample variance σ_x^2, where $x \in \{\text{lon}, \text{lat}\}$. We assume the values of σ_x to be χ^2-distributed, with a variance of $S^2 = \text{Var}(\sigma_x^2) = \frac{2\sigma_x^4}{n-1}$. We estimate the true value of σ_x by our measured values σ_{lon} and σ_{lat}, where n is the number of events considered for this estimation. To determine the fluctuation of σ_x from the width $\sqrt{S^2}$ of the distribution of σ_x^2, we use Gaussian error propagation:

$$(\Delta\sigma_x)^2 = \left(\frac{\partial\sigma_x}{\partial\sigma_x^2}\right)^2 \cdot S^2 = \left(\frac{\partial\sqrt{\sigma_x^2}}{\sigma_x^2}\right)^2 \cdot S^2 = \left(\frac{1}{2\sqrt{\sigma_x^2}}\right) \cdot S^2 = \frac{1}{4\sigma_x^2} \cdot S^2 = \frac{\sigma_x^2}{2(n-1)} \quad (61)$$

ultimately yielding the sought-after uncertainty

$$\Delta\sigma_x = \frac{\sigma_x}{\sqrt{2(n-1)}} \quad (62)$$

Derivation of the RMS of a circular disc

Given a two dimensional circular disc of radius R, on which values are uniformly distributed, the probability density at a given position (r, φ) is:

$$P(r, \varphi) = \frac{r}{\pi R^2} \quad (63)$$

Integration over φ yields

$$P_r(r) = \int_0^{2\pi} \mathrm{d}\varphi P(r, \varphi) = \frac{2r}{R^2} \quad (64)$$

The expected value is then

$$\text{E}[r^2] = \int_0^R \mathrm{d}r\, r^2 P_r(r) = \frac{2}{R^2}\int_0^R \mathrm{d}r\, r^3 = \frac{2}{R^2} \cdot \frac{R^4}{4} = \frac{R^2}{2} \quad (65)$$

Then the RMS, here referred to as σ_{moon}, is

$$\text{RMS} = \sqrt{\text{E}[r^2]} = \frac{R}{\sqrt{2}} \quad (66)$$

Tables and Figures

E [GeV]	Δlon[°]	Δlat[°]	σ_{lon}[°]	σ_{lat}[°]
10^2-$10^{2.5}$	−0.004(8)	−0.010(8)	0.835(6)	0.836(6)
$10^{2.5}$-10^3	−0.005(5)	−0.003(5)	0.708(4)	0.705(4)
10^3-$10^{3.5}$	0.002(4)	0.001(4)	0.542(3)	0.544(3)
$10^{3.5}$-10^4	−0.005(4)	−0.004(4)	0.428(3)	0.430(3)
10^4-$10^{4.5}$	0.001(6)	−0.007(6)	0.333(4)	0.333(4)
$10^{4.5}$-10^5	−0.013(9)	−0.017(9)	0.237(6)	0.233(6)
10^5-$10^{5.5}$	0.043(17)	0.011(20)	0.171(12)	0.205(14)
$10^{5.5}$-10^6	0.05(8)	−0.16(9)	0.27(6)	0.30(7)

Table 2: The values of the CoG and RMS as found by evaluation of the undeflected subtracted maps.

E [GeV]	Δlon[°]	Δlat[°]	σ_{lon}[°]	σ_{lat}[°]
10^2-$10^{2.5}$	−2.41(14)	−0.9(3)	0.69(10)	1.00(15)
$10^{2.5}$-10^3	−1.847(8)	−0.496(8)	0.729(6)	0.712(6)
10^3-$10^{3.5}$	−0.769(4)	−0.193(4)	0.595(3)	0.541(3)
$10^{3.5}$-10^4	−0.256(4)	−0.067(4)	0.424(3)	0.418(3)
10^4-$10^{4.5}$	−0.081(6)	−0.024(6)	0.332(4)	0.336(4)
$10^{4.5}$-10^5	−0.007(9)	−0.001(9)	0.236(6)	0.238(6)
10^5-$10^{5.5}$	−0.030(20)	−0.039(20)	0.207(14)	0.177(12)
$10^{5.5}$-10^6	0.11(9)	−0.13(9)	0.29(7)	0.30(7)

Table 3: The values of the CoG and RMS as found by evaluation of the deflected subtracted maps simulated at the SWGO site.

