UNIVERSITY OF LEEDS

A Search for Photons of Energy above 6×10^{18} eV Using Data from the Water-Cherenkov Detectors of the Pierre Auger Observatory

Lu Lu

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Physics and Astronomy

September, 2014

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

Firstly I would like to express my gratitude to Prof. Alan Watson, who has invented the entity method and guided me through the darkest periods. He could have relaxed and enjoyed life on the golf course but instead he chose to spend the summer correcting prepositions and articles. I would also like to thank Dr. Johannes Knapp, who took me on the beautiful journey to UHECR and taught me the attitude of criticism while appreciating nature. Then I would like to send my thanks to Prof. Karl-Heinz Kampert, who kindly allowed me to stay in his group in Wuppertal for 18 months when the high energy group in Leeds nearly vanished.

Next I would like to thank for the entire Pierre Auger Collaboration for creating the excellent atmosphere to work in. In particular, I would like to thank the Auger groups in Wuppertal, Santiago de Compostela and La Plata for the valuable time shared together. Special thanks go to Dr. Ronald Bruijn, Dr. Carla Bleve, Dr. Inés Valiño, Heather Cook, Nicole Krohm, Dr. Julian Rautenberg and Dr. Piotr Homola for inspiring conversations.

It was very depressing time seeing everyone leaving the high energy astrophysics group in Leeds halfway through my PhD. I am very grateful for receiving help from Dr. Julian Pittard, Dr. Stella Bradbury, Prof. Tom Hartquist and Prof. Peter Olmsted for making it possible for me to move to Germany. I am also very thankful that Prof. René Oudmaijer took me as his student from the Leeds side.

I am lucky to have all my amazing office mates and colleagues in Leeds and Wuppertal. Still feeling guilty for the trouble I brought to people in the astroparticle physics group in Wuppertal due to my nearly non-existing German.

The list could go on endlessly, but next and most importantly I would like to show my greatest respect to my family. Their support and encouragements helped me to pursue my dream of studying astroparticle physics from day one. Thanks for believing in me and your unconditional love.

Last but not the least, is to thank my fiancé, Colin Baus, whom I met because of cosmic rays and Auger. Thanks for understanding me on so many levels and putting up with me always. You are my muse and best friend.

Abstract

It is important to develop methods to search for photons in the ultra-high energy cosmic ray beam. Photons above 6×10^{18} eV can be created if the highest energy cosmic rays are dominated by light nuclei (such as protons) but not if the primaries are dominated by heavy nuclei such as iron. In this thesis a new method is developed to attempt to distinguish showers initiated by photons from those initiated by hadrons.

Data from January 2004 to March 2014 that have zenith angle $0 < \theta < 60^{\circ}$ are used in the analysis. This is the first time that a parameter based on the surface detectors from the Pierre Auger Observatory is valid for the entire zenith angle range and even for energies below 10^{19} eV. 81, 9, 2, 0 and 0 photon-like events were found for the energies above $10^{18.8}$, 10^{19}, 2×10^{19}, 3×10^{19} and 4×10^{19} eV respectively (all photon candidates are considered as hadronic background). The upper limit (95% c.l.) of the photon flux compared to previous analyses that were obtained by other methods has been improved by a factor of 3.

Zusammenfassung

Die Suche nach Photonen in ultrahochenergetischer kosmischer Strahlung und die Verbesserung der Suchalgorithmen ist von besonderer Bedeutung. Photonen mit einer Energie über 6×10^{18} eV können erzeugt werden, wenn der Anteil an leichten Kernen (wie z.B. Protonen) in der kosmische Strahlung überwiegt, jedoch nicht, wenn schwere Kerne, wie Eisen, dominieren. In dieser Arbeit wird eine neue Methode zur Unterscheidung zwischen von Photonen erzeugten Luftschauern und von Hadronen erzeugten Luftschauern vorgestellt.

Hierzu werden Daten des Pierre Auger Observatoriums von Januar 2004 bis März 2014 mit einem Zenitwinkel von $0 < \theta < 60^{\circ}$ analysiert. Zum ersten Mal, ist aufgrund eines auf dem Oberflächendetektorensignal basierenden Parameters der Einsatz in diesem breiten Winkelbereich und zusätzlich für Energien unterhalb 10^{19} eV ermöglicht worden. In den Energiebereichen über $10^{18,8}$, 10^{19}, 2×10^{19}, 3×10^{19} und 4×10^{19} eV wurden jeweils 81, 9, 2, 0 und 0 photonähnliche Ereignisse gefunden. Trotz der konservativen Annahme, dass diese von Hadronen stammen, konnte die obere Vertrauensgrenze (95% c.l.) des Photonflusses um einen Faktor 3 gegenüber vorheriger Methoden verbessert werden.

摘要 (Zhai Yao)

找到探索超高能光子的方法是很重要的。如果超高能宇宙射线是由轻核（例如质子）主导，光子可以通过 GZK 反应产生而且达到 6×10^{18} eV 的能量。如果超高能宇宙射线是重核组成则不然。这项工作提出一个新方法来鉴定由光子带来的大气广延簇射。

本文分析了从 2004 年 1 月到 2014 年 3 月期间在奥格实验台（Pierre Auger Observatory）采集的数据。天顶角使用的范围为 $0 < \theta < 60°$。这是史上第一次有基于表面探测器数据的参数能够在如此广角度范围内使用，也是第一次使表面探测器参数能够寻找能量低于 10^{19} eV 的光子。在能量分别高于 $10^{18.8}$，10^{19}，2×10^{19}，3×10^{19}，4×10^{19} eV 的范围，一共找到 81，9，2，0，0 疑似光子事件。即使假设所有寻找到的时间是强子背景，超高能光子的流量上限（95% CL）仍然可以降低到用其他方法得到结果的三倍以下。

We are the cosmos made conscious and life is the means by which the universe understands itself.

(Brian Cox)

Table of Contents

List of Tables

List of Figures

Personal Contributions

The work on the entity method is original and first proposed by A. A. Watson in 2011. The method has been developed and is shown to be efficient for photon-hadron separation of zenith angle 0 - 60˚ (Chapter 4, 5, 6 and 7). Improved photon limits above 6×10^{18} eV are reported. In the future, the method is likely to contribute to the paper on the results of photon-searches from Auger.

The original work in the thesis, carried out by the author:

<1>. The parameterisations of the mean time-trace and uncertainties (Chapter 5). The shape of the trace is fitted with a partly-normalised Landau function and the uncertainties are fitted with a Moyal-like function.

<2>. The entity likelihood parameter is defined for photon-hadron separations on the event-by-event basis.

<3>. The probability that protons have a neutral pion as the leading particle from the first interaction was investigated (Chapter 8).

<4>. Through section 9.4 it was shown that the mean of thinned showers agrees with the unthinned case. However, the artificial fluctuations make the spread of the distributions larger for the entity parameter. The trace of each individual station is distorted in the thinned simulations.

<5>. Arriving-particle distributions including the energy spectrum, muon fraction, number of particles are explored using thinned and unthinned simulations (section 4.3, 4.4 and 6.5.2.2).

The thesis is of benefit to the Auger collaboration due to:

<1>. The entity method could be used for monitoring the quality of photomultipliers (section 6.4.1 and 6.4.2).

<2>. Outliers found using the entity method show problems with shower reconstructions. A possible solution for improving the start time determination is described in the section 6.3.

<3>. The method could be used for the search of exotic events (section 9.3).

<4>. The method could be combined with other parameters (e.g. X_{max}) for the search of photons with lower energies (section 9.1 and 9.2)

<5>. A class of events has been identified with features of pre-pulses (section 6.4.3). This could be related to shower physics.

List of Abbreviations

AGN	Active Galactic Nuclei
CDAS	Central Data Acquisition System
CIC	Constant Intensity Cut
CMB	Cosmic Microwave Background
CR	Cosmic Rays
EAS	Extensive Air Shower
EeV	10^{18} eV
FADC	Flash Analogue to Digital Converter
FD	Fluorescence Detector
GPS	Global Positioning System
GZK	Greisen-Zatsepin-Kuzmin
ICRC	International Cosmic Ray Conference
LDF	Lateral Distribution Function
LHC	Large Hadron Collider
LPM	Landau, Pomeranchuck and Migdal
NKG	Nishimura-Kamata-Greisen
PMT	Photomultiplier Tube
SD	Surface Detector
SHDM	Super Heavy Dark Matter
TD	Topological Defects
TOT	Time Over Threshold
UHE	Ultra-High Energy
UHECR	Ultra-High Energy Cosmic Ray
VEM	Vertical Equivalent Muon

Chapter 1. Ultra-high energy cosmic rays and photons

Cosmic rays are extra-terrestrial charged particles that were first discovered by Victor Hess in 1912[1]. He measured the rate of production of ionisation as a function of height using electrometers during a flight in a hot air balloon that rose to over 5 km. Later, experiments were carried further out to measure the dependence of ionisation (ion pairs $cm^{-3}\ s^{-1}$) of the atmosphere on altitude, air pressure and latitude confirming that the radiation creating the ionisation is dominated by sources outside of the earth. The invention of the Geiger counter and coincidence circuitry enabled, in the late 1930s, the discovery of extensive air showers[2], which are cascades of secondary nuclei generated through interactions between the primary cosmic ray and the particles in the atmosphere. In detailed studies of cosmic-ray radiation, cloud chambers were used and led to the discovery of the positron[3], the muon[4] and charged pions[5], which are essential discoveries leading to the modern particle physics.

Cosmic rays have a vast range of energies and the flux received at earth decreases, following an approximate power law spectrum. At energies lower than ~ 100 TeV, cosmic rays can be measured directly with equipment carried on balloons and satellites and the mass composition is relatively well-understood. But for cosmic rays with higher energies and lower flux, one has to observe them indirectly via the detection of the cascades created through interactions between the cosmic ray and the particles in the atmosphere. Ultra-high energy cosmic rays (UHECRs) typically carry energy around 10^{19} eV and arrive earth at the rate of only about one particle per km^2 per year. Ground-based arrays of detectors covering large areas and equipped with water-Cherenkov detectors or scintillators have been built for measuring the footprint of the secondary particles on the ground. Fluorescence telescopes are used to observe the cascade development in the atmosphere. In addition, the radio emissions from propagations of the charged particles from air showers have been detected at several ground-based arrays and by the balloon-based project, ANITA above Antarctica. In the future, satellites will be used to look down to the atmosphere to detect the fluorescence light from the air shower.

The highest energy given to particles accelerated at the most powerful accelerator is only a few TeV (at the Large Hadron Collider at CERN) while the highest energy cosmic ray observed so far has an energy > 10^{20} eV (~ 500

TeV in the centre of mass frame for proton-proton collision). UHECR come from sites which exhibit the most extreme conditions in the Universe. The mystery of how these particles are accelerated remains today.

The mass composition and sources of the highest energy cosmic rays are unknown. The cosmic rays interact with the cosmic microwave background (CMB) and with galactic and inter-galactic magnetic fields during propagation from the sources. The existence of ultra-high energy neutrinos and photons are predicted in various astrophysical scenarios, such as the interaction between the charged cosmic rays with the cosmic microwave background photons.

In this chapter, the cosmic-ray spectrum will be discussed (1.1), the possible sources and interactions that occur during propagation will be described (1.2), the physics of extensive air showers will be explained (1.3), and the importance of studying UHE photons will be introduced (1.4).

1.1 Energy spectrum of cosmic rays

The differential cosmic-ray flux falls by 30 orders of magnitude over 11 decades of energy with an approximate power-law

$$dN / dE \propto E^{-\alpha} \qquad 1.$$

α is the index of the spectrum and shows the steepness of the falling of the flux. A combined energy spectrum from measurements made from balloon, space and ground-based observations is shown in Figure 1.1-1. The flux has been scaled by $E^{2.5}$for a better visualisation of some of the detailed features.

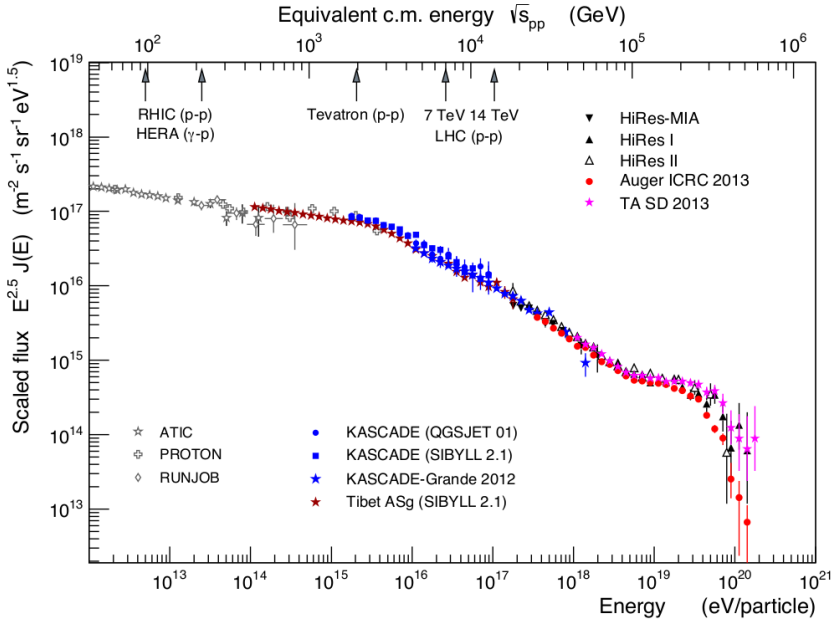


Figure 1.1-1 Energy flux spectrum of cosmic rays from various experiments[6]

The flux drops at the rate of ~ 2.6 orders of magnitude per energy decade from the solar energetic particle origin (the lowest energy region of the spectrum) until to the 'knee', E = 3 × 10^{15} eV. The spectrum gets steeper with α increasing to 3. Cosmic rays at the knee energy arrive at earth at the rate of ~ 1 particle m^{-2} yr^{-1} and are studied with ground based experiments such as KASCADE-Grande[7] in Karlsruhe and the ARGO experiment[8] in Tibet. The reason for the knee feature is now generally accepted to be the change in the mass composition from light to heavy nuclei in the galactic cosmic rays. Light nuclei have relative higher rigidity and escape the galaxy at a lower energy than heavy nuclei. For example the cut-off energy of helium is twice as large as hydrogen due to the rigidity effect. A 'second knee' has been reported[9] to occur at ~ 4 × 10^{17} eV which makes the spectrum even steeper and corresponds to the cut-off energy of uranium[10] (Z = 92). At ~ $10^{18.5}$ eV (the 'ankle'), the spectrum index starts to decrease from α = 3.5 to 2.7. This could mark the transition between galactic and extragalactic cosmic rays[11] but could also be the signature of the interaction between ultra-high energy protons with the cosmic microwave background which then produces electron-positron pairs. The latter assumption requires proton-rich scenario. Beyond energy ~ $10^{19.5}$ eV, the spectrum becomes even steeper. This is the region where pion-production and photo-disintegration starts to be important. The flux is expected to be ~ 0 if there is no source located in the near Universe within ~ 100 Mpc. The interactions are described in section 1.2. However, the drop of the spectrum could also be due to a feature of the source if this is the maximum energy of acceleration from the source[12].

