Reconstructing the cosmic-ray energy spectrum using HiSPARC data

Keerthana Umesh
The HiSPARC experiment, Nikhef, Science Park 105, 1098 XG, Amsterdam, The Netherlands
Email: keerthanaumesh25@gmail.com
December 2020

Supervised by—
Kasper van Dam[a]
Prof. dr. ing. Bob van Eijk[a,b]

[a] *Nikhef National Institute for Subatomic Physics*
[b] *Faculty of Science and Technology, University of Twente*

1 Introduction

1.1 Cosmic rays

Cosmic rays are energetic particles that move through space at very high speeds. These rays are primarily made up of charged subatomic particles (~90% protons, ~9% alpha particles), gamma particles and atomic nuclei [1]. Lower energy cosmic rays are thought to originate from the Sun during periods of high solar activity, but the origins of higher energy cosmic rays are still unknown. Although we can reconstruct the incoming direction of these rays from the earth's surface, charged rays travelling such long distances are often accelerated and diverged by magnetic fields in the interstellar medium, and so the direction of arrival does not always indicate the direction of the source [2]. These high-energy rays are one of the great mysteries of modern physics, and the subject of much ongoing research.

When cosmic rays reach the earth's atmosphere, they interact with the nitrogen and oxygen nuclei present, which results in the production of secondary cosmic ray particles. Gamma particles interacting with the electromagnetic field of a nucleus results in the pair production of an electron and positron, which then interacts with the field once more to result in new photons that continue this cycle. Protons that interact with nuclei generate pions, which decay into muons, and then electrons successively [3]. The results of these two processes (known as electromagnetic and hadronic respectively), results in a huge cascade of particles, each with less energy than the particle before it. This known as a cosmic air shower. At a certain height, the number of particles reaches its maximum, after which they start to decay.

Most cosmic air showers "die out" before they can reach the earth's surface, but the ones that do are are usually the result of higher energy cosmic rays ($> 10^{14}$ eV) [1]. These showers are called Extensive Air Showers (EAS), and their remnants can be detected on the earth's surface. For primary rays perpendicular to the earth's surface, the corresponding footprint of secondary particles reaching the earth's surface can range from a few meters to several kilometers in diameter [1].

1.2 Cosmic ray energies

Cosmic rays are particularly interesting because of the large energies that they can have. Energies of primary cosmic rays range from 10^9 eV to about 10^{21} eV. The latter is known as the GZK cut-off, and is thought to be the upper limit for the energies of cosmic rays [4]. Lower energy rays are the most abundant, and the flux decreases rapidly with an increase in energy. Rays with the highest energies far exceed the capacity of particle accelerators on the earth's surface.

When cosmic rays reach the earth's atmosphere and interact, the resultant new particles are diverged and scattered from the incident axis. The lateral profile of the cosmic shower *perpendicular* to this axis grows linearly with each successive interaction, until the critical energy is reached, and no new particles can be produced [2]. The lateral size of the shower is proportional to the energy of the primary cosmic ray. Moving away from the shower core, the number of particles falls drastically [2]. This distribution of these particles along the shower front, i.e. their later profile, can be described using a lateral distribution function (LDF).

Because cosmic particles move very close to the speed of light, they experience time dilation, and as a result, a much larger percentage of these particles are able to reach the earth's surface before they decay than their lifetimes would otherwise permit. The shower front of incident particles is usually very slightly curved along the edges as lower energy particles "lag behind", but it can be approximated to an infinitely thin plane for simplicity when reconstructing arrival directions [4].

The flux of cosmic rays, measured in $[\mathrm{m}^2 \text{ sr s GeV}]^{-1}$ decreases logarithmically with an increase in energy, and can be described using the function:

$$F(E) = E^{-\gamma} \tag{1}$$

γ can take two distinct values: $\gamma \approx 2.7$ for energies smaller than 10^{15} and bigger than 10^{18}, and is ≈ 3.1 for the energies between them [5]. The two turning points on the corresponding cosmic ray energy spectrum are called the "knee" and the "ankle" respectively.

In this study, an effort was made to reconstruct the flux of cosmic rays using their corresponding footprint on the earth's surface. In particular, the arrival direction of the shower, and its lateral distribution of particles were used for the reconstructions. The data used in this study was collected by the HiSPARC detector array located at the Amsterdam Science Park.

1.3 The HiSPARC project

The High School Project on Astrophysics and Research with Cosmics (HiSPARC) is a research initiative that aims to involve highschools, together with academic institutions, in the search for cosmic rays. Since its establishment in 2003, the project has expanded from its base in the Netherlands to also include schools in Denmark, England and Germany [1]. Schools and universities who choose to take part in the HiSPARC project build, install and run detectors on their rooftops which collect a multitude of data on cosmic ray activity in the region. The data from this array of detectors is then assembled and saved in the HiSPARC servers at Nikhef, from where it can then be accessed through a public database and used in cosmic ray research.

HiSPARC detectors are made up of three parts: a scintillator, a photomultiplier tube (PMT), and an electronics box. The detection plate itself is a rectangular piece of plastic of about $0.5\mathrm{m}^2$ doped in scintillating material (anthracene) [4]. When a small percentage of charged particles from cosmic rays pass through the scintillator, their energy is absorbed through ionization. This excites the electrons in the fluor, which then release the energy again when returning to the ground state in the form of blue-purple light. These resultant photons are guided into the PMT, which makes use of the photoelectric effect to convert the light energy into electrical signals. The electrical signals are finally read and interpreted by the HiSPARC electronics into digital signals through the use of an analogue digital converter (ADC). The size of these digital pulses is proportional to the number of minimum ionizing particles [1] (MIPs) passing through the scintillator, and their mean energy loss approximately follows that of the Landau distribution: a large peak caused by single MIPS, followed by a longer "tail" made up of atmospheric photos and particles that arrived later in the shower [1].