E [GeV]	Δlon[°]	Δlat[°]	σ_{lon}[°]	σ_{lat}[°]
$10^{2.5}$-10^3	−1.909(9)	0.311(9)	0.670(6)	0.690(6)
10^3-$10^{3.5}$	−0.866(4)	0.131(4)	0.610(3)	0.544(3)
$10^{3.5}$-10^4	−0.291(4)	0.040(4)	0.437(3)	0.426(3)
10^4-$10^{4.5}$	−0.100(5)	0.014(6)	0.331(4)	0.343(4)
$10^{4.5}$-10^5	−0.025(8)	0.003(8)	0.213(6)	0.222(6)
10^5-$10^{5.5}$	−0.03(3)	0.021(19)	0.213(15)	0.193(14)
$10^{5.5}$-10^6	0.03(3)	−0.01(4)	0.15(3)	0.20(3)

Table 4: The values of the CoG and RMS as found by evaluation of the deflected subtracted maps simulated at the HAWC site.

E [GeV]	σ_{undef}[°]	$\Delta\sigma_{\text{undef}}$[°]	σ_{def}[°]	$\Delta\sigma_{\text{def}}$[°]
10^{2}-$10^{2.5}$	1.17	0.14	1.2	0.7
$10^{2.5}$-10^{3}	0.98	0.10	1.00	0.13
10^{3}-$10^{3.5}$	0.75	0.08	0.78	0.08
$10^{3.5}$-10^{4}	0.58	0.07	0.57	0.07
10^{4}-$10^{4.5}$	0.44	0.07	0.44	0.07
$10^{4.5}$-10^{5}	0.28	0.07	0.28	0.07
10^{5}-$10^{5.5}$	0.20	0.08	0.21	0.08
$10^{5.5}$-10^{6}	0.4	0.3	0.4	0.3

Table 5: The values of σ_{angres} with and without deflection in the geomagnetic field with their respective errors.

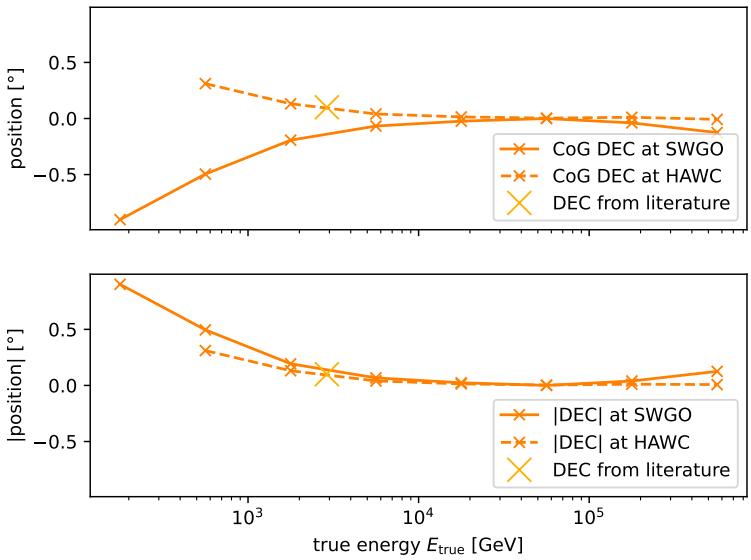

Figure 59: The CoG of the latitude ($\hat{=}$ declination) for simulations at the SWGO and HAWC site. The absolute values are shown for easier comparison.