1.2 Sources and propagation

The journeys of UHECRs from the source to the end of their lives are summarised in Figure 1.2-1. Red curves represent the Cosmic Microwave Background (CMB) photons and act as 'targets' of collisions in scenario (1) and (3), which are discussed in this section. The proton from the (2) case reaches earth and interacts in atmosphere, which is described in details in section 1.3.

Cosmic rays propagating in the galaxy go through spallation and suffer deflections due to the galactic magnetic field (~3 μG). A typical residence time of cosmic rays with GeV energy in the galaxy is measured as 15 × 10^6 years from the CRIS experiment[13]. The energy density of cosmic rays is ~ 1 eV cm^{-3} if one assumes the flux over energy of cosmic rays follows an isotropic and uniform distribution in space. The volume of the Milky-Way is

can be approximated as a cylinder with diameter 26 kpc and height 0.3 kpc, which means the total energy of cosmic rays in the galaxy is ~ 5×10^{54} ergs. It was pointed out[14] that supernovae which happen ~ twice per century in our galaxy only need < 10% of their kinetic energy to produce the energy density of cosmic rays. A recent publication from the Fermi collaboration[15] showed the first indirect evidence that cosmic rays are produced from supernovae via hadronic interactions.

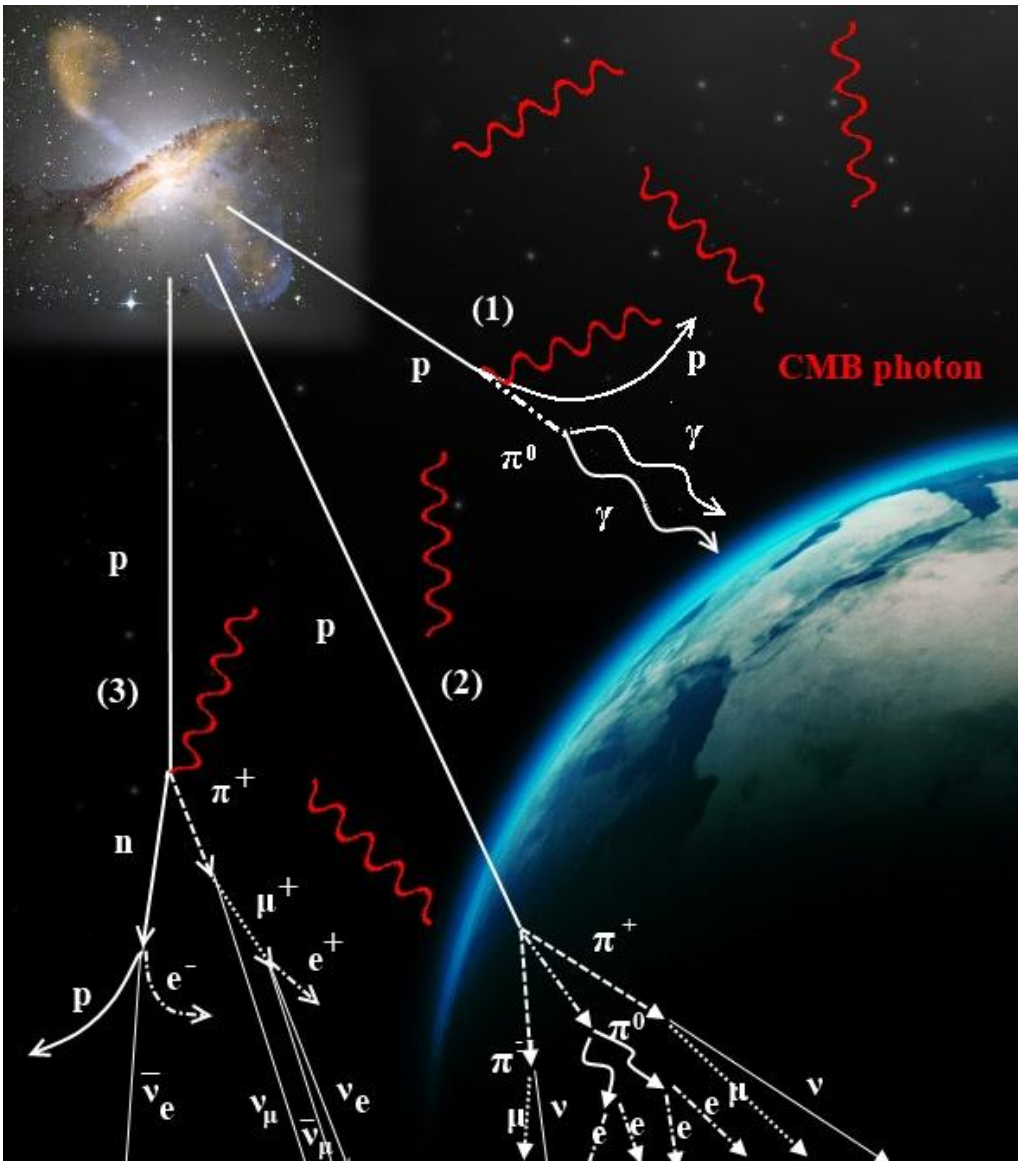


Figure 1.2-1 An artist's impression of the journeys of UHECRs. (1), (2) and (3) are all from pion-interaction of proton and cosmic microwave background photons. Scenario (1) produces UHE photons, (2) doesn't interact with CMB but creates an air shower in the atmosphere (3) produces UHE neutrinos.

The stochastic process of the first order Fermi acceleration (also known as the diffuse shock acceleration) of particles in the shock waves and turbulent magnetic fields has been proven to be sufficient of producing a power law spectrum with index $\alpha = 2$[16,17,18,19]. Although the mechanism to accelerate

particles to 10^{20} eV is still unclear, the maximum energy of charged particles accelerated from an object with radius R and magnetic field B can be estimated[20] by requiring the gyro-radius of the particle to be within the size of object

$$E_{max} \cong 10^{18} \text{ eV } Z \text{ (R / kpc) (B / } \mu\text{G)} \quad 2.$$

The maximum energies for various known astrophysical objects are shown in Figure 1.2-2. Only objects beyond the solid line are able to accelerate protons to 10^{20} eV.

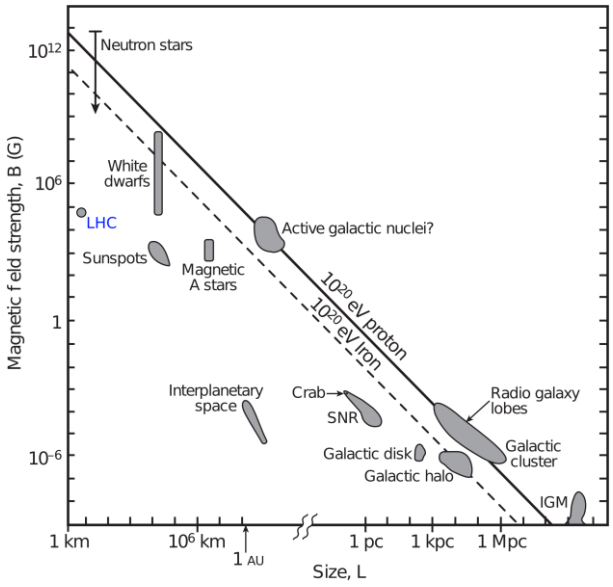


Figure 1.2-2 Hillas plot[20] of maximum energies of particles accelerated from various objects[21]

As a reference the radius of the LHC is ~ 8 km and the magnetic field is ~ 10^5 G. To accelerate iron to 10^{20} eV, one has to build a collider with radius on the order of the distance from the planet Mercury to the Sun.

Alternatively the particles might be accelerated not directly from sources. The popular models are top-down models and the Z-burst model. The common feature of those theories is that they predict a high fraction of UHE photons at the highest energies. Details of such models are described in section 1.4.1.

The Universe is filled with black body radiation at a temperature of ~ 2.7 K with an energy density of ~ 0.25 eV m^{-3}. The average energy of the CMB photons is ~ 6×10^{-4} eV. As shown in Figure 1.2-3, protons interact with photon with different wavelengths through pair-production

$$p + \gamma \rightarrow p + e^- + e^+ \quad 3,$$

and pion-production

$$p + \gamma \rightarrow \Delta^{+} \rightarrow p + \pi^{0} \rightarrow p + \gamma + \gamma \quad 4,$$

$$p + \gamma \rightarrow \Delta^{+} \rightarrow n + \pi^{+} \rightarrow n + \mu^{+} + \nu_{\mu} \quad 5.$$

Then muons decay to electrons and more neutrinos are produced (could also be from neutron decay).

The energy threshold of photons measured in the laboratory using fixed target protons is found to be ~ 1 MeV for pair-production and ~ 145 MeV for pion-production. The threshold energy for protons to interact with cosmic background photons can be calculated by applying Lorenz transformation,

$$\varepsilon = \varepsilon_0 \, \Gamma \left(1 + \frac{v}{c} \cos\theta\right) \quad 6,$$

where ε is the energy of photons in the laboratory frame, ε_0 is the mean energy of the cosmic background photons. Γ is the Lorenz factor of the cosmic protons with ultra-high energy. In the simplest scenario one could assume θ is 0˚ and v / c ~ 1, therefore the equation becomes

$$\varepsilon = 2\varepsilon_0 \, \Gamma \quad 7.$$

The Lorenz factor Γ of pair-production and pion-production are ~ 10^9 and ~ 10^{11} respectively. Therefore indicative energies for pair-production and pion-production by protons to interact with CMB photons are ~ 10^{18} eV and ~ 10^{20} eV. These calculations are based on the assumption of a head-on collision between proton and photon with the mean energy 6 × 10^{-4} eV. Furthermore the energy distribution of the CMB is not accounted for.

The cross-section of the pair-production at relativistic energies is ~ 10 mb (10^{-30} m^2). Since the matter density of CMB photon is ~ 400 cm^{-3}, the mean free path is in the order of ~ 0.1 Mpc[22]. The energy loss per interaction is only ~10^{-3} of the energy of the proton so the attenuation length is ~ 1000 Mpc when the proton has energy 10^{19} eV. The cross-section of pion-production is ~ 0.25 mb (2.5 × 10^{-32} m^2) and on average for a proton travelling one interaction length (~ 8 Mpc), the energy that carried away by pions is ~ 1/5 of the energy of the primary proton. The attenuation length for a proton is thus as the order of tens of Mpc and corresponds to a lifetime of ~ 10^8 years. This means that if we observe cosmic rays with energy > 10^{20} eV and if they are protons, they have to originate from within tens of Mpc from our galaxy. This argument is known as the GZK-effect[23,24]. The processes discussed also take place through interactions with other photon fields (for example infrared / radio / UV backgrounds) but at different energies. In Figure 1.2-3, the attenuation lengths[25] of proton through various interactions are shown.

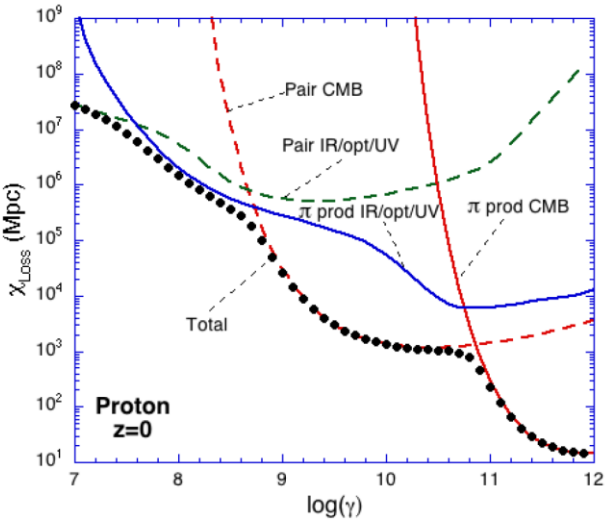


Figure 1.2-3 Attenuation length of pair-production and pion-production for interactions between proton and cosmic background photons with various wavelengths[25]

If the cosmic-ray has heavier mass A >>1, it undergoes pair-production and photo-disintegration (photon energy is ~ 30 MeV in the rest frame of proton) following:

$$A + \gamma \rightarrow A + e^+ + e^- \qquad 8,$$

$$A + \gamma \rightarrow (A - 1) + n \qquad 9,$$

$$A + \gamma \rightarrow (A - 2) + 2n \qquad 10.$$

Each nucleon can be treated as the superposition of Z protons and A-Z neutrons with the total energy ~ 1/A. Figure 1.2-4 shows the fraction of elements that survive after travelling distance D[21].

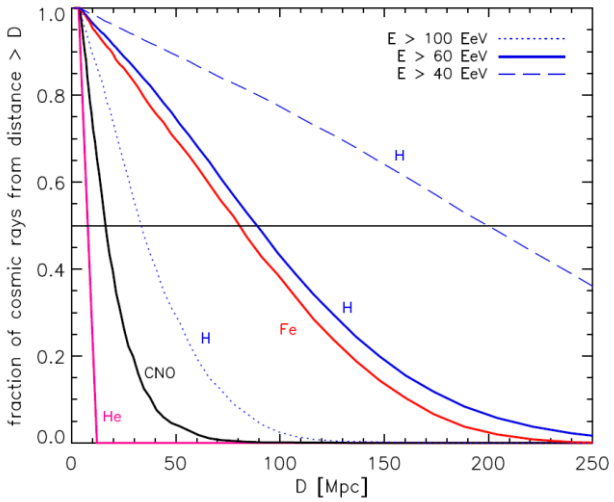


Figure 1.2-4 Fraction of nucleons survive after travelling distance D for three different initial energies[21]

The secondary products of the GZK-interactions carry important information of the primary mass of the cosmic ray. Some of the example interactions are illustrated in Figure 1.2-1. Scenario (1) produces UHE photons with energy ~ 20% of the primary proton energy. Details of how the photons propagate to the earth are introduced in section 1.4.2 and identifying them in the cosmic ray flux is the main purpose of the thesis. Scenario (3) generates neutrons which have a lifetime ~ 900 s in the rest frame and travel ~ 1 Mpc at 10^{20} eV before decaying. Neutrinos are generated through the neutron decay

$$n \rightarrow p + e^- + \nu_e \qquad 11.$$

The energy of the neutrino is typically 10^{-3} of the energy of the neutron. Therefore for neutrons from the GZK-interaction between proton and the CMB, the energy of the neutrinos is ~ 10^{16} eV.

Neutrinos can also be generated via the decay of pions as described in equation 5 and by the decay of the produced muons

$$\mu^+ \rightarrow \nu_e + \bar{\nu}_\mu + e^+ \qquad 12.$$

The energy of the neutrinos is ~ 25% of the energy of pions and therefore ~ 10^{18} eV.