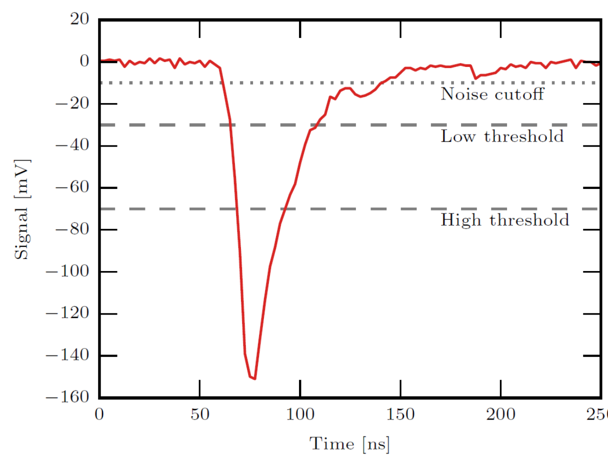

Fig. 1: An example of a signal as captured by the HiSPARC detectors with its characteristic peak and "tail" [2]

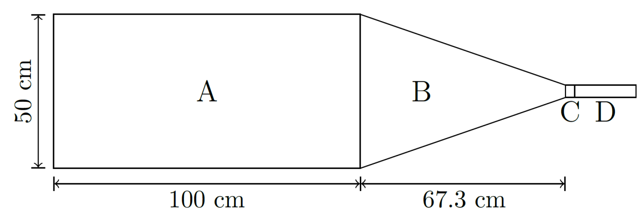

Fig. 2: Schematic diagram of a detector: The scintillator is denoted by the letter A, and the light guide and its adapter piece with B and C. The photomultiplier tube (D) is mounted onto the adapter piece [1].

A detector "station" typically involves 2 or 4 detectors in either a triangular or a rhombic set-up, a weather station that measures the local temperature, air pressure, humidity etc. and a GPS module that measures the times and locations of the detectors. Each station has its own unique station number. The detectors are set up in such a way that each of them have a higher, and a lower threshold level (usually 70mV and 30mV respectively), and a signal is registered when the pulse from the scintillator passes the higher threshold in at least two detectors, or passes the lower in three. An "event" is only recorded as such if two or more detectors in a station measure a signal within 1.5 μs of each other; this way,

[1] A MIP is a particle whose mean energy loss through a medium is at a minimum. Most relativistic particles (like cosmic particles) are minimum ionizing

single background muons and random gamma particles are eliminated [2].

When two or more stations measure an event in quick succession, a "coincidence" is said to have taken place. This is usually the result of a larger cosmic ray shower. Signals are sampled every 2.5 ns per detector, so a single shower passing through in several microseconds is sampled several times, and the shape of the resulting trace over time is similar to the shape of the shower front [4]. The pulse heights, pulse integrals and arrival times of the signals, along with the weather data is saved and can be accessed from the central database at Nikhef.

Groups of geographically close stations are arranged into “clusters”, and if they are particularly large, can be broken up into “subclusters”. In this study, only data from four-detector stations in the Amsterdam Science Park subcluster were used, since they typically improve the shower sampling and allow for better direction reconstructions [2]. Much of the analysis was done in Python using the SAPPHiRE package, a library of scripts and methods developed specifically for HiSPARC data analysis .

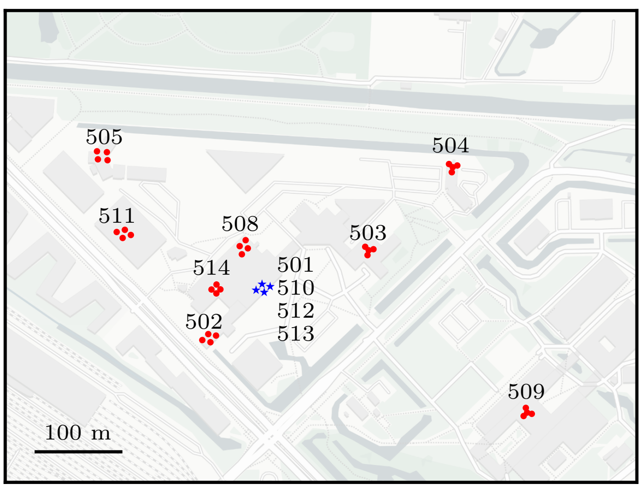

Fig. 3: An aerial view of the Amsterdam Science Park array containing all the stations used in this study. The red dots represent detectors, and the blue stars each represent a station with four detectors. Figure courtesy of K. van Dam [6]

2 Data quality and selection

2.1 Time period

Before beginning with the reconstructions of direction and energy, it was first important to determine a suitable time period, and the quality of data available during it. The flux for higher energy cosmic rays can drop to single events per square meter *per year*, so in order to be able to reconstruct them accurately, a large amount of data ranging over several months is needed [1]. This is complicated by the fact that stations are regularly being changed, updated or repaired, and that during certain time periods, they were not operational at all. Sometimes, one or more detectors in a station can fail, resulting in fewer events being recorded on that day.

The Amsterdam Science Park cluster is made up of 11 stations, and in order to be able to reconstruct the directions accurately, at least six of them need to be functional and active. The best time period, with the most number of stations operational was determined to be from Nov 2018-Sep 2020. This almost two-year time period has the most stable configuration of detectors, while still being large enough to include measurements of high-energy cosmic rays. Fig.4 shows the amount of data recorded during this time period on a day-to-day basis.

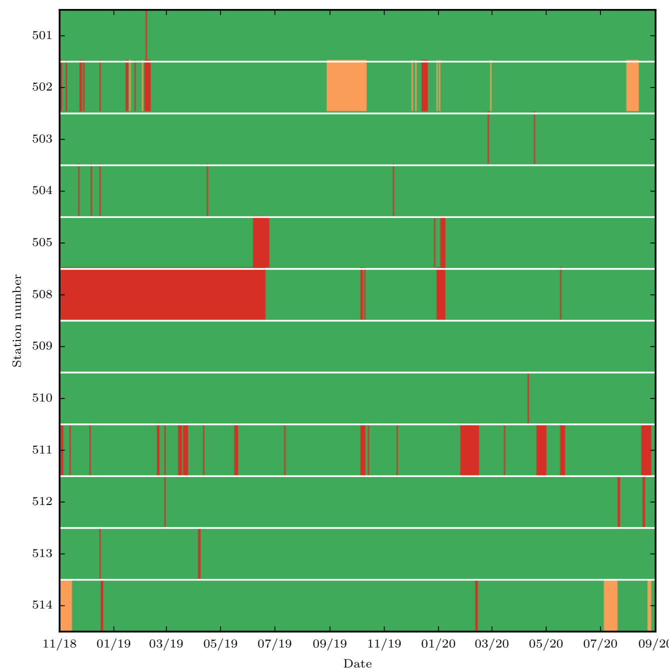

Fig. 4: Daily status of stations from Nov 2018 to Sep 2020. This is based on the number of events detected per day- if somewhere around the accepted 50000 mark, then the station was assumed to be active throughout and coloured green. If very few events were detected, the station was assumed to be partially active and coloured yellow, and if no events were detected at all for that day, it is coloured red.