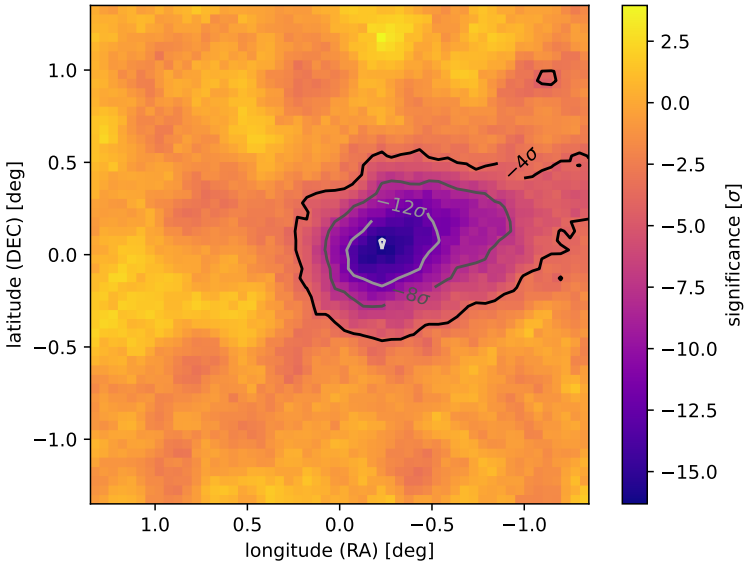

Figure 60: An all-energy sky map of the moon shadow for a 100 h observation of the moon at the HAWC site, with particles being deflected in the geomagnetic field. The contour lines are meant to visualize areas of the same significance level.

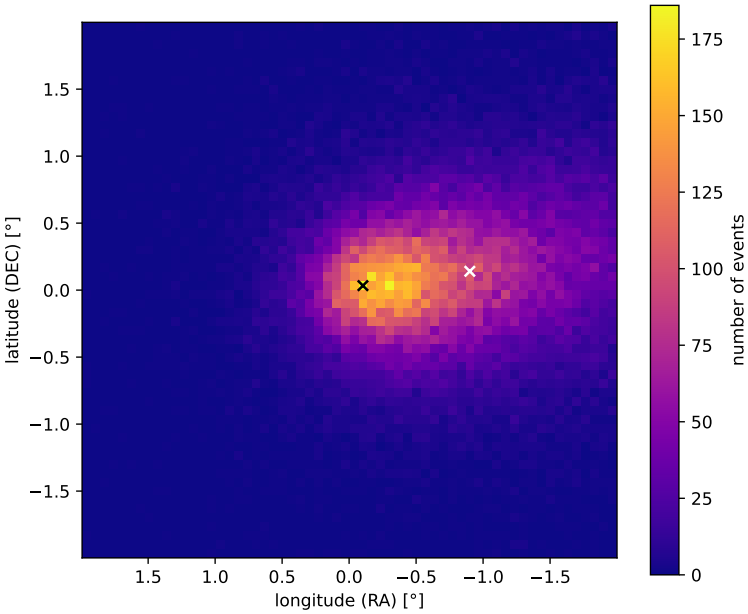

Figure 61: An all-energy subtracted map of the moon shadow for a 100 h observation, with particles being deflected in the geomagnetic field. The markers depict the CoG and the peak of the distribution in longitude and latitude.