For heavy elements, neutrinos come from the decay of neutrons that were produced in photo-disintegration and carry energy ~ 10^{16} eV. When E/A is larger than the GZK-threshold of proton, the heavy nuclei could undergo pion-production and produce neutrinos the same way as discussed in the proton case. The energy spectrum of neutrinos from cosmic rays with primary of proton/iron[26] is shown in Figure 1.2-5. The primary energy is $10^{21.5}$ eV, which is sufficient for iron to experience photo-disintegration and pion-production.

The suppression feature observed in the energy spectrum using Auger data is at energy $10^{19.5}$ eV, which arguably could be due to the GZK-interaction of light nuclei, the photo-disintegration of heavy nuclei or just because of the limitations of the accelerations from sources. The measured spectrum has been fitted[27] with various astrophysical scenarios and is shown in Figure 1.2-6.

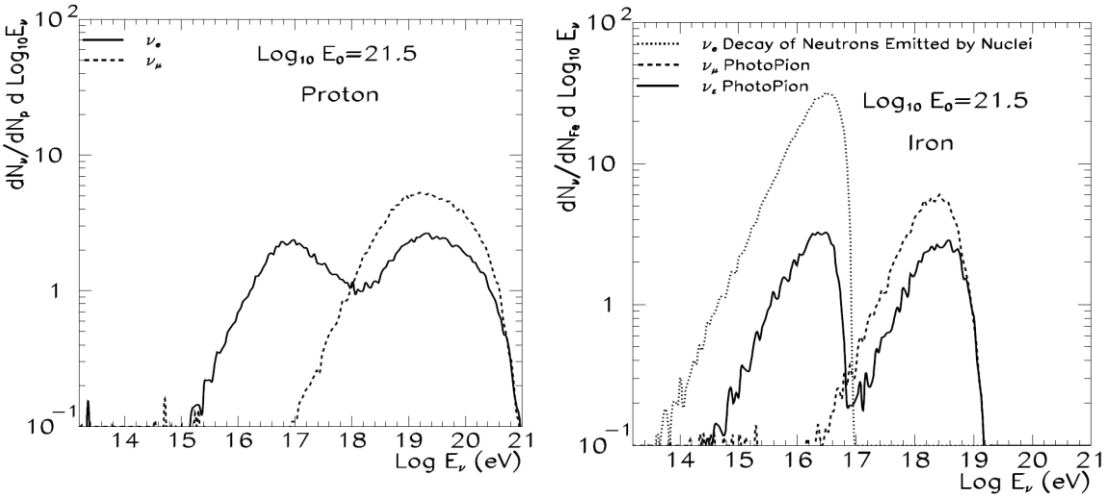


Figure 1.2-5 Energy spectrum of neutrinos produced from interactions between proton (left) / iron (right) and the cosmic microwave background photons. The cosmic ray has the primary energy $10^{21.5}$ eV and travelled a distance of 300 Mpc. For proton, the lower energy electron neutrinos are from the decay of neutrons that were produced in the pion-production and the higher energy electron / muon neutrinos are from the charged pions. For iron, the low energy electron neutrinos (dotted line) are from the decay of neutrons that created from photo-disintegration. The low energy electron neutrinos are also from decay of neutrons but the neutrons are created through pion-production. The high energy electron / muon neutrinos are from pion decays.[26]

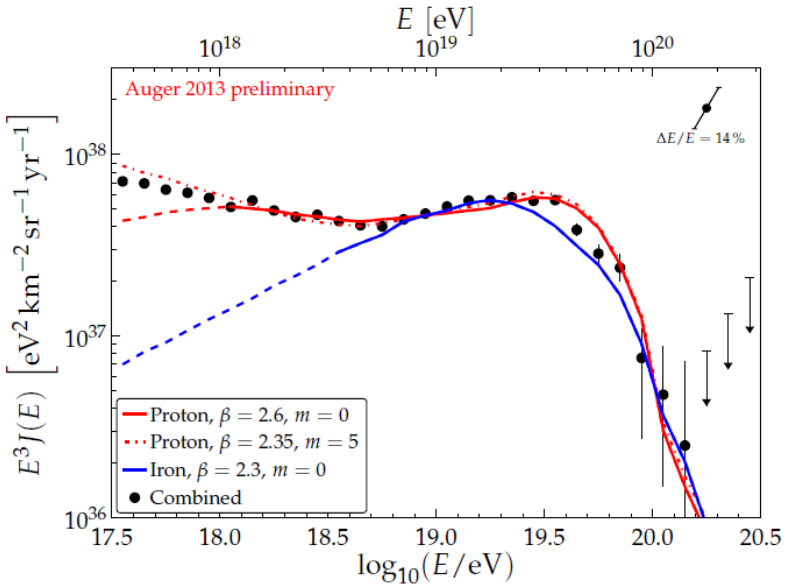


Figure 1.2-6 Auger spectrum fitted with various astrophysical scenarios[27]. Protons (red) interact with CMB photons to produce e^-/e^+ pairs and pions. β is the spectrum index of the source injection spectrum and m is the source evolution parameter, which is related to how the density of the source evolves with the redshift via $(1 + z)^m$.

The spectrum is the combined result using data from surface detectors and fluorescence detectors. The model predictions are from proton and iron simulations[28,29]. β is the spectral index of the source injection spectrum and

m is the source evolution parameter. The 'ankle' in the spectrum is due to pair-production in the pure-proton scenario and due to the transition of cosmic rays from galactic to extragalactic in the pure-iron scenario. All three scenarios can fit well to data and it is clear that without knowing the mass composition of the cosmic rays, it is impossible to understand which model is correct.

UHE cosmic rays do not lose large fractions of their energies while travelling in the galaxy but they could be deflected in the magnetic field. The Larmor radius describes the radius of the charged particle travelling in the magnetic field

$$r_L = 110kpc \cdot \frac{1}{Z} \cdot \frac{E}{10^{20}eV} \cdot \frac{\mu G}{B} \qquad 13.$$

If one assumes the magnetic field is constant perpendicular to the travelling path and the galactic magnetic field is ~ 3 μG. A proton with energy 10^{18} eV then has a Larmor radius ~ 300 pc, which is about the thickness of the galactic plane. The effect of the magnetic field is not significant if the Larmor radius is at the same order with the travelling distance of the particle. The deflection angle α for a particle that travels a distance d typically is

$$\alpha \cong \int \frac{\mathrm{dL}}{\mathrm{r_L}} = 0.52° \cdot \mathrm{Z} \cdot \left(\frac{\mathrm{E}}{10^{20}\mathrm{eV}}\right)^{-1} \cdot \left(\frac{\mathrm{B}}{10^{-9}\mathrm{G}}\right) \cdot \left(\frac{\mathrm{d}}{\mathrm{Mpc}}\right) \qquad 14.$$

So for proton with energy 10^{20} eV travelling in the magnetic field of μG and distance of kpc, the deflection is ~ 1˚, which is in the same order as the proton travels in the magnetic field of nG and distance Mpc. Cosmic rays with higher charges suffer large deflections and therefore it is more difficult to study the sources from the arrival directions. The importance of the mass composition study has been demonstrated again. Figure 1.2-7 shows a classic[20] illustration of the deflection of proton and iron in the galaxy.

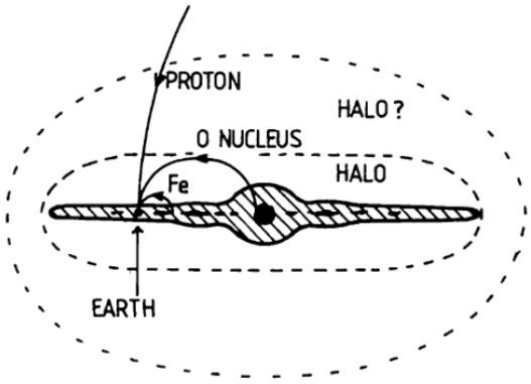


Figure 1.2-7 Deflection of proton and iron in the galactic magnetic field[20]

Cosmic rays that manage to travel to the earth will then enter the atmosphere and interact. They generate cascades of secondary particles in the air as shown in the second scenario of Figure 1.2-1. In the next section, the developments of the air shower are discussed and the recent result on the mass composition of the cosmic rays using data from the Pierre Auger Observatory is shown.

1.3 Extensive air showers

As introduced in section 1.1, the rate of arrival of cosmic rays with energy above 10^{14} eV is too low to be detected by direct measurements made using balloons or satellites. When high energy cosmic ray particles enter the atmosphere, they interact with nuclei in the air and produce cascades of secondary particles. The phenomenon has also been observed in cloud chambers when a low energy cosmic ray passing through. Shown in Figure 1.3-1, a proton interacts with lead plates [30] and generates cascades of secondary particles. The ionisation processes create tracks in the cloud chamber which are similar to what can be observed during the propagation of UHECR in the atmosphere. The first interaction typically takes place at an altitude of 15 to 35 km (~ of 50 g cm^{-2} vertically) which depends on the energy and the type of the primary particle. The secondary and tertiary particles will keep interacting with nuclei in the air and eventually form an air shower. An example of an air shower simulated[31] using a proton primary with energy 10^{15} eV is also shown in Figure 1.3-1. The similarity between the cloud chamber image and the simulated shower is remarkable.

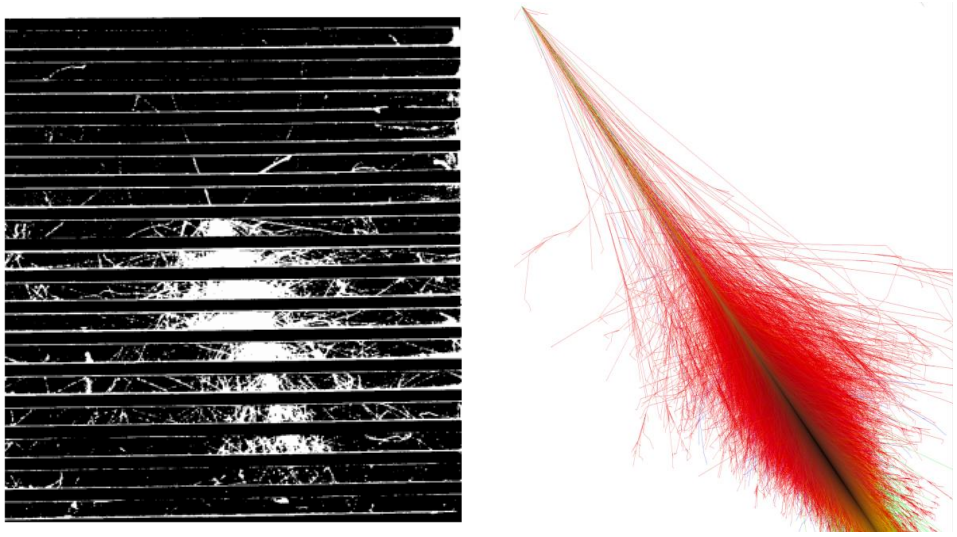


Figure 1.3-1 Cascades from a cloud chamber[30] and from simulation of air shower[31]

The depth at which the number of particles reaches the greatest value is defined as the depth of shower maximum, X_{max}. At ground level, the footprint of the shower can be observed using particle detectors. For UHECR such as a vertical proton with energy ~ 10^{19} eV, the footprint is about 10 km^2. An extensive air-shower (EAS) describes the development of secondary particles and carries properties of the primary cosmic ray, such as the direction, energy and mass composition.

Air showers initiated by protons and nuclei are composed of hadronic, muon, electromagnetic and neutrino components. The neutrino contribution to the shower is typically <4% for air showers with zenith angle less than 60˚. Details of how each component could be detected calorimetrically are briefly discussed in section 2.3.

For the electromagnetic components in the shower, there are mainly three processes for energy loss, namely pair-production of photons, and bremsstrahlung and ionisation losses of electrons and positrons**.** The cross-section for the pair production is ~ 60 mb, which is 7/9 of the cross section for bremsstrahlung. The opening angle for electrons and positrons has a typical value $\theta \sim mc^2 / h\nu$ through every pair-production covers a large range of angles, where m is the mass of the electron and hν is the incident energy of the photon. The electrons and positrons lose energy by ionisation at the rate of ~ 2 MeV g^{-1} cm^2 while at the same time radiating photons through bremsstrahlung. The radiation length of electrons in air is λ_r ~ 37 g cm^{-2}, which defines the exponential loss length of the energy by bremsstrahlung. The relationship between the energy E of the electron (positron) and the depth of the atmosphere it has travelled through (x) is

$$E=E_0\, e^{(-x/\lambda_r)} \qquad 15.$$

Therefore the energy loss due to bremsstrahlung, – dE/dx is proportional to E, which means that the energy loss due to bremsstrahlung is dominant at the highest energies. The energy spectrum of photons being emitted follows a nearly uniform distribution[32]. The electromagnetic cascade keeps growing until the energy loss due to ionisation is larger than that of bremsstrahlung. The critical energy at which the electron loses energy at the same rate from the two processes is ε_c and is ~ 86 MeV in the air.

The Heitler model[33] is a simple model that describes the basic process of the electromagnetic shower development. Shown in Figure 1.3-2 (left), a photon with energy E enters the atmosphere and interacts the first time at depth ~ 37 g cm^{-2} which has been marked as n =1. It splits its energy to two equal

halves and each travels the same fixed distance d. Then at the next level which is marked as n = 2, these two particles both split and resulting in 4 particles each carrying energy E/4. So after the n^{th} splitting, the total number of particles is 2^n and the energy carried per particle is $E/2^n$. At X_{max}, the energy per particle is at the critical energy ε_c as introduced previously.

According to equation 15, the depth of the atmosphere which the particle has to travel before interacting is $d = \lambda_r \ln 2$, where λ_r is the radiation length and the approximated mean free path for pair-productions. Therefore the number of particles at X_{max} is

$$N=2^n=2^{(x/\lambda_r)} \quad 16.$$

The depth of maximum therefore can be expressed as

$$X_{max}=\lambda_r \ln\left(\frac{E}{\varepsilon_c}\right) \quad 17.$$

This model predicts that the maximum number of particles is proportional to the initial energy and the shower of maximum is proportional to the logarithm of the primary energy, as

$$E = N_{max}\, \varepsilon_c \quad 18.$$

The elongation rate which was firstly proposed by Linsley[34] is defined as the rate of change of X_{max} with respect to change of E,

$$\Lambda = dX_{max} / d\log_{10}E \quad 19,$$

which is ~ 85 g cm^{-2} for electromagnetic air showers.

The electromagnetic component of the shower also suffers multiple scattering and can be described by Molière radius. The angle of the scattering after an electron with the critical energy travelling one radiation length is ~ 14˚. The spread of the electromagnetic particles after one radiation length is[35]

$$R_{\text{Molière}} = \lambda_r \times 21\ \text{MeV} / \varepsilon_c \quad 20.$$

This quantity is important because it characterises the lateral distribution of the electromagnetic particles in the air shower. It increases with the altitude and at depth of 875 g/cm^2, the Molière radius is ~ 70 m.