2.2 Data quality

Now that a time period was chosen, histograms of the number of events recorded per day were plotted and compared for every station (see Fig 5). From this, two points became clear. The first is that some stations, like station 509 and 511, tended to have two peaks- i.e. the mean number of events recorded per day was unclear. And secondly, even if the station had only one peak, it still had a large spread in the number of events recorded- it varied significantly from day to day. In order to determine what was causing this large spread, two possible factors were tested: fluctuations in the weather (atmospheric temperature and pressure), and fluctuations in the voltage gain of the PMT.

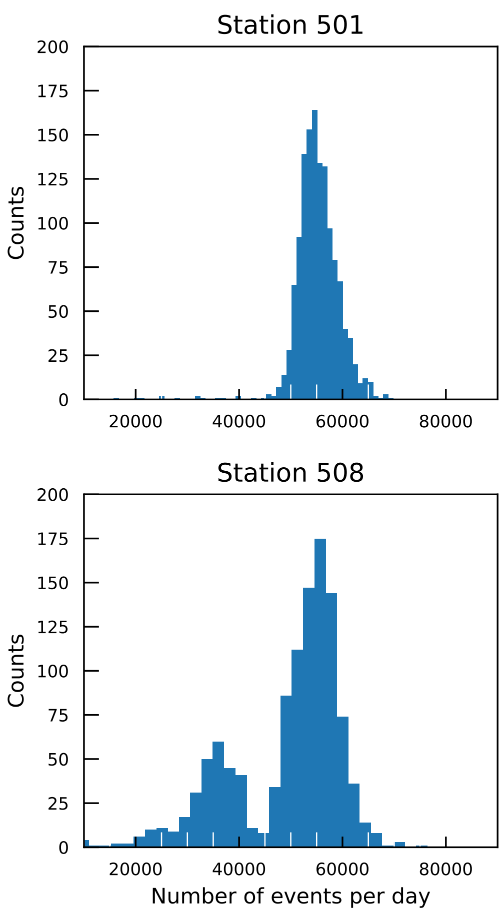

Fig. 5: An example of the large spread (top) and "double peak" (bottom) in the number of events recorded per day. Possible reasons include fluctuations in the weather and PMT voltage.

HiSPARC detectors make use of two types of PMT bases, one that is developed by Nikhef, and another commercial version. The temperature of HiSPARC PMTs can drop to as low as -30°C and go as high as +60° with changes in the atmospheric temperature [1]. For the commercial PMT bases, a higher temperature results in a lower gain, but for Nikhef-created bases, a higher temperature leads to a higher gain. Usually, this is corrected for by averaging out the MIP-peaks every 4 hours, and using the resultant value to calibrate the PMT output [1]. However, it is still possible that the fluctuations in the gain were so high that they affected the number of signals that were able to pass the threshold per day, even with the corrections applied. This would explain the deviations in the number of events recorded per day.

Furthermore, with an increase in the atmospheric pressure, there is also a corresponding increase in the number of interactions between atmospheric and cosmic particles. Hence, cosmic particles travelling longer distances tend to lose their energy before they can reach the Earth's surface. An increase in the atmospheric pressure leads to a corresponding decrease in the number of detected particles on the Earth's surface [1]. If the number of events varies significantly with changes in atmospheric pressure, it could be another explanation for the large spread of data in Fig 5.

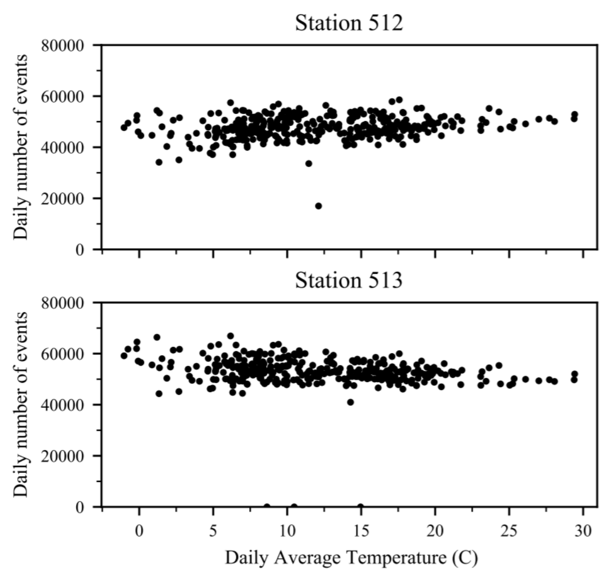

Fig. 6: Graphs showing a slight variation in the number of events depending on the daily average temperature. Both 512 and 513 have Nikhef-based PMTs, and show a slight increase in the number of events on hotter days.

And lastly, it is also possible that over this two year time period, the configurations of the detectors were changed or updated, particularly the voltage gain in the PMT. Lower gain means fewer signals pass the lower threshold, and so fewer events are recorded. It is also possible that sometimes, a failure in the electronics would result in a gain so low that only two-detectors were active- this would explain the "double-peak" around 20000 and 50000 events.

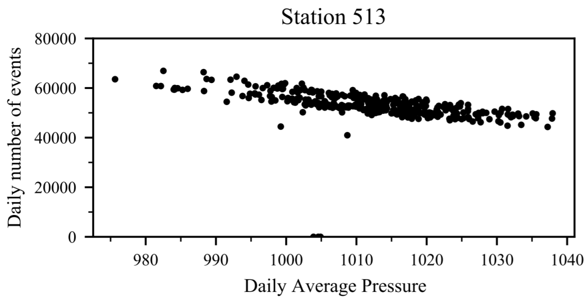

Fig. 7: A much more pronounced variation in the number of events can be seen with changes in the average daily pressure.