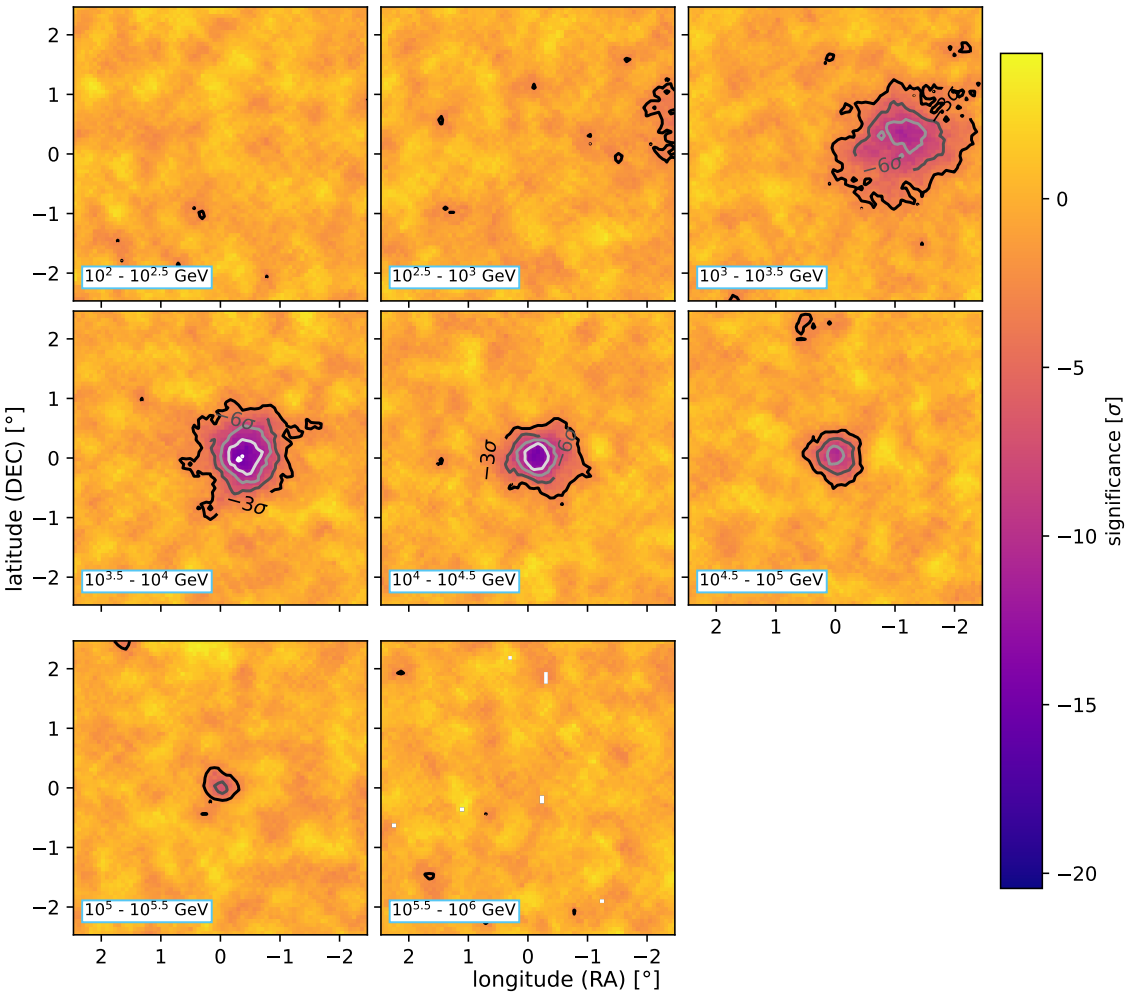

Figure 62: Sky maps for a 100 h observation of the moon at the HAWC site, with particles being deflected in the geomagnetic field. The same color scale was applied to all maps. The contour lines are meant to visualize areas of the same significance level.