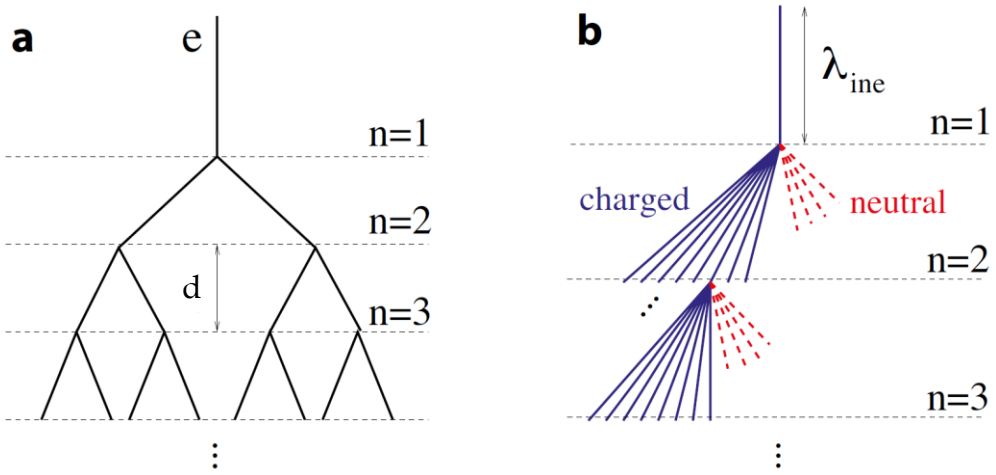


Figure 1.3-2 Heitler model for EM showers (a) and the model modified by Matthews (b) for hadronic showers[6]

The hadronic component of the air-shower has been described by the modified Heitler model by Matthews[36]. Shown in Figure 1.3-2 (right), when a hadron with energy E enters the atmosphere, it travels ~ one interaction length λ_{int} and collides with a nucleus in the air,

$$p + p \rightarrow p + p + N(\pi^+ + \pi^0 + \pi^- \ldots) \qquad 21,$$

Through the interaction at the level where n = 1, if one assumes multiplicity N = 3 (in the graph N = 12) the primary hadronic energy is split into three equal parts to two charged pions and a neutral pion. The neutral pion has a very short lifetime ~ 8.4×10^{-17} s and decays to two photons. Photons will then create electromagnetic showers like introduced before (Figure 1.3-2 left). The charged pions are assumed to travel a constant distance which is related to the interaction length λ_r and then conduct hadronic interactions as indicated at level n = 2. If the number of the charged particles produced at each level is N_{ch}, the average energy per pion $E\pi$ is

$$E_\pi = \frac{E}{\left(\frac{3}{2}N_{ch}\right)^n} \qquad 22,$$

The size of the shower keeps growing since there are more and more hadronic interactions. However due to the lifetime of charged pions is τ ~ 2.6×10^{-8} s, low energy pions have larger probability of decaying than interacting. The energy of pions at this level is called the critical energy ζ_c^π, and if one assumes Γ is the Lorenz factor of the pion, one gets

$$\lambda_{int} / \rho = \Gamma c \tau \qquad 23,$$

ρ is the density of the atmosphere and ζ_c^π is ~ 30 GeV[36]. The interaction length can be calculated from the cross-section and here can be treated

roughly as ~ 80 g cm^{-2}. Pions decay to muons and also create neutrinos in this process

$$\pi^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu} / \overline{\nu_{\mu}} \qquad 24,$$

Hence the number of muons at the critical level is

$$\ln N_{\mu} = \ln N_{\pi} = n \ln N_{ch} = \beta \ln \left[\frac{E}{\zeta_c^{\pi}}\right] \qquad 25,$$

where $\beta = \frac{\ln[N_{ch}]}{\ln\left[\frac{3}{2}N_{ch}\right]}$ and is shown to be ~ 0.85 from simulations[36]. Therefore the number of muons in the hadronic shower is proportional to the logarithmic primary energy. If one uses a toy model to where each iron nucleus is composed of 56 nucleons that each with energy E/56, the number of muons is ~ 1.5 times that from a proton primary shower with the same initial energy E.

It should be noted that the elasticity has not taken into account in the above calculations. The elasticity κ_{ela} which is the fraction of the energy carried away by a single 'leading particle' is poorly known from particle physics experiment because the energy of the first interaction of the ultra-high energy cosmic rays is significantly higher than the highest energies reached by particles accelerated in accelerators. The change of elasticity and cross-section has great impact on X_{max}. The number of muons on the ground is sensitive to the multiplicity of hadronic interactions[37]. At the moment, extrapolations that based on the measurements from the LHC are used in the hadronic simulations. However there are still disagreements between mostly commonly used models. This topic is explored in section 8.1.

Using Monte Carlo simulations, it is possible[6] to see the longitudinal development of different components in the air (Figure 1.3-3, right) and the lateral spread of each component on the ground (Figure 1.3-3, left). The hadron component is composed of long-lived secondary particles (baryons, charged mesons) that travelling along shower axis and contributes to signals at the core. This component is often ignored in the later chapters because of the low density. The shower is initiated by proton-primary and at energy of 10^{19} eV. It was simulated using CORSIKA.

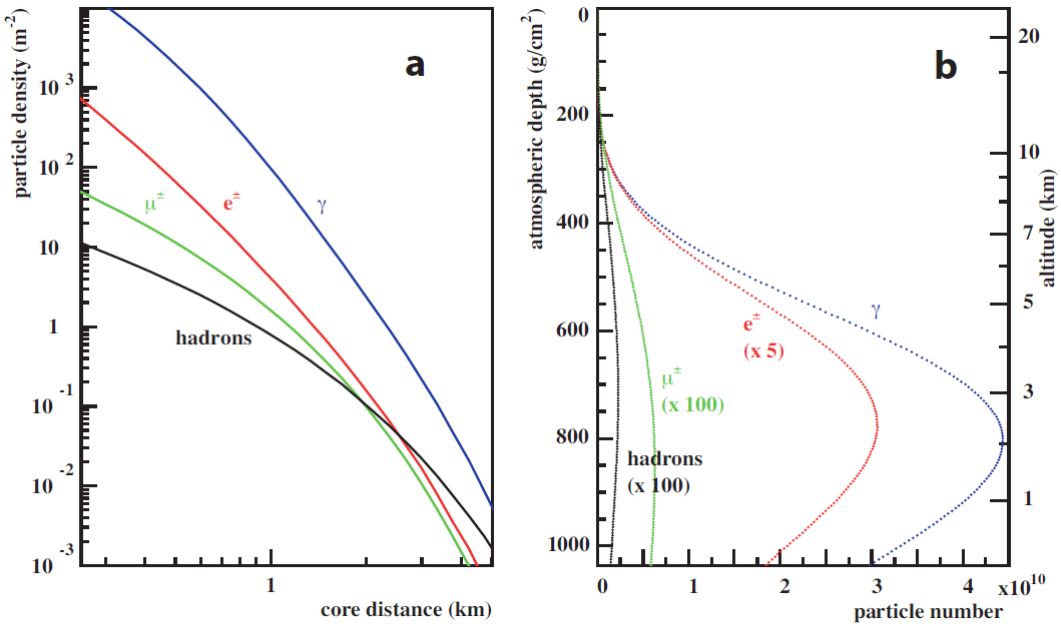


Figure 1.3-3 Lateral (at ground level) / longitudinal distribution of particle density[6] of four components for a proton of 10^{19} eV

The ground is set at the altitude of the Pierre Auger Observatory, as 875 g cm^2 for the plot on the left. Number of hadrons, muons and electrons has been scaled respectively in the longitudinal plot. It can be concluded that at the ground, the slope of the change of signals with respect to core distance is larger for photons, electrons than muons. It also can be seen that the measurement of X_{max} mainly describes the maximum number particles from the electromagnetic component.

X_{max} is closely related to the first interaction point in the atmosphere, which is determined by the cross-section. It is also sensitive to the fraction of energy that goes to the electromagnetic channel. For showers with the same energy of primary proton and iron, according to the toy model the energy per nucleus for iron is smaller than proton, such that on average the X_{max} of proton is larger than that of iron.

A recent result from the Pierre Auger Observatory about the mass composition based on measurements of the mean and the fluctuations of X_{max} is shown in Figure 1.3-4[38]. Data from Jan. 2004 to Dec. 2012 were used. The precision of X_{max} on average is ~ 20 g cm^{-2}. Data are compared to updated simulations based on EPOS-LHC[39] and the QGSJetIII-04. The mean and fluctuation of X_{max} indicate for $E < 3 \times 10^{18}$ eV that the mass is light. The reason that the fluctuation is higher than predicted for protons is perhaps because of a mixed composition. It is evident there is a change of the mass and that it appears to be heavier at higher energies.

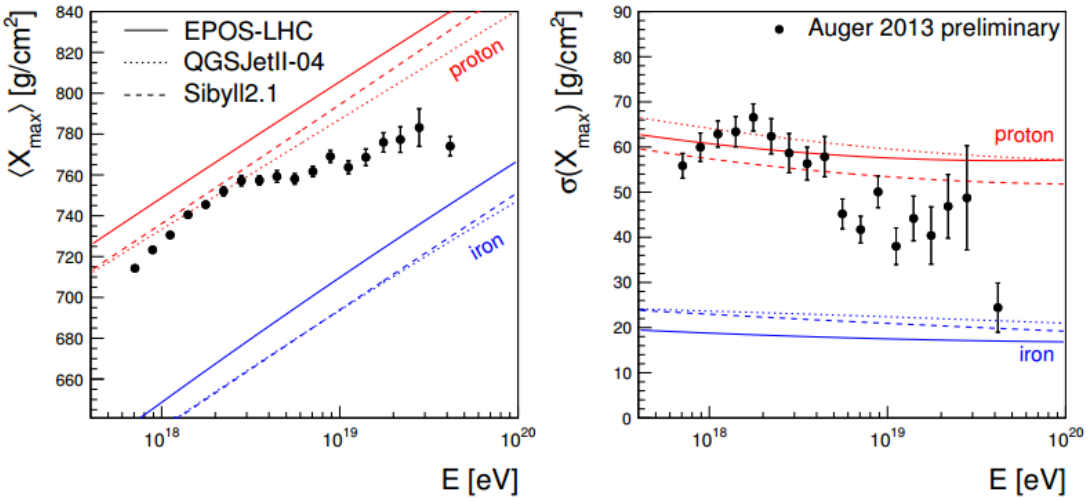


Figure 1.3-4 Updated[38] mass composition measurements from Auger using mean and fluctuations of X_{max}

The most recent result[40] from the mass composition study in Auger is interpreted using different hadronic models. It suggests a strong dependence of the proton fraction with respect to energy but does not favour a significant contribution of iron nuclei. Disagreements in the interpretation using hadronic interaction models are also found.

The aim of the thesis is to look for ultra-high energy photons although the data are dominated by hadronic primaries. In the next section, the importance of studying UHE photons is introduced and the energy loss during propagation is shown. In the end the photon fraction measured from previous experiments are compared to predictions from various astrophysical scenarios

1.4 Ultra-high energy photons

The mass composition of the ultra–high energy cosmic ray is a key problem to solve as part of the progress to identify the origin of these particles. The current publications from the Pierre Auger Collaboration using X_{max} suggest a mixed mass composition (Figure 1.3-4). As the energy of the primary particle is beyond the highest energy that the particle accelerator could reach, extrapolations of the hadronic interaction models are needed for the predictions which are used for the mass composition analysis. Some of the recent studies[41] by the Pierre Auger Collaboration suggest deficit of the muon number detected compared to results from simulations, which has raised the question whether one should trust the hadronic simulations

blindly. It is therefore important to study the mass composition using the methods that depend less on the hadronic interaction models.

Photons of ~ 10^{19} eV could be an indicator of the mass composition without relying on hadronic interaction models. In this section, scenarios and mechanisms for the production of UHE photons are introduced, the energy loss processes during the propagation of photons are described and a brief summary of the photon fraction expected from various astrophysical scenarios is included.

1.4.1 The production of ultra-high energy photons

As introduced in section 1.2, neutral pions can be produced through interactions between ultra–high energy protons and the cosmic microwave background. Photons are generated from decays of neutral pion

$$\pi^0 \rightarrow \gamma + \gamma \qquad 26,$$

These photons are known as cosmogenic photons and carry ~ 20% of the primary proton energy. If the primary cosmic rays are heavier, the threshold for the pion-production increases (e.g. ~ 5×10^{21} eV for Fe). The dominating interaction for heavy nuclei of energy 5×10^{19} eV is photo-disintegration which does not produce photons (equation 8, 9 10). The photon fraction is therefore sensitive to the primary mass of the cosmic ray.

Alternatively, there are top-down models that predict the existence of UHE photons. For example the Z-burst model[42] proposes UHE neutrinos from sources annihilate with the relic neutrino background to produce Z-resonance (Figure 1.4-1).

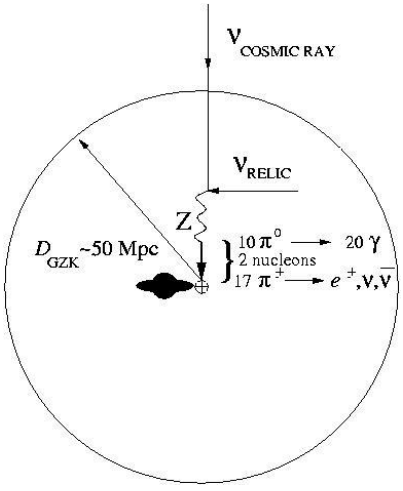


Figure 1.4-1 Z-burst model and the production of UHE photons[42]

Here, the energy threshold of the UHE neutrino in the rest frame of the neutrino background is

$$E = m_Z^2 / (2 \times m_\nu) = 4 \times (eV / m_\nu) \times 10^{21} \text{ eV} \quad 27.$$

Then Z-boson decays to showers of nucleons and pions. Charged pions generate neutrinos and neutral pions decay to photons. This model predicts a photon fraction higher than what is expected from the GZK-interaction of protons and about 0.015% at 10^{19} eV. The radius shown in the plot indicates the attenuation length of the GZK-interaction for protons.

The other two top-down models include the decay of Super Heavy Dark Matter (SHDM), which are metastable particles and produced in the early Universe; and Topology Defects (TD) which propose the decay of a super-massive particle X with energy 10^{22}-10^{25} eV. The decay products[43] are leptons and quarks. Quarks then hadronize into hadronic jets (3%) and pions (97%) that produce UHE photons.

1.4.2 The propagation of UHE photons

Our universe is not transparent but is filled with cosmic background radiations of photons ranging from the radio to infrared wavelength. UHE photons can interact with them and create e^+/e^- pairs

$$\gamma + \gamma_{radio} \rightarrow e^+ + e^- \quad 28,$$

The distribution of the energy density[44] of photon background (cosmic radio background CRB, cosmic microwave background CMB and cosmic infrared background CIB) is shown on the left of Figure 1.4-2. The cross-section[44] for pair-production of photon with GeV energy in the laboratory frame is shown in the right of Figure 1.4-2.