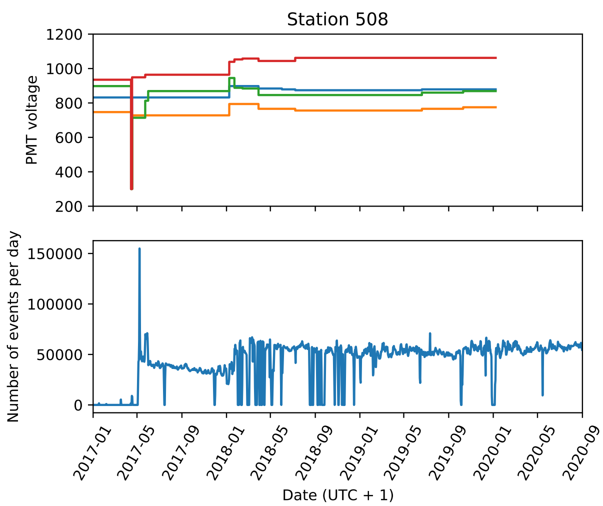

Fig. 8: The "double peaks" in the histograms were most likely caused by changes in the voltage gain of the PMTs. In this graph, a direct correlation can be seen between the voltages of each detector's PMT (coloured lines on top) and the number of events.

2.3 Timing offsets

Cosmic rays are isotropic, i.e. there is equal probability of incidence in every direction [2]. This then means that every detector in a station has equal probability of being hit first. Hence if a histogram of all of the the arrival time differences between two detectors was plotted, its peak would lie at the zero second mark. However, systematic errors in the electronics can sometimes lead to this histogram being "shifted" by a few nanoseconds. This is the timing offset of that particular set of detectors, and was corrected for, as the direction reconstructions relies heavily on the arrival times of particles.

For bigger cosmic rays that resulted in multiple stations being in coincidence, the same offset correction was applied at the station level. Unlike with the detector offsets, the timing offsets between stations seemed to fluctuate often over the two year time period. This could be for a variety of reasons, the most probable one being changes in GPS configurations during repairs or reboots. In order to work around this, the timing offset was determined at regular intervals, and the average of these intervals were chosen. Anytime there was a "dip" or huge jump in the timing offset, a new average timing offset was calculated and corrected for. Fig. 9 shows these offset fluctuations.

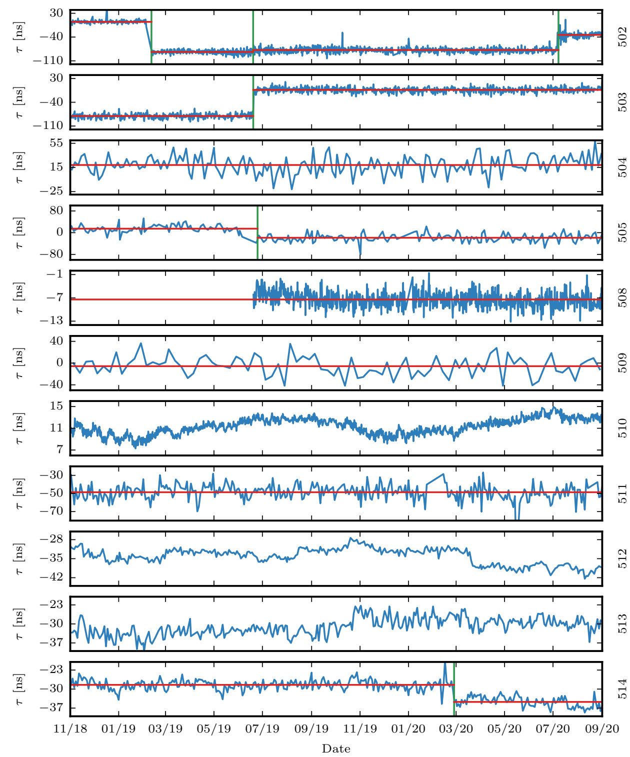

Fig. 9: Fluctuations in the timing offset between station 501 and all the others used in this study. The red lines show the average offset. Figure courtesy of K. van Dam [6]

3 Energy Flux diagram

3.1 Energy reconstructions

The energies of cosmic rays was determined from HiSPARC data in three steps:

1. Using the arrival times of stations in coincidence, the azimuthal and zenith angle of the incoming cosmic rays were calculated.

2. The particle densities measured in each detector were projected onto the reconstructed shower fronts.

3. A lateral density distribution function was fit to the projected data and used to predict the shower core and initial energy of the ray.

Direction reconstruction

The incoming directions of cosmic rays was determined by using their corresponding arrival times in each of the detectors that measured them. The shower front was assumed to be a long, infinitely thin plane passing through the array of detectors. Typically, only three detectors are needed for the method used in this study, but for the sake of accuracy, a minimum of four stations in the "core" (i.e. stations 501, 502, 503 and 514) was set as a requirement for the events to be reconstructed. Furthermore, each of the detectors involved in the event needed to have detected a minimum signal of 2 MIPs to be considered. A simplified version of the algorithm used is given below:

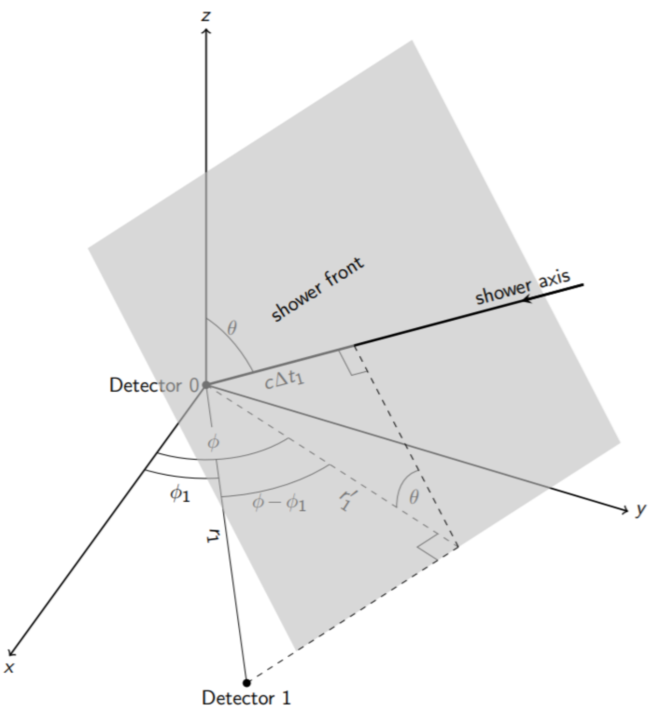

Fig. 10: Three-dimensional schematic diagram of an event, sourced from [2]

For the event in Fig 10, assume that the incoming shower axis has a zenith angle θ and an azimuthal angle ϕ with respect to the positive x and z axes. Since cosmic ray particles are relativistic, assume that the shower front is moving at speed c.