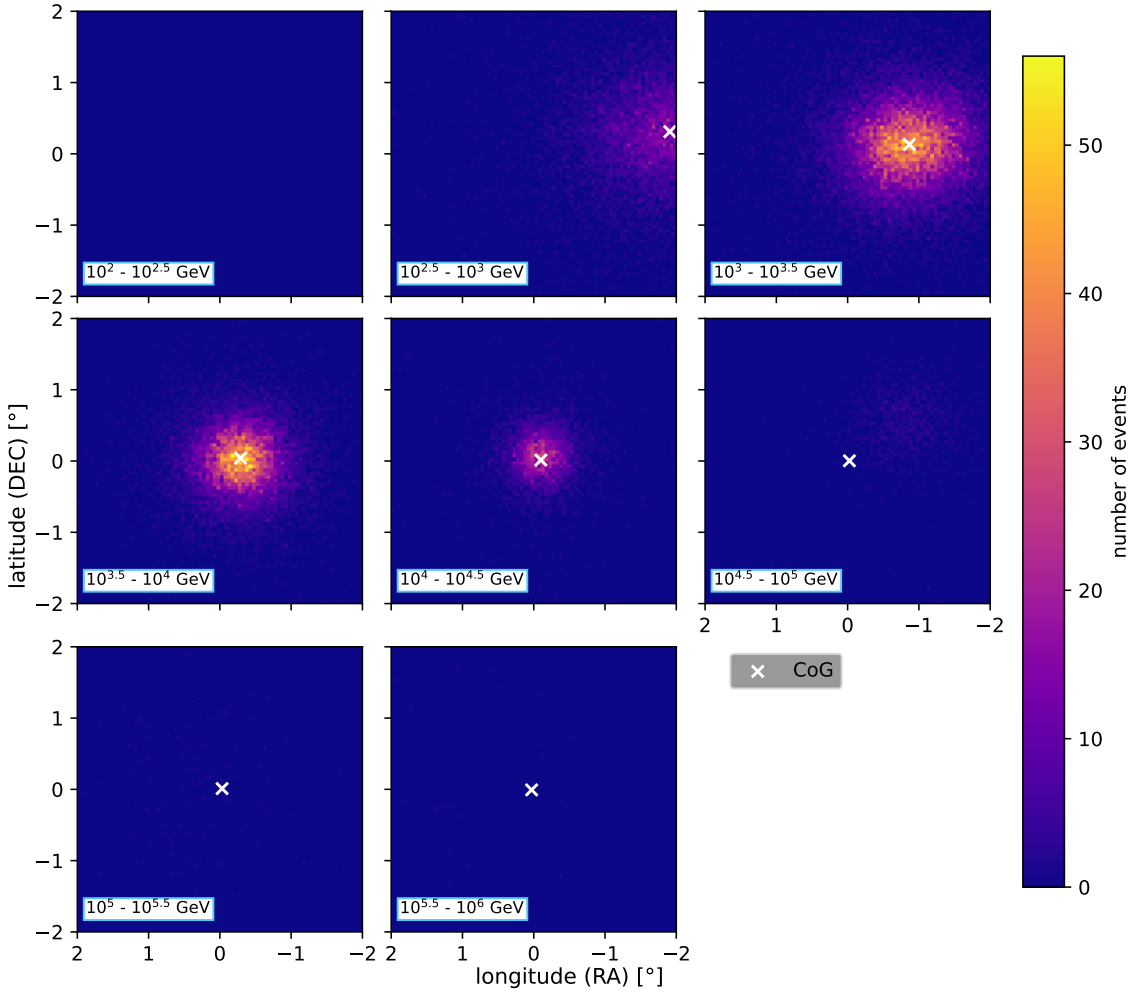

Figure 63: Subtracted maps for a 100 h observation at the HAWC site, binned in primary energy, with particles being deflected in the geomagnetic field. The same color scale was applied to all maps. The CoG of the first energy bin exceeds the shown window, which is why it is not seen here.

Bibliography

[1] *A simple figure explaining latitude and longitude*. [Online; accessed 16-January-2026]. 2005. URL: https://commons.wikimedia.org/wiki/File:FedStats_Lat_long.svg.

[2] H. Abdalla, A. Abramowski, F. Aharonian, F. Ait Benkhali, et al. "The H.E.S.S. Galactic plane survey". In: *Astronomy & Astrophysics* 612 (Apr. 2018), A1. ISSN: 0004-6361, 1432-0746. URL: https://www.aanda.org/articles/aa/abs/2018/04/aa32098-17/aa32098-17.html (visited on 02/04/2026).

[3] A. Abdul Halim, P. Abreu, M. Aglietta, I. Allekotte, et al. "Large-scale Cosmic-ray Anisotropies with 19 yr of Data from the Pierre Auger Observatory". In: *The Astrophysical Journal* 976.1 (Nov. 2024). Publisher: The American Astronomical Society, p. 48. ISSN: 0004-637X. URL: https://doi.org/10.3847/1538-4357/ad843b (visited on 01/13/2026).

[4] A. U. Abeysekara, A. Albert, R. Alfaro, C. Alvarez, et al. "Observation of the Crab Nebula with the HAWC Gamma-Ray Observatory". In: *The Astrophysical Journal* 843.1 (July 2017), p. 39. ISSN: 0004-637X. URL: https://ui.adsabs.harvard.edu/abs/2017ApJ...843...39A/abstract (visited on 01/20/2026).