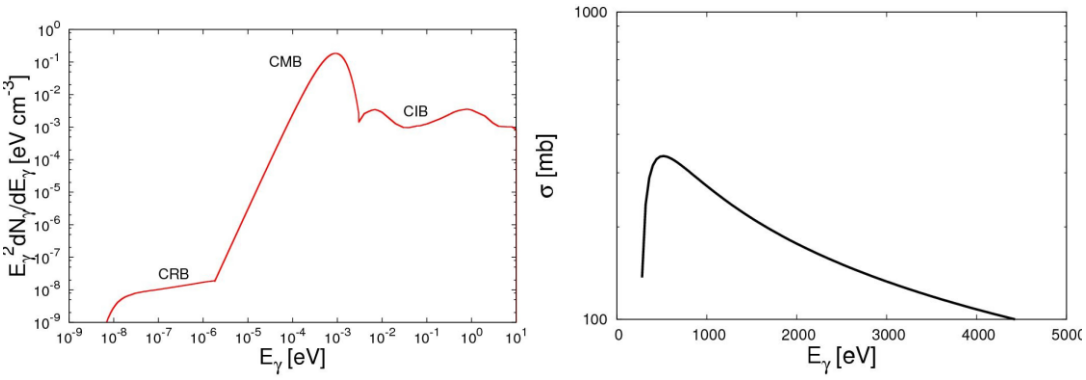


Figure 1.4-2 Energy density for cosmic background radiations (left) and the cross-section for pair-production when the photon is colliding with a photon with GeV energy in the laboratory frame (right)[44]

For photons with energies ~ 10^{19} eV, they are more likely to interact with radio background and generate e^+/e^-. The leading electron/positron typically carries 90% of the primary energy of the photon if the background photon is of energy ~ 10^{-6} eV. Electrons and positrons then could generate

UHE photons through inverse Compton scattering with the background photons (~ 10^{-8} eV cm^{-3}) or lose energy through synchrotron radiations and are deflected in the extragalactic magnetic field (~ nG, ~ 0.25 eV cm^{-3}).

If the cosmic microwave background is the only radiation field, the dominant energy loss for e^+/e- is synchrotron radiation. However, if the radio background has a larger energy density, the cross-section of the inverse Compton scattering follows the classical Thomson approximation and results in the inverse Compton scattering as the dominant source of the energy loss. No matter in which scenario, as a consequence, UHE photons lose energy gradually and are restricted to come from the near Universe. The attenuation length (the distance that the primary photon loses ~63%, (1-e^{-1}), of its original energy after travelling through) is plotted against the primary energy of the photon[45] and shown in Figure 1.4-3.

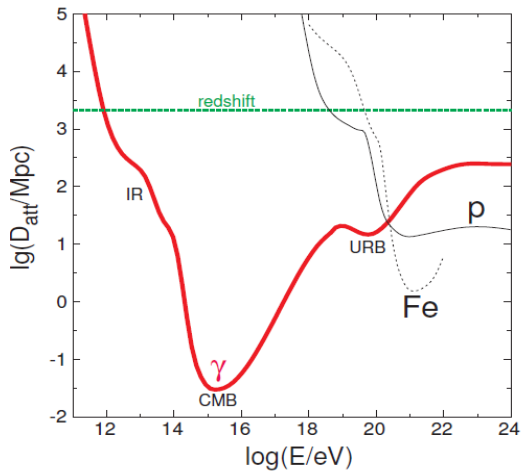


Figure 1.4-3 Attenuation length of pair-production for photons with different energies[46]. URB is the cosmic radio background, which interacts with UHE photons.

The dip in the centre is due to the relatively high density of the cosmic microwave background. It is interesting that as the bi-products of the these interactions, it might be possible to use gamma-ray astronomy telescopes (e.g. the Cherenkov Telescope Array, CTA) for the multi-messenger studies of UHE photons since flux of photons at < 10^{12} eV are not suppressed compared to the flux of the highest energy photons.

1.4.3 Predictions of photon fraction from various astrophysical scenarios

No UHE photons have been claimed to be discovered so far. Several experiments have results giving upper limits to the photon fractions at energy > 10^{19} eV such as the AGASA[47], Yakutsk[48] and the Haverah Park[49] experiments.

Data taking in Yakutsk started in 1973. It was designed to observe cosmic rays with the highest energies. It used Cherenkov detectors but only had one muon detector before 1978. Then it was replaced by six detectors with the largest one of area 36 m^2. AGASA was a ground-based experiment with 111 scintillation detectors covering an area of 100 km^2 and 27 muon detectors. It was operating from 1993 to 2003. By pre-selecting events with light muon signals, muon and scintillation signals of 10 events observed by Yakutsk and AGASA were analysed and concluded[50] that at 95% of the confident level, the upper limit of photon fraction for cosmic rays with energy > 10^{20} eV is 36%.

The Haverah Park experiment was located close to the University of Leeds and was an array of water-Cherenkov detectors. The total area is 12 km^2 and was recording data during 1974 to 1987. Data with inclined angle (60˚ to 80˚) were analysed and there were 46 events with energy > 10^{19} eV. The muon density maps were generated[49] because at such high zenith angles the electromagnetic components only contribute a small fraction to the signal measured on the ground. The muon map obtained from Monte Carlo simulations was compared to data as the photon-hadron separation parameter. It was concluded that at 95% confidence level, the photon fraction has the upper limit of 55% for energy larger than 4 × 10^{19} eV.

A summary of results prior to Auger works, of upper limits of photon fractions[51] is shown in Figure 1.4-4. HP stands for Haverah Park, Y stands for Yakutsk, A1 and A2 are from AGASA, AY is the combined analysis between Yakutsk and AGASA, Auger HYB is the Auger hybrid and Auger SD is the data from surface detector of Auger. Details of results from Auger are introduced in Chapter 3.

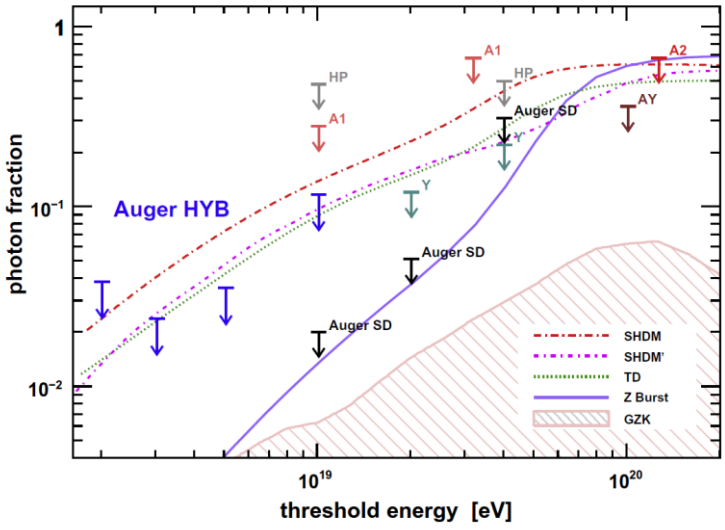


Figure 1.4-4 Upper limits to photon fraction at integrated energy bins[51]

Constraints on some of the top-down models had been set by experiments before the Auger Observatory. However, as Auger started taking data, most of top-down models have been eliminated except for the Z-burst model at the time this plot was made in 2009. It should be noted that the results from Haverah Park, AGASA and Yakutsk are model-dependent.

Except for constraining top-down models, the measurement of the photon fraction and flux also offers complementary information to the UHECR spectrum for restraining parameters of astrophysical scenarios assuming UHECR are accelerated from sources. For example if the UHECR sources are isotropically distributed, various combinations of values for the spectra index α of the source injection, extragalactic magnetic field strength B, the cut-off energy of the source E_{max} and the mass composition can be fitted to satisfy the measured UHECR spectrum and mass composition results from Auger. Figure 1.4-5 shows two possible scenarios[52]. The grey band shows the photon fraction if all UHECR are composed of protons and the red band shows if the composition is iron. The widths of bands are due to uncertainties in the radio background. Black points are measured values from SD results that were published by the Pierre Auger Collaboration, 2008. Radius of the curvature and risetime of the FADC trace from SD stations are used in this analysis (details see Chapter 2). The blue line shows the expectation of photon fraction after 20 years of operation of Auger if the analysis does not change and no photon candidates are found.

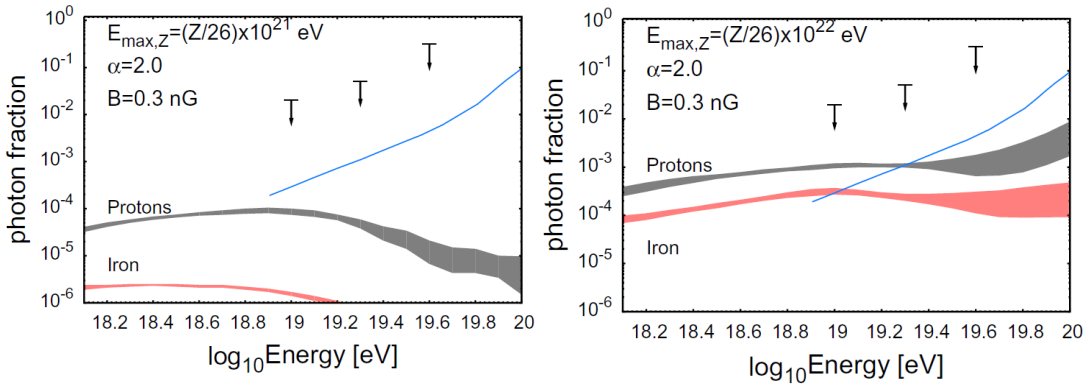


Figure 1.4-5 Predictions of photon fraction for different astrophysical scenarios[52]

It could be inferred that a proton-dominated mass composition is likely to allow the detection of UHE photon from Auger if the cut-off energy of the source is not limited to the threshold energy of the GZK effect. The photon fraction for the heavy mass-composition is ~ suppressed by one order of magnitude compared with that of the proton–dominated UHECR.

This chapter described the journey of a UHECR and the UHE photons from the source, propagating through space and the Earth atmosphere, to the detector. The importance of studying UHECR has been demonstrated. The next chapter contains an introduction of the Pierre Auger Observatory, which is the largest experiment for detecting UHECR in the world.

Chapter 2. The Pierre Auger Observatory and air shower simulations

The Pierre Auger Observatory was designed to study the energy spectrum, mass composition and the distribution of arrival directions of ultra-high energy cosmic rays. It is the largest experiment in the world for cosmic ray measurements and covers an area of 3000 km^2 in the province of Mendoza, Argentina (Figure 2.1-1). The construction was completed in June 2008 and data have been recorded since 1st Jan. 2004. It contains multiple types of detectors: fluorescence detectors (FD) that detect longitudinal information of the air shower and enable the calculation of calorimetric energy of the shower. An array of surface detectors (SD) filled with water that detect the Cherenkov radiation from charged particles with high energies on the ground with 1500 m spacing, muon counters made of plastic scintillators and buried 3 m underground, which are surrounded by an array of SD with spacing 750 m and 435 m (Auger Muons and Infill for the Ground Array[53], AMIGA, total area 30 km^2), radio antennas that measures the radio pulses (30 – 80 MHz) from the air shower (Auger Engineering Radio Array[54], AERA) and various instruments as prototypes to detect microwaves with few GHz[55]. The radio detection is still in the development stage and is not discussed in this work. The muon detectors could be useful for photon-hadron separations but due to the small area of the array, they are only sensitive to showers with energy ~ 10^{17} eV and are also not discussed. The fluorescence telescopes can only be operated during dark, moonless and clear nights which reduce the duty cycle to ~ 13%. The analysis introduced in this thesis mainly uses data from the surface detectors, which have almost 100% duty cycle.

In this chapter, information to show how the surface detectors are calibrated and triggered is introduced. The event trigger and reconstruction based on data from the SD is explained. A short introduction to the fluorescence detector is included. The philosophy behind energy calibration of SD events based on the calorimetric energy measured by FD is shown. Finally the techniques and models needed for simulating air showers are summarized.

2.1 The surface detectors: descriptions, calibrations and triggers

There are 1600 surface detectors arranged in a hexagonal grid on the ground with 1500 m spacing. The layout of the detectors has been proved to be fully sensitive for the detection of showers from hadronic primaries with energy above $10^{18.5}$ eV and zenith angle 0 – 60°. Figure 2.1-1 shows a map[56] of the SD array where each of the points represent a SD station.

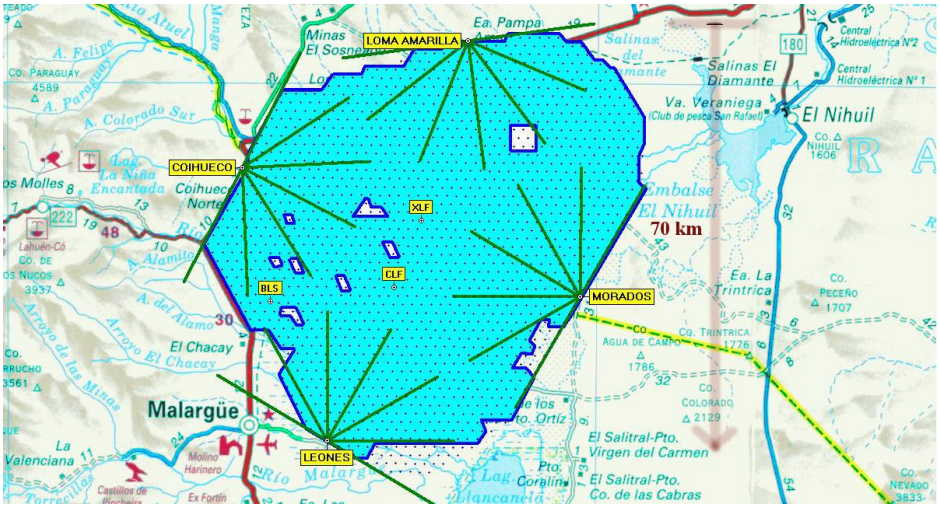


Figure 2.1-1 Map of the Pierre Auger Observatory and the control centre is in Malargüe. Los Leones, Los Morados, Loma Amarilla and Coihueco are four fluorescence stations (FD). CLF , XLF and BLS are laser facilities that are used for calibrations.[56]

In particular, there are a few pairs of 'twin' stations built in the left top corner of the array and close to the Coihueco station. The distance between each station of the twin is 11 m and they are used to measure the uncertainty of the signals (see section 5.2). The surface detectors were designed with inspiration from the Haverah Park array. Each detector has an area of 10 m^2 and is filled with 12 m^3 purified water. When charged particles go through the water (refractive index n) with relativistic speed βc that is larger than the speed of light in the water c / n, the polarization before and at the rear of the charged particle is asymmetric. This leads to a varying dipole moment and emission of photons, which form a coherent wave front. The angle between photons and the travel direction of the particle is θ (Figure 2.1-2, left). From the simple geometry,

$$\cos\theta = 1 / (\beta n) \qquad 29.$$

The maximum opening angle θ is reached when $\beta \sim 1$ and for particles in the water the Cherenkov cone is with $\theta \sim 41°$. The energy needed for such particle to create Cherenkov photons is

$$E=\Gamma mc^2=\frac{mc^2}{\sqrt{1-\beta^2}}=\frac{mc^2}{\sqrt{1-\left(\frac{1}{n\cos\theta}\right)^2}}\geq\frac{mc^2}{\sqrt{1-\left(\frac{1}{n}\right)^2}} \qquad 30.$$

The threshold energy is ~ 0.77 MeV for electrons and ~ 158 MeV for muons.