In order to pass from detector 1 to detector 0, the plane will have to travel a distance of cΔt_1, where t_1 is the difference in arrival times of the particles as measured by the HiSPARC detectors. In the xy-plane, this distance is represented as r_1'. From here, two equations can be derived:

$$\sin\theta = \frac{c\Delta t_1}{r_1'} \tag{2}$$

and

$$\cos(\phi - \phi_1) = \frac{r_1'}{r_1} \tag{3}$$

which, upon combining becomes:

$$c\Delta t_1 = r_1 \cos(\phi - \phi_1)\sin\theta \tag{4}$$

where r_1 is the distance between the two detectors, and ϕ_1 is the azimuthal angle of this distance. If data from another detector, say detector 2 was used, then two equations can be derived from equation 4 and solved simultaneously to get:

$$\tan\phi = \frac{r_1\Delta t_2\cos\phi_1 - r_2\Delta t_1\cos\phi_2}{r_2\Delta t_1\sin\phi_2 - r_1\Delta t_2\sin\phi_1} \tag{5}$$

and

$$\sin\theta = \frac{c\Delta t_1}{r_1\cos(\phi - \phi_1)} \tag{6}$$

where Δt_1, ϕ_1 and r_1 is between detector 0 and 1, and Δt_2, ϕ_2 and r_2 is between detector 0 and 2. This method for finding the zenith and azimuthal directions is called triangulation. [[2], Chapter 4] includes a more in-depth description of this algorithm.

The reconstructions are complicated by the fact that the data used in this study came from detectors that were *not* placed on a horizontal plane i.e, the altitude varies from station to station. This was taken into account by initially assuming every detector to be placed on the same xy-plane, from which and an approximation of the incoming direction was calculated. Using this approximate direction, the position coordinates and arrival times that the detectors would have had, if they all were to lie on the same plane was determined. These "predicted" positions and the original positions are lined up in such a way that the incident axis passes through both of them (see Fig. 11).

Using these new detector positions and arrival times, the direction is again repeatedly calculated using the triangulation method and groups of detector stations. These repeated iterations quickly converge onto one value, which gives the direction of incidence of the cosmic ray.

A histogram of the reconstructed zenith angles was then plotted. Although the distribution of cosmic rays is isotropic, experimental data shows that when detected on the earth's surface, cosmic rays with an incoming angle of ~20° are the

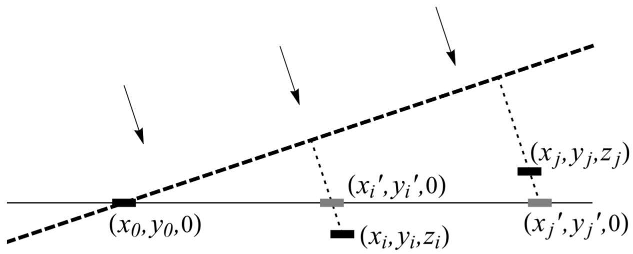

Fig. 11: 2D diagram on how the estimated direction of incidence passes through both the real detector coordinates, and the "predicted" coordinates $(x_i', y_i', 0)$ and $(x_j', y_j', 0)$

most abundant [4]. For angles below this value, the effective surface area gets increasingly smaller, thereby decreasing the probability of cosmic rays with smaller zenith angles. For angles greater than 20°, the rays have to travel a much larger distance to reach the earth's surface. The attenuation is greater, and hence fewer particles reach the surface before they decay. The histogram provides a good estimate for the accuracy of the reconstructed directions from HiSPARC data.

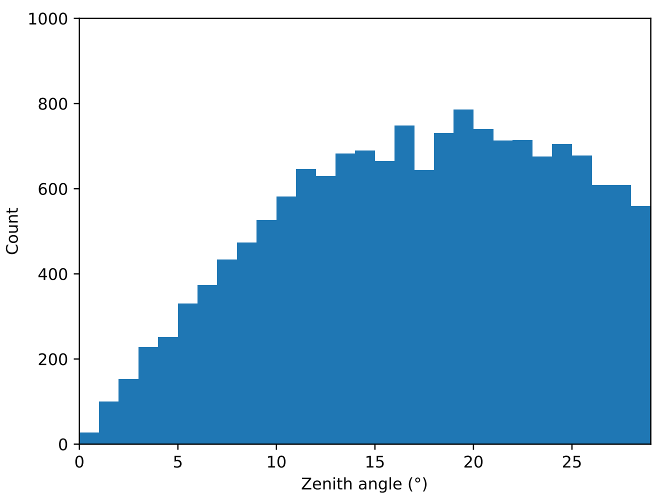

Fig. 12: Histogram of the reconstructed zenith angles with a clear gradual peak at ~20°. This provides a crude estimation of the reconstruction accuracy.

Arrival plane projections

At the end of each day, the HiSPARC software calculates the average signal strength of a single MIP, as recorded on that day. This is done by plotting a histogram of all of the pulseheights recorded on that day, the peak of which gives the most probable signal value for one particle [1]. By dividing the pulse integrals in each detector by the value for the single MIP, a rough idea of the number of particles that passed through the scintillator can be derived per shower. This is a useful estimate, because it gives a picture of the particle density throughout the array for each event. These particle densities are then projected onto the arrival planes of the cosmic rays, so that the lateral density distribution function can be applied, and the initial energy can be determined. The projection method is described below:

Using the zenith (θ) and azimuthal angle (Φ) calculated from the direction reconstructions, a unit vector parallel to the direction of incidence is defined with:

$$\vec{n} = \begin{pmatrix} \sin\theta \cdot \cos\phi \\ \sin\theta \cdot \sin\phi \\ \cos\theta \end{pmatrix} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \tag{7}$$

From here, a new set of axis e_1 and e_2 are defined, which is perpendicular to the normal vector from above.

$$\vec{e_1} = \begin{pmatrix} -n_2 \\ n_1 \\ 0 \end{pmatrix}, \quad \vec{e_2} = \begin{pmatrix} -n_1 \cdot n_3 \\ -n_2 \cdot n_3 \\ n_1^2 + n_2^2 \end{pmatrix} \tag{8}$$

$\vec{e_1}$, $\vec{e_2}$ and $\vec{n}$ together make up the new set of axes on which the particle densities are projected, with $\vec{e_1}$ and $\vec{e_2}$ lying on the shower front plane. So a detector at p = (x,y,z) would need to be projected in the direction of $\vec{n}$ over some distance d, such that it lies on the $\vec{e_1}, \vec{e_2}$ plane.