[5] Ag2gaeh. *Kugelkoordinaten: lokale Basis (engl. Vers.)* https://commons.wikimedia.org/wiki/File:Kugelkoord-lokb-e.svg. [Online; accessed 07-January-2026]. 2023.

[6] Hannu Karttunen et al. *Fundamental Astronomy*. 5th ed. 2007.

[7] A. Cillis and S. J. Sciutto. *Geomagnetic Field and Air Shower Simulations*. arXiv:astro-ph/9712345. Dec. 1997. URL: http://arxiv.org/abs/astro-ph/9712345 (visited on 01/22/2026).

[8] HAWC Collaboration. *HAWC. The High-Altitude Water Cherenkov Gamma-Ray Observatory*. [Online; accessed 15-January-2026]. 2011-2026. URL: https://www.hawc-observatory.org/.

[9] H.E.S.S. Collaboration. *The full H.E.S.S. Array in Namibia*. [Online; accessed 15-January-2026]. 2004-2026. URL: https://hess-experiment.eu/telescopes/.

[10] NASA/DOE/Fermi LAT Collaboration. *The Compton Legacy: A Quarter-century of Gamma-ray Science*. https://svs.gsfc.nasa.gov/12194/. [Online; accessed 30-December-2025]. 2016.

[11] SWGO Collaboration, P. Abreu, R. Alfaro, A. Alfonso, et al. *Science Prospects for the Southern Wide-field Gamma-ray Observatory: SWGO*. arXiv:2506.01786 [astro-ph]. June 2025. URL: http://arxiv.org/abs/2506.01786 (visited on 10/13/2025).

[12] The SciPy Community. *SciPy Documentation*. [Online; accessed 31-January-2026]. 2008. URL: https://docs.scipy.org/doc/scipy/index.html.

[13] The Astropy Developers. *The Astropy Project*. [Online; accessed 31-January-2026]. 2011-2025. URL: https://www.astropy.org/.

[14] Daniel W Fiorino, Segev Benzvi, and James Braun. "Observation of the Moon Shadow and Characterization of the Point Response of HAWC-30". In: *33rd International Cosmic Ray Conference* (2013).

[15] J. Glombitza, M. Schneider, F. Leitl, S. Funk, and C. van Eldik. "Application of graph networks to a wide-field water-Cherenkov-based Gamma-ray Observatory". In: *Journal of Cosmology and Astroparticle Physics* 2025.02 (Feb. 2025), p. 066. ISSN: 1475-7516. URL: https://iopscience.iop.org/article/10.1088/1475-7516/2025/02/066 (visited on 01/20/2026).

[16] Zigfried Hampel-Arias. *Cosmic Ray Observations at the TeV scale with the HAWC Observatory*. 2017. URL: https://hawc-observatory.org/publications/docs/HampelThesis.pdf (visited on 10/14/2025).

[17] William F Hanlon. *The Energy Spectrum of Ultra High Energy Cosmic Rays measured by the High Resolution Fly's Eye Observatory in stereoscopic mode*. 2008.

[18] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, and T. Thouw. *CORSIKA: A Monte Carlo code to simulate extensive air showers*. 1998. URL: https://publikationen.bibliothek.kit.edu/270043064 (visited on 01/16/2026).

[19] Inductiveload. *Spherical Coordinates (Colatitude, Longitude)*. [Online; accessed 07-January-2026]. 2019. URL: https://de.wikipedia.org/wiki/Datei:Spherical_Coordinates_%28Colatitude,_Longitude%29_%28b%29.svg.

[20] Karl M. Laundal. *Pure Python International Geomagnetic Reference Field*. [Online; accessed 08-January-2026]. 2021. URL: https://github.com/IAGA-VMOD/ppigrf.

[21] CMG Lee. *A diagram of the equatorial coordinate system*. [Online; accessed 04-February-2026]. 2025. URL: https://commons.wikimedia.org/wiki/File:Equatorial_coordinates.svg?lang=en.