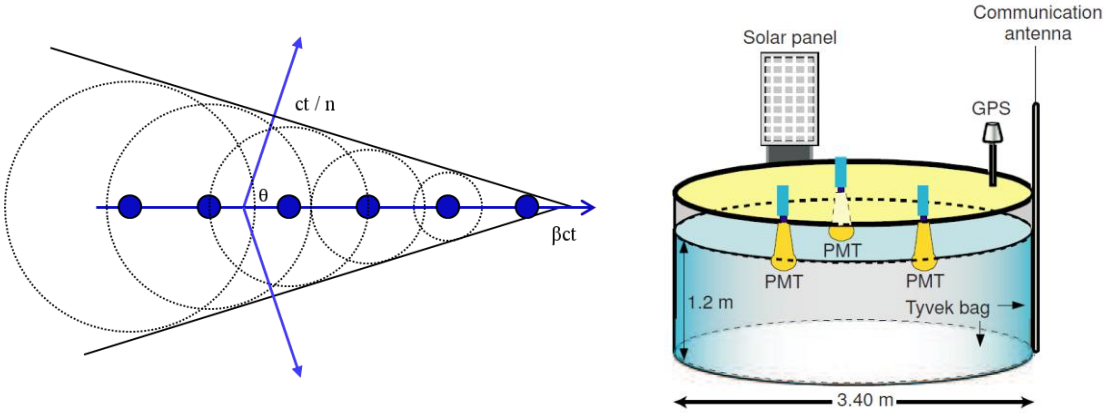


Figure 2.1-2 A cartoon of Cherenkov emission of photons and details of a SD station[57] (For illustration only, not to scale)

The water in each detector is 1.2 m deep and contained in the plastic liner with diameter 3.4 m. The liner is covered with Tyvek® to increase the diffuse reflections of the Cherenkov light. Each station is powered by a battery, which is charged with a solar panel that gives a power of 10 W. Three 9-inch photomultipliers (PMT) are installed at the top of the liner to detect the Cherenkov photons produced in the water. The end-to-end gain of the PMT is ~ 2 ×10^5 and two outputs are obtained from each PMT, namely from the last dynode and the anode. The last dynode signal is amplified by a factor of 32 (5 bits) and is called the high-gain channel (HG). The gain ratio of dynode over anode is continuously monitored using non-saturated pulses. The time delay is in the order of 5 ns between two gains due to the amplification process. The HG channel is used for the determination of signal sizes unless it is saturated, in which case the anode signal (the low-gain channel, LG) is then used. Flash analogue-digital converters (FADC) are used to digitise readings from PMTs in 10 bits (1024 channels) and with sampling rate of 40 MHz (25 ns). By combining both outputs, one is able to measure in a range of a few to 10^5 photoelectrons. Signals are then processed by a fast programmable logic device (PLD), which has been programmed with the triggering algorithms. When a station is triggered, 256 bins (25 ns per bin, 6.4 μs) of FADC traces before and 512 bins (12.8 μs) after the trigger are saved. Each station is connected to the Global Positioning System (GPS) which has an internal 100 MHz-clock for time synchronisations and uncertainty of 7 ns. To transfer recorded information, stations are connected with the Central Data Acquisition System (CDAS) via signals transmitted via radio towers that located in the array. The bandwidth is ~ 1200 bits per second. The communication system was designed and implemented by a

team from the University of Leeds. The monitoring data for each station such as the temperature of the water, the baseline of the PMT, the voltage of the remaining battery is sent to CDAS every 10 minutes

Atmospheric muons hit each detector with a frequency of ~ 2500 Hz and mean energy of ~ GeV, which allows the signal calibration to be based on nature. The energy deposited in the water from muons is proportional to the track length. When a muon travels through the detector vertically from the top, a pulse is formed in each PMT with the typical decay time ~ 60 ns for the tail of the response. The energy deposit of such muon is ~ 240 MeV, which is calibrated as one Vertical Equivalent Muon (VEM). A reference detector was selected and used to measure the Single Muon Response (SMR) by a muon telescope (two scintillators, one on the top and one below the detector). For three PMTs (3-fold), three charge distributions can be obtained. The integrated charge of the three distributions is defined as 3 VEM (Figure 2.1-3, left). The mean of the charge distribution is known as the VEM peak, which is also called the amplitude of the ADC channel (Figure 2.1-3, right).

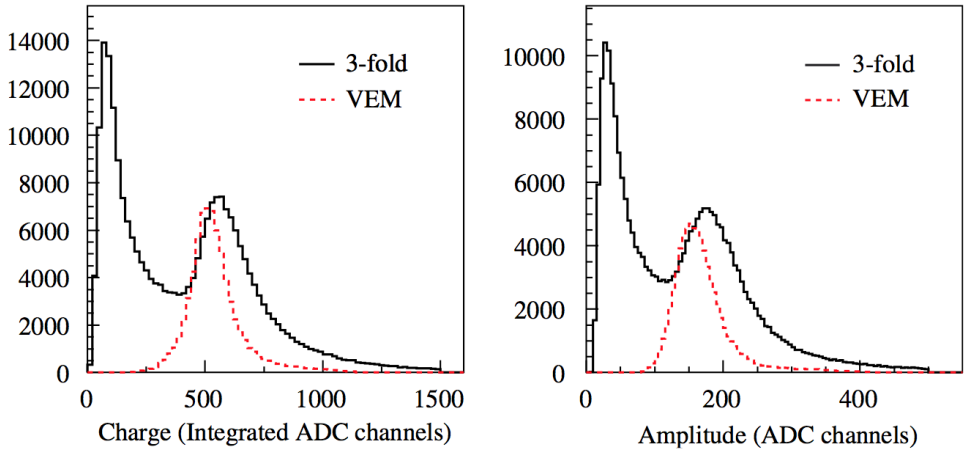


Figure 2.1-3 The integrated charge (left) of three PMTs, the mean (right) of the charge distribution (each PMT) from atmospheric muons (black) and only-vertical muons (red)[58]

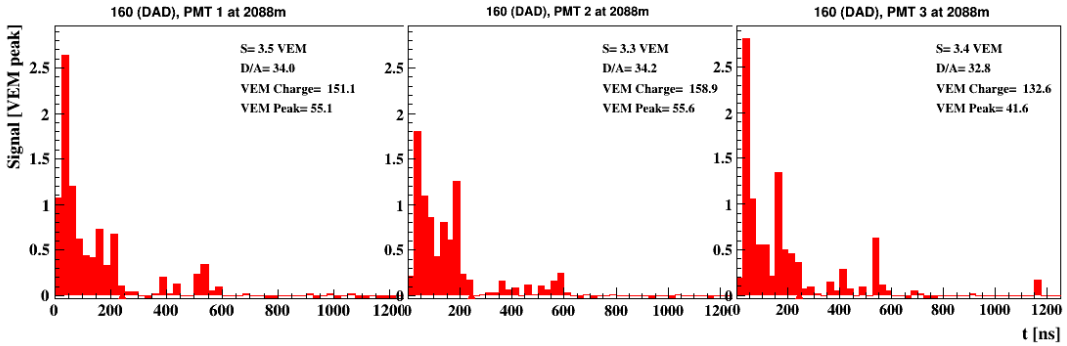
Atmospheric muons come from all angles and with a spectrum of energies. Distributions of the integrated charge and the peak (amplitude) are shown in Figure 2.1-3. The large tail of the distribution is because muons arrive at inclined angles having longer path lengths in the detector. The cause of the first peak from the black curve has been studied in details[59], which suggests it is due to the triggering effect of low energy particles. The peak of the summed charge distribution of three PMTs is ~ 1.09 VEM and the mean of the peak of charge distribution of each PMT is 1.03 ± 0.02 VEM. The difference of the two values is because the sum of three PMTs describes the

total signal in the station while each individual PMT only measures a fraction of the signal that are geometrically close to it.

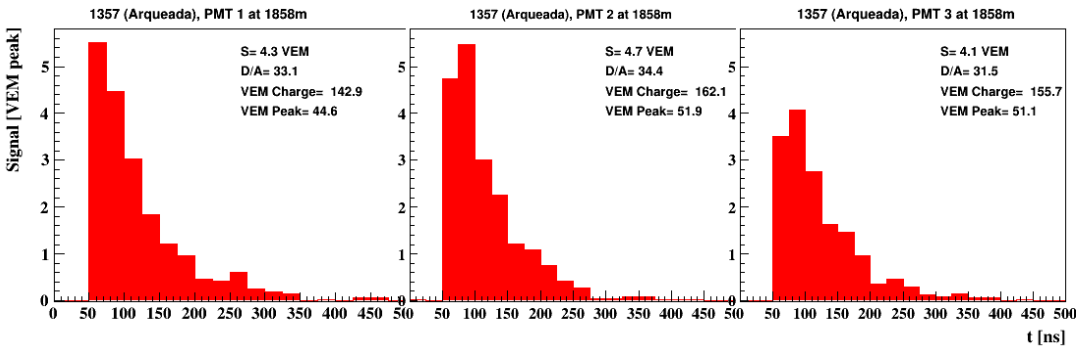
It is faster to calculate the peak of the distribution than do the integration; therefore triggers carried by PLD are designed to use the peak values. There are two levels for the local station trigger, namely T1 and T2. To reach the trigger level, three types of triggers are defined:

- Th1r trigger: Signal threshold trigger. At least one time bin with signal larger than 1.75 VEM peak (Figure 2.1-4 a).
- Th2r trigger: Also signal threshold trigger but requests at least one time bin with signal larger than 3.2 VEM peak (Figure 2.1-4 b).
- TOT trigger: Time Over Threshold trigger. It requests at least 13 continuous time bins (25 ns per bin, 325 ns in total) with signal larger than 0.2 VEM peak (Figure 2.1-4 c).

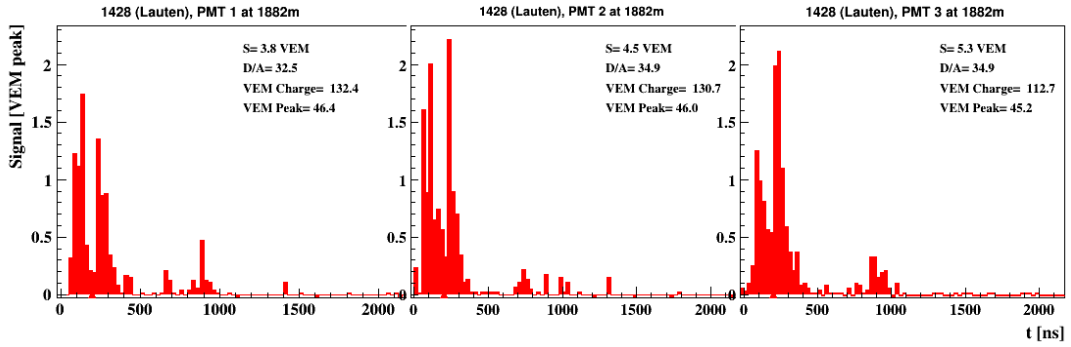
The threshold triggers are designed to select fast traces and are sensitive to muon signals. The TOT trigger is more useful for long and slowly developed traces, which is more sensitive to the electromagnetic component of the showers.



a. Th1r trigger



b. Th2r trigger (also survives Th1r trigger)



c. TOT trigger

Figure 2.1-4 Example of FADC trace that has Th1r / Th2r / TOT trigger

It should be noted that the VEM charge are used to describe the signal of the PMT. This is because the fluctuations of the integrated charge are smaller than the mean of the charge distribution (Figure 2.1-3). The peak value is ~ 3 times the value of the VEM charge. Through this thesis, the default unit of the signal of the FADC traces is VEM charge unless specified.

When it is determined that a station was triggered at T2 level, the time stamp and the trigger type is sent to CDAS. The rate of the T2 trigger and TOT trigger is used for monitoring the status of the SD station. Patterns of the spatial and timing of stations are continuously searched using CDAS to find out if there is such combination of stations to satisfy the next trigger level – T3 trigger. T3 trigger is designed to seek for detectors with a spatial and temporal coincidence. The philosophy of this trigger is to make use of the coincidence rate within a time window of 50 μs that ≥ 3 neighbouring stations were flagged as triggered. Various patterns are allowed for the trigger[60], such as a three-fold event, which has a station of the highest signal with T2, and two other stations located at 1500 m. The information about PMT traces are delayed to be transferred for 10 s and only will transfer if a T3 trigger is found.

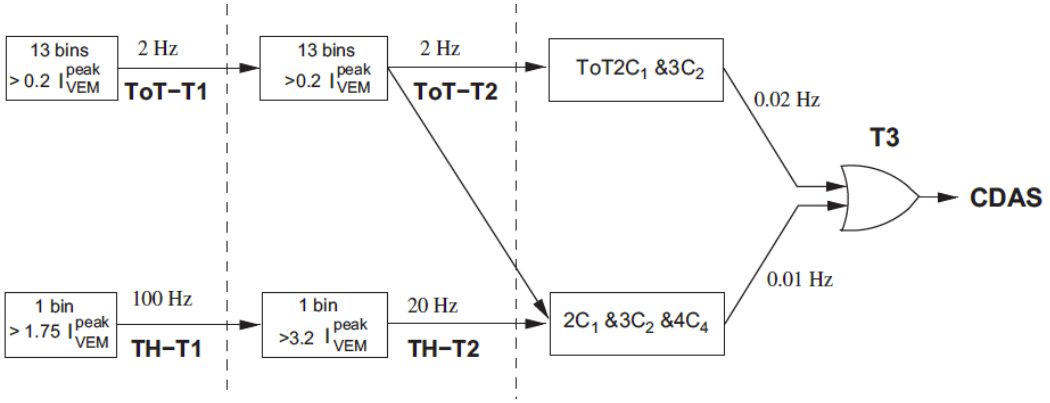


Figure 2.1-5 Flowchart diagram of T1, T2 and T3 triggers for SD stations[61]

Events selected with T3 trigger are still mixed with a fraction of background events due to coincidences of accidental muons from the atmosphere. The physics trigger – T4 trigger is therefore implemented to provide further selection of cosmic-ray events. All possible layouts of detectors that satisfy the T4 selection are shown in Figure 2.1-6. The first category includes a minimum of three stations that are TOT and therefore named 3TOT, which is shown in (a). The second category (b) allows stations with any type of the T2 trigger but in addition the station in the centre has to be surrounded by another three stations on the closest hexogen. This trigger is called 4C1.