If the shower front plane is assumed to pass through the (0,0,0) mark on the x-y-z axes, then the dot product formula can be simplified so that d is calculated with:

$$d = \vec{n} \cdot \vec{p} \tag{9}$$

Hence the coordinates of a detector projected onto the $\vec{e_1} - \vec{e_2}$ is given using:

$$\vec{p}_{\text{projected}} = \vec{p} - d \cdot \vec{n} \tag{10}$$

and,

$$\left.\begin{aligned} e_1 \text{ coordinate} &= \overrightarrow{e_1} \cdot \vec{p}_{\text{projected}} \\ e_2 \text{ coordinate} &= \overrightarrow{e_2} \cdot \vec{p}_{\text{projected}} \end{aligned}\right\} \tag{11}$$

This method was used to define the $\vec{e_1}$ and $\vec{e_2}$ coordinates of all of the detectors involved in each event. The number of MIPs in each detector was also projected onto the plane.

Lateral density distribution

A cosmic ray vertically incident to the earth's surface grows laterally as new particles are produced when moving through the atmosphere. The density of particles is highest near the shower axis or the "core", and decreases when moving further away [4]. The number of particles produced, and the overall radius of the shower directly depends on the initial energy of ray. This fact can therefore be used to predict the energies of cosmic rays.

Using the projected particle densities from above, the position of the shower core is approximated onto the arrival planes. This approximation is done by fitting a lateral density distribution function, which describes the spread of the particle densities, onto the projected shower. The density function used in this study is a modified version of the Nishimura-Kamata-Greisen (NKG) formula, which states that the number of particles N at a distance from the shower core r is given by:

$$N(r) = A\left(\frac{r}{r_o}\right)^a \left(1+\frac{r}{r_o}\right)^b \tag{12}$$

where r_0 = 29.6, a = -0.566, b =-2.56 and A is a fit parameter that is dependent on the energy [1]. Hence by describing the distances between detectors and the number of particles in each, the value of A and the core position can be determined.

From here, the intial energy of the ray can be estimated by using:

$$\log(E) = c \cdot (\log(A) + d) \tag{13}$$

where c = 0.797, d = 17.62 and E gives the energy of the cosmic ray. Fig. 13 shows an example of the lateral density distribution function fit to projected particle densities.

Inclined showers need to travel a much longer path to reach the earth's surface, and so their particle numbers are reduced greatly [1]. This attenuation needs to be corrected for before the energies can be predicted. This is done using:

$$A_\perp = A \cdot \exp\left(p\left(\frac{1}{\cos\theta} - 1\right)\right) \tag{14}$$

where p = 6.937, θ is the zenith angle and $A_\perp$ is the attenuated fit parameter used in place of A in equation 13. The attenuation algorithm, however, only holds for showers with average zenith angles $< 30°$, and applying this to higher inclinations results in energy calculations beyond the GZK cut-off. Hence only the energies of rays with a zenith angle smaller than this value were reconstructed.

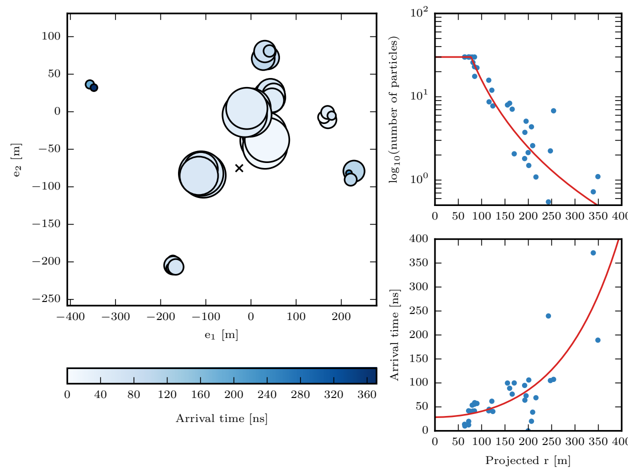

Fig. 13: The graph on the left shows the projected positions of the detectors that measured a signal- the size of the circle is dependent on the number of MIPS, and the colour is dependent on the arrival time of the particles. The shower core is denoted with the small "x" in the center. The top-right figure shows the particle density of the shower with increasing distance from the shower core, to which the lateral density function (red line) is applied. Typically, shower fronts begin to curve slightly when moving away from the core, as can be seen from the arrival times in bottom-right graph. A catenary function (red) was fit to this curvature. Figure courtesy of K. van Dam [6]

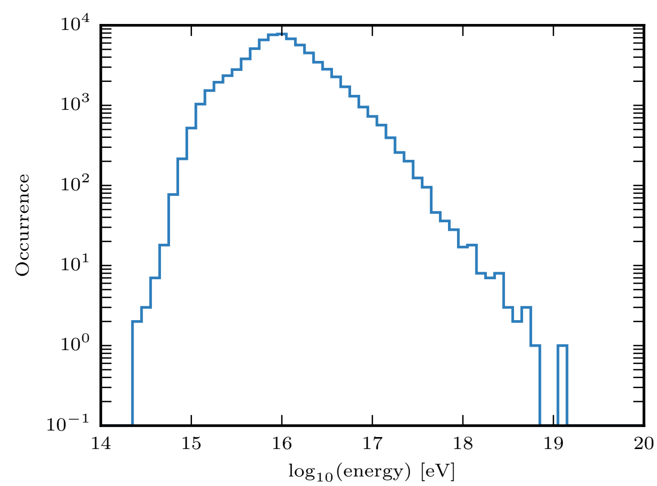

Fig. 14: Histogram of the reconstructed cosmic ray energies. The peak can be explained by the fact that below ~10^{16} eV, the effective surface area decreases; furthermore, there is a smaller chance of these lower energies being able to meet the trigger conditions of the HiSPARC detectors. The rarity of rays increases with energy, which explains the drop in occurrence after ~10^{16} eV. In the whole two year period, only one shower with an energy greater than 10^{19}eV was detected. Figure courtesy of K. van Dam [6].