[22] Ti-pei Li and Yu-qian Ma. "Analysis Methods for results in Gamma-Ray Astronomy". In: *The Astrophysical Journal* (1983).

[23] Malcolm S. Longair. *High energy astrophysics*. 3rd ed. Cambridge: Cambridge University Press, 2011. ISBN: 978-0-521-75618-1.

[24] J. Matthews. "A Heitler model of extensive air showers". In: *Astroparticle Physics* 22.5-6 (Jan. 2005), pp. 387–397. ISSN: 09276505. URL: https://linkinghub.elsevier.com/retrieve/pii/S0927650504001598 (visited on 11/07/2025).

[25] MikeRun. *Magnetic field around a spherical dipole magnet like the earth with a schematic dipole magnet drawn to its center*. [Online; accessed 29-January-2026]. 2020. URL: https://commons.wikimedia.org/wiki/File:VFPt_Earths_Magnetic_Field_Confusion_overlay.svg.

[26] M. Tanabashi et al. (Particle Data Group). *29. Cosmic Rays*. 2019. URL: https://pdg.lbl.gov/2019/reviews/rpp2019-rev-cosmic-rays.pdf (visited on 12/29/2025).

[27] Frank M. Rieger, Emma de Ona-Wilhelmi, and Felix A. Aharonian. *TeV Astronomy*. Feb. 2013. URL: https://arxiv.org/abs/1302.5603v1 (visited on 02/04/2026).

[28] LHAASO Team. *LHAASO. The Large High Altitude Air Shower Observatory.* [Online; accessed 16-January-2026]. 2002-2026. URL: https://www.hawc-observatory.org/.

[29] *Theo_Phy_I_Kapitel3_Koordinatensysteme.* URL: https://www.uni-rostock.de/storages/uni-rostock/Alle_MNF/Physik_Qms/Lehre_Scheel/theo_I/Theo_Phy_I_Kapitel3_Koordinatensysteme.pdf (visited on 01/07/2026).

[30] TWCarlson. *A schematic diagram of the terms "Azimuth" and "Altitude" as they relate to the viewing of celestial objects.* https://commons.wikimedia.org/wiki/File:Azimuth-Altitude_schematic.svg. [Online; accessed 07-January-2026]. 2020.

[31] R. Ulrich, J. Blümer, R. Engel, F. Schüssler, and M. Unger. "On the measurement of the proton-air cross section using air shower data". In: *New Journal of Physics* 11.6 (June 2009), p. 065018. ISSN: 1367-2630. URL: https://iopscience.iop.org/article/10.1088/1367-2630/11/6/065018 (visited on 02/04/2026).

Acknowledgements

I want to wholeheartedly thank anyone who supported me throughout writing this thesis and during my studies:

- My supervisor, Markus Pirke, who proved to have a lot of patience to answer "stupid" questions and who has been the best support ever since starting this thesis.
- My friends and my partner for their emotional support and help in proofreading this thesis. Any university degree is not possible to accomplish alone and I would certainly not have come this far without them.
- Nintendo, for creating the best music to listen to while writing a thesis.
- My physics teachers, without whose great classes I certainly would not have decided to study physics.
- My family, for always supporting my decision to study physics and enabling my studies.

Declaration of Originality

I, Lena Jäger, student registration number: 23114201, hereby confirm that I completed the submitted work independently and without the unauthorized assistance of third parties and without the use of undisclosed and, in particular, unauthorized aids. This work has not been previously submitted in its current form or in a similar form to any other examination authorities and has not been accepted as part of an examination by any other examination authority.

Where the wording has been taken from other people's work or ideas, this has been properly acknowledged and referenced. This also applies to drawings, sketches, diagrams and sources from the Internet.

In particular, I am aware that the use of artificial intelligence is forbidden unless its use an aid has been expressly permitted by the examiner. This applies in particular to chatbots (especially ChatGPT) and such programs in general that can complete the tasks of the examination or parts thereof on my behalf.

Any infringements of the above rules constitute fraud or attempted fraud and shall lead to the examination being graded "fail" ("nicht bestanden").

Place, Date	Signature