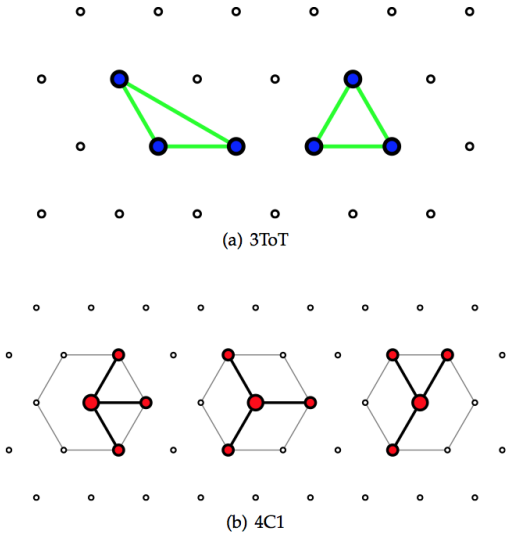


Figure 2.1-6 Patterns of stations that satisfy requirements of T4 trigger[62]

In the end, to ensure the quality of the event, the station with the highest signal (often called the 'hottest' station) has to be surrounded by 6 stations that are working normally. This is the T5 trigger and with the rate of 3×10^{-5} Hz. The flowchart that describes T3, T4 and T5 trigger is shown in Figure 2.1-7.

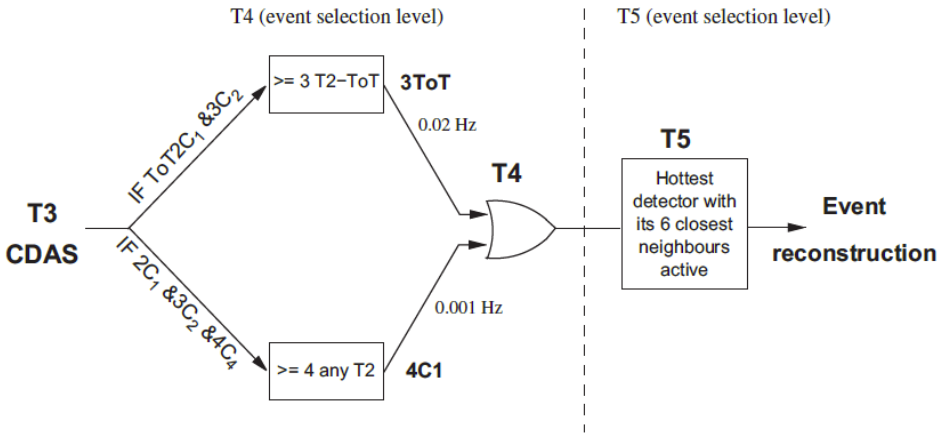


Figure 2.1-7 Flowchart diagram of T3, T4 and T5 trigger for event selections[61]

2.2 Event reconstruction using surface detectors

The aim of the SD event reconstruction is to obtain the geometry of the shower including the zenith angle, the azimuth angle, the core position and the size of the shower. When the shower front reaches the ground, firstly it is fitted using a plane model (Figure 2.2-1, left). A more realistic model of the shower front is a sphere, which is related to the determination of the radius of curvature (Figure 2.2-1, right). The barycentre of all stations selected in the event is set as the initial value of the core position of the plane fit.

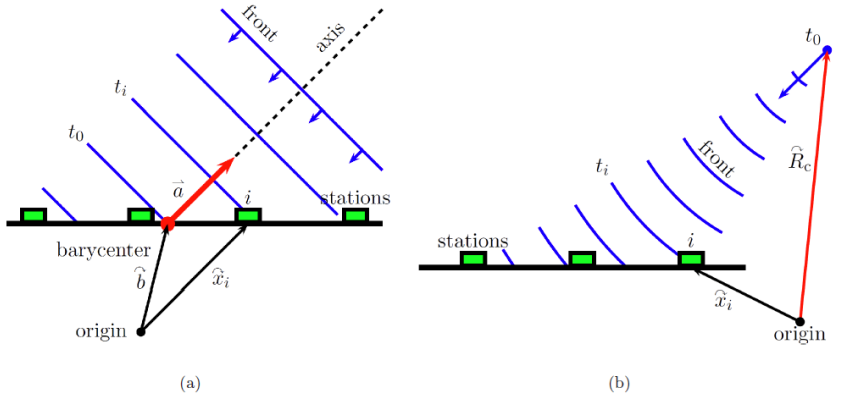


Figure 2.2-1 Shower front fitted with a plane/sphere for calculating arrival directions[63]

The delay of the trigger time of each station with respect to the plane is fitted with respect to the distance from core using a second-order polynomial. An example of the plane fit from an event with zenith angle ~ 15° and energy ~ 8×10^{18} eV is shown in Figure 2.2-2. There are 6 stations triggered in this event. The resolution of the arrival direction is typically of the order of ~ 1°.

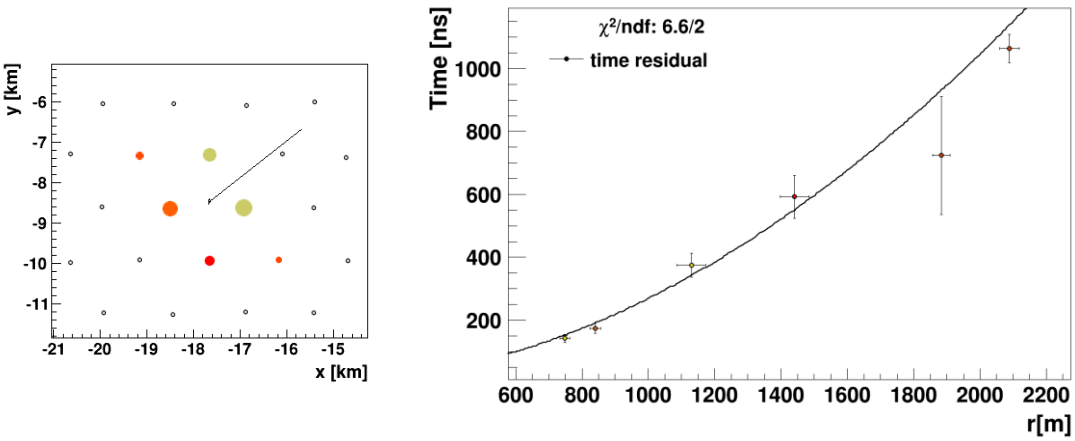


Figure 2.2-2 The geometry and plane fit (start time delay with respect to the plane vs core distance) of a near-vertical event. The size of the detector drawn indicates the size of the signal and the colour of the stations are related to the trigger time of each station. The first triggered detectors are marked in yellow, and then orange and the last is red.

The other important ingredient for the event reconstruction is to fit the Lateral Distribution Function (LDF), which describes the change of signal size in each station as a function of the distance from the shower axis. The LDF fit is found empirically[64] and the following function has been used

$$S(r)=S(1000)\left(\frac{r}{1000}\right)^{\beta}\left(\frac{r+700}{1700}\right)^{\beta+\gamma} \qquad 31.$$

Both β and γ are slope parameters that describe how rapid the signal falls along the core distance (the example event has β = - 2.4 and γ = 0). These are not free parameters during the LDF fit but have been parameterised as functions of zenith angle and S(1000), the signal estimated at 1000 m from the core. The LDF of the example event is shown in Figure 2.2-3.

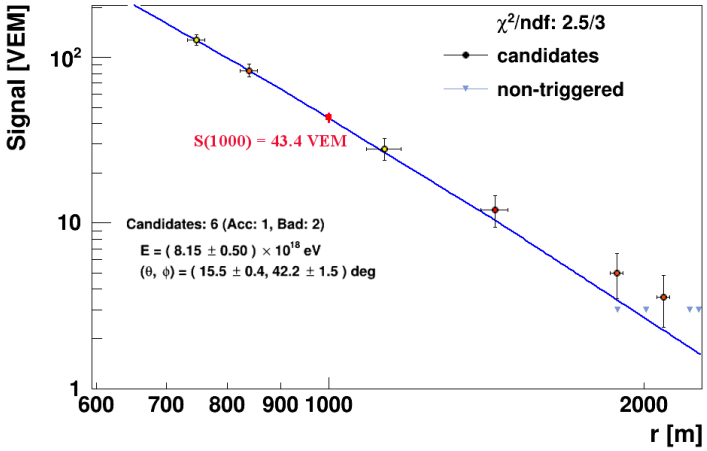


Figure 2.2-3 The lateral distribution function of a nearly-vertical event with six stations. The square at 1000 m shows the fitted value of S(1000) is ~ 43 VEM.

The core position (x and y) and S(1000) are found through the reconstruction. The fitting is achieved by using the maximum likelihood method and accounts for neighbouring stations that have not been triggered. The uncertainties of signal sizes are obtained by measuring the signal variance[65] from 'twin' stations that are 11 m apart (similar technique is used in section 5.2 and described in details).

S(1000) is found[66,67,68] to be a good estimator of the size of the shower and is used to evaluate the energy. At a fixed primary energy, S(1000) depends on the zenith angle of the shower since the shower goes through different depth of the atmosphere. It is possible to convert values of S(1000) empirically at any zenith angle to S(1000) measured at zenith angle θ. The median of the distribution of zenith angle is 38° and S_{38} is used for the energy conversion. The relationship between S(1000) and S_{38} is

$$S_{38} = \frac{S(1000)}{1+ax+bx^2} \qquad 32,$$

where $x = \cos^2 38^\circ - \sin^2 38^\circ$ and values have been parameterised[27] using data. Instead of using Monte Carlo simulations, a calorimetric energy determined from the fluorescence detectors, is used to calibrate the SD energy. Details are introduced in the following section.

2.3 Fluorescence detectors and the energy calibration

The charged particles in the extensive air shower excite nitrogen molecules in the atmosphere. A fraction of the energy deposited in the air is then re-emitted during the de-excitation of the nitrogen molecules. The emission is a band spectrum with many lines. A recent measurement of the emission spectrum of the ultraviolet fluorescence light from nitrogen measured by the AirFly collaboration is shown in Figure 2.3-1.

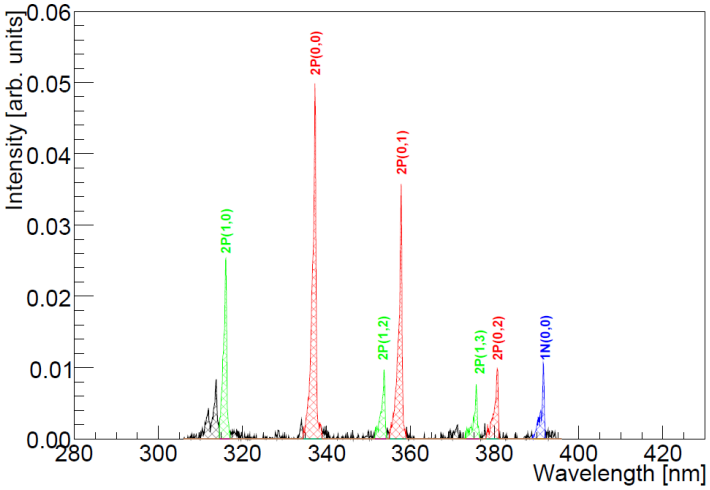


Figure 2.3-1 Emission spectrum of ultraviolet photons from excited nitrogen measured from the AirFly collaboration[69]

The number of fluorescence photons emitted is proportional to the energy that is deposited by charged particles from the air shower in the atmosphere. The fluorescence yield Y^f quantifies the relationship between the intensity of the fluorescence photons and the energy loss of the charged particle. A typical value for the standard environment is that for each MeV of energy deposited, 5 photons are emitted in the wavelength range of 300 – 400 nm in the dry air[70].

The emission of the fluorescence photons is isotropic and only a fraction of them can be observed using, the Fluorescence detectors (FD). The expected signal at a FD with aperture A is

$$N_\gamma = N_e Y^f c \Delta t Q_e (A/4 \pi r^2) \exp(-r/r_0) \quad 33.$$

The exponential term takes account of the transmission of photons to reach the detectors (distance is r), $c\Delta t$ is the track length of ionised electrons from the air shower and Q_e is the quantum efficiency of the optical devices.

Several techniques have been used in the Pierre Auger Observatory for monitoring the atmosphere and aerosols above the telescopes such as using LIDARs to receive laser shots[71] to determine the amount of clouds using the cloud cameras (sensitive to photons with wavelength 7 – 14 μm) and to obtain atmospheric conditions (pressure, temperature and humidity) from the Global Data Assimilation System[72] (GDAS).

There were originally in total four FD stations and 24 telescopes. Each telescope has the field of view of 30 × 30° from 0° to 30° in elevation. In 2009 another three telescopes were built as the low-energy enhancement[73] for Auger. The project is named 'High Elevation Auger Telescopes' (HEAT). Each telescope has the field of view from 30° to 60° in elevation. No data from HEAT are used in this thesis.

All telescopes are equipped with a filter so that only ultraviolet light with wavelength 300 – 410 nm can pass through. The filter is proven to be important for avoiding noise from other light sources and also helps prevent dust going into the camera system. Figure 2.3-2 shows the structure of a FD telescope.

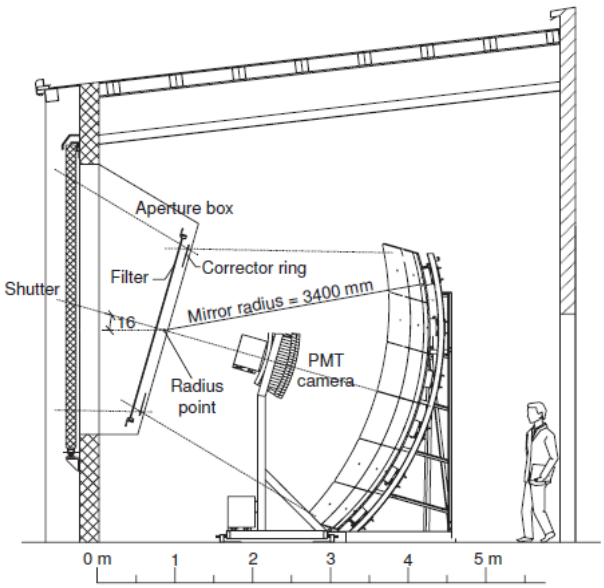


Figure 2.3-2 erre Auger Observatory[69]

It is a camera with 440 (20 × 22) pixels (PMTs) each of which gives a signal that is digitised with a 100 MHz FADC. The start time of each triggered PMT

is used to determine the arrival direction of the air shower by treating the shower development as a series of point source travelling along the shower axis. The plane that contains both the shower axis and the telescope is named the shower detector plane (SPD) and shown in Figure 2.3-3[74].