3.2 Cosmic ray flux

Cosmic ray flux is usually given in terms of the time period of observation (~22 months), and the part of the sky visible, i.e. the effective solid angle and surface area subtended by the HiSPARC detector array. In theory, the field of view subtended by the HiSPARC array is 2π rad, but with an increase in path length and zenith angle, the attenuation of rays increases exponentially, leading to a large decrease in efficiency. Furthermore, only energies with a zenith angle <30° were reconstructed. Below 10^{17}, the solid angle was estimated from experimental data. The maximum solid angle (Ω) can be approximated to:

$$\Omega = 2\pi(1 - \cos(30)) = 0.84sr \tag{15}$$

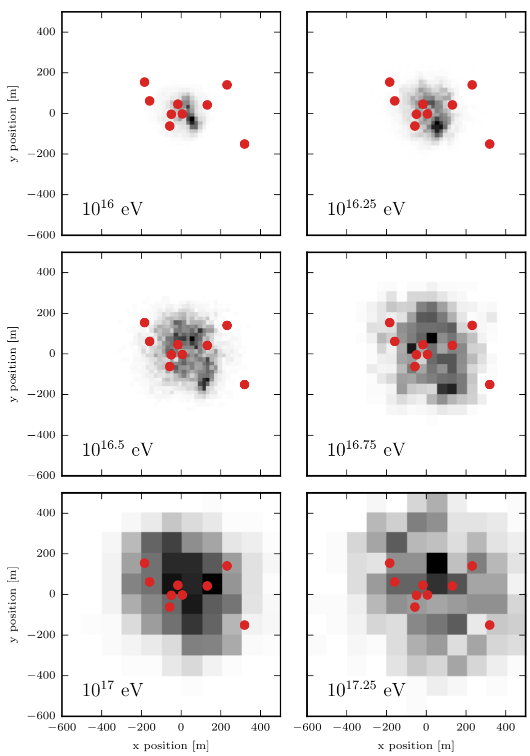

Fig. 15: Histogram of reconstructed core positions per energy bin projected back onto the x-y plane. The red dots show the positions of the detectors used in this study. The higher the density of cores, the darker the pixel. Figure courtesy of K. van Dam [6].

The effective surface area of the cluster was determined using the reconstructed core positions. Fig.15 shows the spread of core positions for different energies, re-projected back onto the earth's surface. The cores form an approximately circular shape, and the maximum radius increases with an increase an energy. From here, a "center of mass" for each of the projected energies was defined.

In order to determine the effective surface areas, the efficiency of the array needs to be known at each distance from the center of mass. With Fig.15, this can be thought of as the number of cores that fall on concentric circles, each with increasing radii from the center. Firstly, geometric effects caused by the increase in surface area for larger radii was removed by dividing the counts by 2πr. At r = 0, and up until some distance r_m, the efficiency is assumed to be 100%, i.e. detectors at these position are very unlikely to miss being hit by particles. Beyond this radius r_m, the efficiency decreases, and is best described with a modified Gaussian distribution. The efficiency can therefore be estimated by:

$$p(r,\theta) = \begin{cases} f(r_m, \alpha, \mu(\theta,\chi,\rho), \sigma, \lambda) & \text{for } r < r_m \\ f(r, \alpha, \mu(\theta,\chi,\rho), \sigma, \lambda) & \text{for } r \geq r_m \end{cases} \tag{16}$$

where μ, σ and λ are the mean, standard deviation and rate of distribution of the Gaussian functions, θ is the zenith angle and α, χ and ρ are further fit parameters. The full version of Eq.16 can be found in section 8.2 of [1].

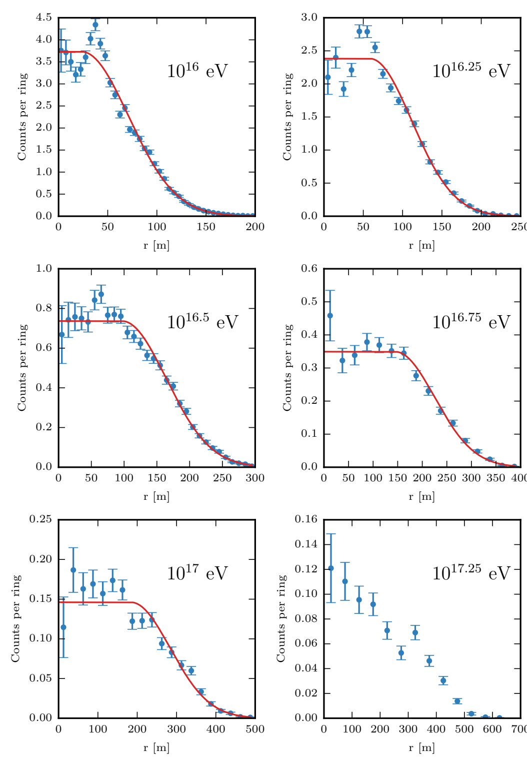

Fig. 16: Distribution of number of cores as a function of distance from the center of mass. The red lines show the Guassian fits. There is a clear plateau where the number of cores is approximately constant at the beginning. After r_m, the counts drop. Figure courtesy of K. van Dam [6].

This parameterization was applied to each energy in Fig.16, as seen from the red curves. Using these efficiency calculations, the effective surface area of each energy can be calculated by integrating the circumference at r, i.e.:

$$\mathrm{A} = \int_0^\infty 2\pi r \, \frac{p(r)}{p(0)} \, dr \tag{17}$$

The calculated effective surface areas were then interpolated for the intervening energies, and described using the quadratic function:

$$\mathrm{A}(E) = a(E - E_0)^2 \tag{18}$$

where a and E_0 are both fit parameters. The effective surface area increases quadratically with the energy of the cosmic ray.

By dividing out the time period (in seconds), and the effective solid angle and surface area as calculated above, the cosmic ray energy spectrum is derived.

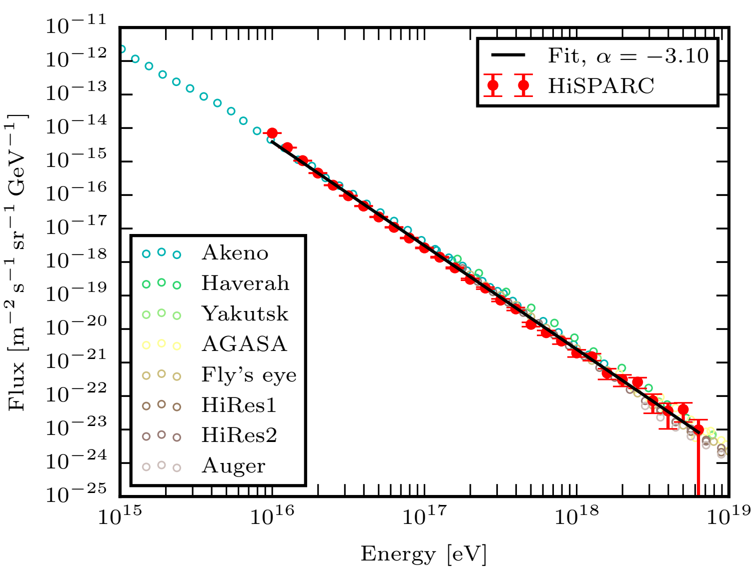

Fig. 17: The cosmic ray energy spectrum derived from HiSPARC data, compared with findings from other detector arrays. The reconstructed energies start just after the "knee" at 10^{15}eV and ends just after the ankle at 10^{18}eV. This turning point is not visible. The spectral index α = -3.10 matches literature values [2]. Figure courtesy of K. van Dam [6].

4 Discussion

Although the energy spectrum was successfully reconstructed and matched literature values, there still exists some possible errors, both within the methods used in this study, and in the HiSPARC electronics themselves. These potential limitations are discussed below.

Firstly, several cuts had to be made in the data:

1. Zenith angles $> 30°$ could not be reconstructed, as the attenuation correction was invalid for these higher angles.
2. Only events that involved the 4 "core" stations (501, 502, 503, 514) were reconstructed.
3. A minimum of 2 MIPs were required for stations involved in the direction reconstruction, and a minimum of 0.5 MIPs for the energy reconstructions. This way, random noise was eliminated.
4. Only stations which measured less than 30 MIPs were considered. Beyond this value, the corresponding voltage is too high for HiSPARC detectors to measure accurately.

9% of the overall data set did not pass through the cuts during the direction reconstruction stage, and a further ~28% had to be discarded since they had a zenith angle larger than 30°. The energies of almost all of the remaining data could be reconstructed.

Secondly, there were also uncertainties caused by the HiSPARC electronics themselves. The maximum ADC voltage that the detectors are capable of reading is 2.4 V. If a shower were to fall directly onto detectors, it can result in hundreds, if not thousands of particles, thereby being far too high to be accurately measured by HiSPARC detectors. Furthermore, with such huge showers, the differences between the Nikhef and commercial PMT bases becomes more pronounced. Random background muons are also indistinguashable from particles appearing later in shower, which again limits the amount of information captured by HiSPARC detectors. The uncertainity in pulse signals is also quite high, which can then affect the number of MIPs calculated in each detector. There may also be additional random errors caused by equipment, e.g. slight differences in cable lengths, arrival time delays etc.

And finally, the errors may also have been caused by the methodology. The stations at the Amsterdam Science Park form an irregular array, and when fitting the LDF algorithm, certain points were preferred over others. This can be seen in Fig.15, and Fig.16 where the projected cores are skewed, especially for lower energies. There also a corresponding deviation in the flux for lower energies in Fig.17. And lastly, the shower front was assumed to be a flat plane, but in reality, particles that "lag" behind and arrive later in the shower results in a slight curvature along the edges [2]. This is seen in the arrival times of particles far away from the four "core" stations, which would have affected direction reconstructions.

5 Conclusions

In this study, Amsterdam Science Park cluster, was used to calculate the flux of primary cosmic rays. Data from 11 four-detector stations was examined, and a significant disparity was found in the number of events recorded per day. This was resolved by examining fluctuations in weather and voltage gain in the PMTs, and finding a rough correlation. Timing offsets between stations in the data set (ranging from November 2018 to September 2020) was corrected for.

The direction of incoming cosmic rays was calculated by assuming an infinitely thin shower front using the arrival times of particles in each detector (triangulation). The differences in altitudes of the detectors was corrected for by performing repeated iterations of the triangulation procedure, and seeing what value they converged to. The reconstructed directions agreed with experimental data, and peaked at approximately 20°. The particle densities were then projected onto the arrival plane. Attenuation through the earth's atmosphere was corrected for. A modified version of the Nishimura-Kamata-Greisen function was applied to the projected data, and used to predict the position of the core and the original energy of the rays. Because of the nature of the attenuation correction formula, only rays with a zenith angle $<30^{\circ}$ were reconstructed.

From here, the solid angle and effective surface area subtended by the HiSPARC array was calculated. The former was approximated to have a maximum value of 0.84sr. In order to calculate the effective surface area, a "center of mass" was determined for the core positions of each energy value, and the efficiency at regular intervals from this center was mapped. By integrating with reference to these distances, the effective surface areas were calculated, and interpolated for the intervening energy values. By dividing the reconstructed energies with the time period, effective solid angle and surface area, the flux was calculated and the cosmic ray energy spectrum was mapped from ~10^{16}eV to ~10^{19} eV. The spectral index (=-3.10) matches literature values.

6 Acknowledgements

This report contains the research that I did during my 5 month internship with the HiSPARC department at Nikhef. I would like to thank Bob van Eijk for giving me the opportunity to have such a wonderful experience at Nikhef. I have learnt more than I could ever fill in these few pages. I also want to thank my daily supervisor Kasper van Dam for letting me join him in this project that was part of his PhD thesis. His patience and guidance when it came to teaching me such difficult topics was invaluable. And finally, I want to thank the whole HiSPARC team for the morning chats and words of wisdom that I will remember for the rest of my life.

References

[1] Kasper van Dam, B van Eijk, DBRA Fokkema, JW van Holten, APLS de Laat, NG Schultheiss, JJM Steijger, and JC Verkooijen. The hisparc experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 959:163577, 2020.

[2] DBRA Fokkema. The hisparc cosmic ray experiment: data acquisition and reconstruction of shower direction. 2012.

[3] Stephan Runderkamp. Investigating post-lhc hadronic interaction models and their predictions of cosmic ray shower observables. 2018.

[4] Johannes Martinus Cornelis Montanus et al. *The observability of jets in cosmic air showers*. PhD thesis, Universiteit van Amsterdam, 2017.

[5] Andreas Haungs. Cosmic rays from the knee to the ankle. *arXiv:1504.01859*, 2015.

[6] K. van Dam. HiSPARC, A Scintillating Experiment. 2021. *in preparation*.