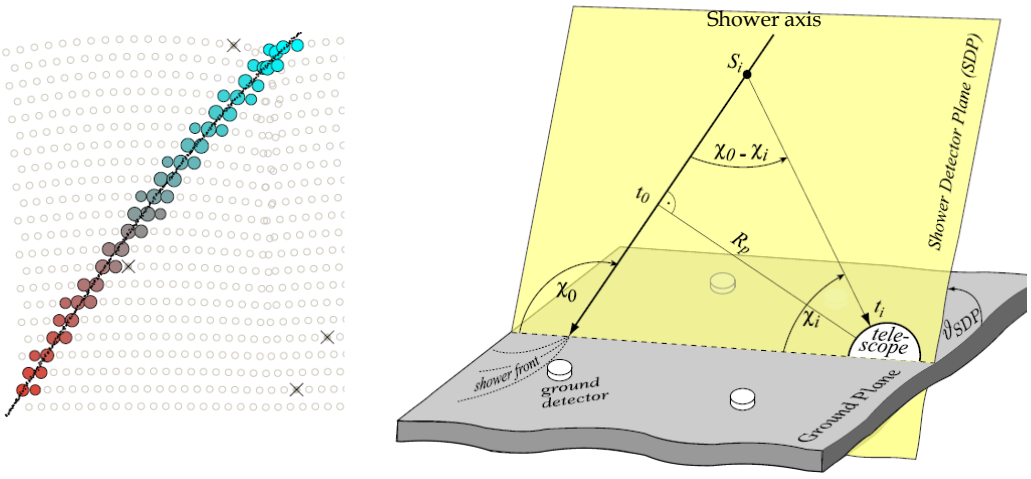


Figure 2.3-3 Light track on the camera (left) & illustration of shower geometry (right). The colour of the light track indicates the sequence of trigger time on each pixel. The crossed pixels are with signals but rejected during the reconstructing process. The shower detector plane is painted in yellow. The angle between the shower axis and the ground is χ_0, the viewing angle of pixel i within the plane is χ_i. And the distance between the shower axis and the FD station on the plane is R_p. The time that a photon travels along the shower axis and reaches the telescope via R_p is t_0.[74]

R_p is the impact parameter and shows the distance between the FD and the shower axis, t_0 is the time that particles travel along the shower axis and then reach the telescope via R_p. For each particle travelling at position S_i on the shower axis to the telescope, the angle between the shower axis and the FD on the SDP is $\chi_0 - \chi_i$. If the particles travel in straight lines with speed of light, the time it takes to reach the telescope is

$$t_i = t_0 + \frac{R_p}{c} \tan\frac{\chi_0 - \chi_i}{2} \qquad 34.$$

The predicted values of t_i are then compared to measured values by calculating the χ^2/ndf. The best fitted values of t_0, R_p and χ_0 are found when χ^2/ndf is minimized. So far the reconstruction of the shower is only based on FD signals, which is named as the mono-reconstruction. In addition, it is possible to combine SD and FD for the reconstruction, which is called the hybrid reconstruction. The hybrid analysis requires at least one FD and one SD station to be triggered. The shower axis therefore can be reconstructed with a better precision by including the start time of the SD stations that are not on SDP in the fitting process. An example of the improvement of the fitting result is shown in Figure 2.3-4.

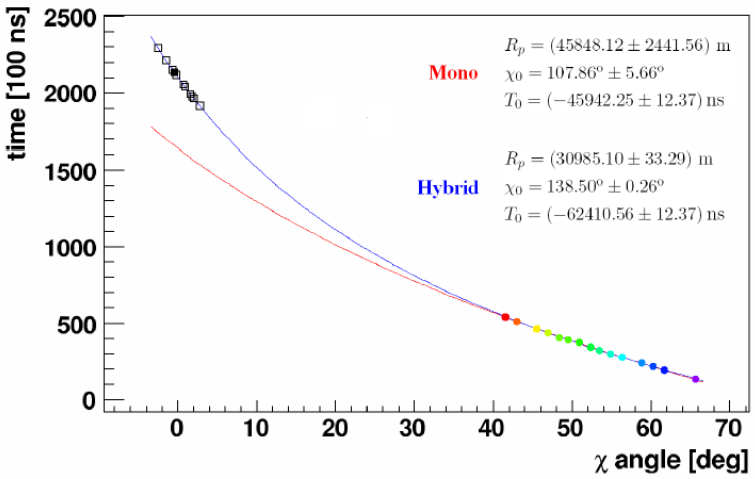


Figure 2.3-4 Fitting of shower geometry using data from FD-only (mono) and hybrid[75]

The blue line is from hybrid-reconstruction and the black points are from start time of SD stations. The uncertainties of R_p, and χ_0 have been decreased several orders of magnitudes by using the hybrid reconstruction. This is one of the unique features of using hybrid technique. By using cross-checks with laser shots from the Central Laser Facility, CLF), the core position is found typically with a resolution of 50 m and the arrival direction is with a resolution of 0.6˚.

Once the shower geometry is known, the signal in each pixel is converted to energy loss: an example of a shower profile which describes the energy deposit along the slant depth (the depth of atmosphere) is shown in Figure 2.3-5.

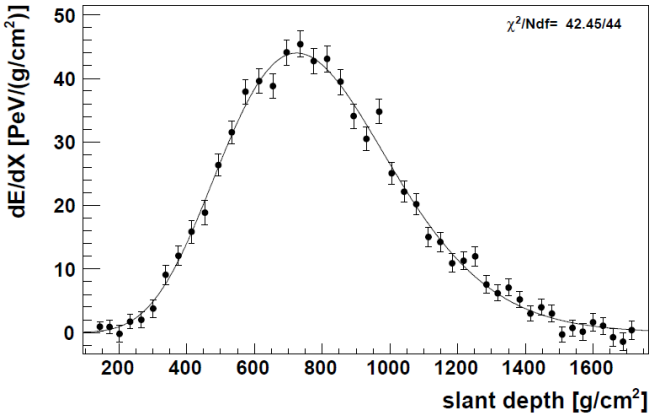


Figure 2.3-5 Energy deposit of the air shower with respect to the slant depth of the atmosphere (shower profile). The calorimetric energy of the shower is calculated by integrating the area under the profile.

The shower profile is fitted with the Gaisser-Hillas function[66]. The calorimetric energy is reconstructed by integrating the curve. However,

during the shower development, the energy of neutrinos, neutrons and almost all energy of muons cannot be measured by using FD. This fraction of energy is known as the invisible energy and the fraction depends on the primary energy and the mass composition of the cosmic ray. Using Monte Carlo simulations, Figure 2.3-6 illustrates the correction that has to be made with respect to different assumptions.

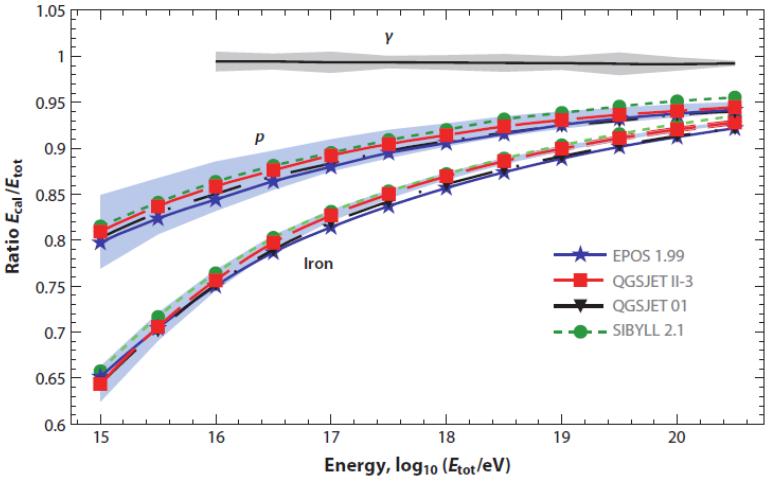


Figure 2.3-6 The fraction of calorimetric energy over the total energy of the shower for different primaries at various energies[6]

Photon showers are mostly electromagnetic so not surprisingly the correction is nearly 0. Iron showers have more invisible energy than proton showers because of a smaller fraction of electromagnetic component in the shower. Hadronic showers with lower energies produce pions with lower energies in the first interactions, which lead to a more rapid decay than interacting and therefore the number of generations of hadronic interactions is smaller. As a consequence the fraction of energy deposited in the hadronic component is larger than in case of higher energy showers. By assuming a 50:50 contribution of proton and iron, the measured calorimetric energy is corrected to compensate the problem brought by the invisible energy. The example of the event shown in Figure 2.3-5 is finally reconstructed with energy of 3×10^{19} eV.

The energy calibration of events with only SD data makes use of the unique hybrid design instead of relying on Monte Carlo simulations which introduce uncertainties from hadronic interaction models. Events that both have independent triggers from SD and FD are selected for the energy calibration. These events are called golden hybrid events. Observables that are obtained from SD such as S(1000) are used to estimate the energy of the shower by parameterising the correlations with the calorimetric energy seen

from FD. Shown in the left of Figure 2.3-7 is the relationship between the FD energy with S_{38}, which is the S(1000) if the shower is with zenith angle 38˚.

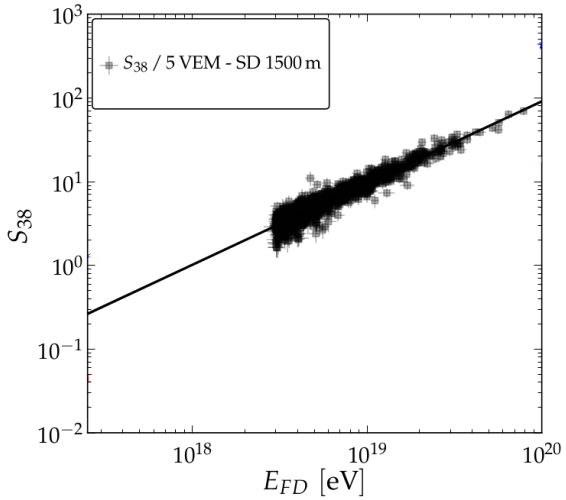


Figure 2.3-7 The energy calibration for SD based on hybrid data[27]

So far the techniques that are used in the Pierre Auger Observatory for detecting UHECR have been introduced. High quality data are continuously taken. Until December 2012, for zenith angle 0 – 60˚, there are 82,318 SD events with $E > 3 \times 10^{18}$ eV.

2.4 Simulations of extensive air showers

Monte Carlo simulations are essential for the interpretations of the mass composition of the UHECR. The simulations are complementary to the hardware of the Observatory. Software packages of increasing complexity have been developed to generate cosmic ray showers during the past 50 years. The COsmic Ray SImulations for KAscade (CORSIKA)[76] programme is currently the most commonly used package for simulating air showers with the highest energies. It computes interactions between cosmic ray particles in atmosphere in detail and tracks the shower development numerically. Additionally the CONEX[77] code tracks particles with energies over a threshold energy and then describes the shower development through solution of cascade equations[78]. CONEX improves the speed of simulations and still provides agreements of variables based on the longitudinal profile with results from CORSIKA[79]. However on the ground level for SD studies, only simulations from CORIKA are used as CONEX only works in one-dimension along the direction of the shower axis. In this thesis, most of the simulations were made using CORSIKA except for the analysis on π^0 as the leading particle of the hadronic interaction in Chapter 8 and the relationship between X_{max} and muon number in section 3.6.

Since the energies of the cosmic rays are much higher than energies of particles accelerated at the large hadron collider, hadronic interaction models need to be extrapolated by a factor of ~ 30 (in the centre of mass frame) for the highest energy cosmic rays from the experimental results. Two hadronic interaction models are included in this thesis, namely EPOS[80] and QGSJetII[81,82]. Both of the models are based on the Gribov-Regge theory[83,84] which considers soft hadronic interactions as multiple scattering processes of microscopic parton (quarks and gluons) cascades that are described using Pomeron[85] exchange in a phenomenological way. The difference between the two models is that QGSJetII is based on the Quark-Gluon String model[82] and can treat loop diagrams of Pomeron exchange to infinite order in the Gribov-Regge picture. The EPOS model on the other hand makes use of energy sharing between the involved partons but treats the higher order loop diagrams with an effective theory. It can however correctly predict many observables measured with data from e.g. the Relativistic Heavy Ion Collider. In their latest version, both models have been tuned to the latest LHC data. Comparisons of some of the quantities from the two models are shown in Chapter 8. Simulations used in Chapter 8 were produced on the Wuppertal Pleiades cluster by the author of the thesis.

It is important to be aware of the uncertainties arising from the hadronic interaction models when simulating hadronic showers. However, UHE photons entering the atmosphere firstly produce electron and positron pairs, which is a pure quantum electrodynamics process and is better understood than hadronic interactions. In this thesis, photon simulations used for photon-hadron separations by applying the entity method were simulated in the alicec cluster in Wuppertal. QGSJetII-03 was used and showers were simulated using CORSIKA by C. Bleve.

A proton with energy 10^{19} eV produces more than 10^{10} particles at X_{max}: it would take ~ 1.5 years to track all the particles for the full simulation using one CPU with 2 GHz. The storage of such shower takes ~ 5 TB which is not practical for producing a library of showers. As a solution, the statistical thinning was introduced by Hillas[86]. The idea is to use an algorithm that follows a fraction of particles with energies over a threshold and assign weights to represent particles that are not tracked. The thinning level ε is the ratio of the energy threshold over the initial energy of the air shower, i.e.

$$\varepsilon = E_{threshold} / E_{intitial} \qquad 35.$$

A value for ε of 10^{-6} is widely used in ultra-high energy cosmic ray analysis[87] and reduces the simulation time to only a few hours for a proton shower

with 10^{19} eV. This means that for this proton shower, all interactions and secondary particles are tracked if particle has the energy above 10^{13} eV. The probability of a particle being tracked is

$$F_i(E_i) = \begin{cases} E_i / \Sigma_k \overline{E_k}, & E_i < E_{threshold} \\ 1, & E_i \geq E_{threshold} \end{cases} \quad 36,$$

where E_i is the energy of the current particle and $\Sigma_k \overline{E_k}$ is the total energy of the sub-shower. The weight assigned to the particle to carry information of particles that are not followed is $w_i = w_{i-1} / F_i$.

Figure 2.4-1 shows a schematic illustration of the thinning procedure. A primary particle carries weight $w_0 = 1$ and interacts in the atmosphere. The secondary particles both have energy over $E_{threshold}$ and therefore F_1 and F_4 are 1. However during the next interaction, particles marked by the thin lines have energies below the threshold. As a consequence the particle drawn with a thick line (energy over the threshold) is kept and assigned the weight with $w_2 = w_1 / F_2$ and $w_5 = w_4 / F_5$ respectively. The same procedure goes on until the shower is propagated to the ground level.

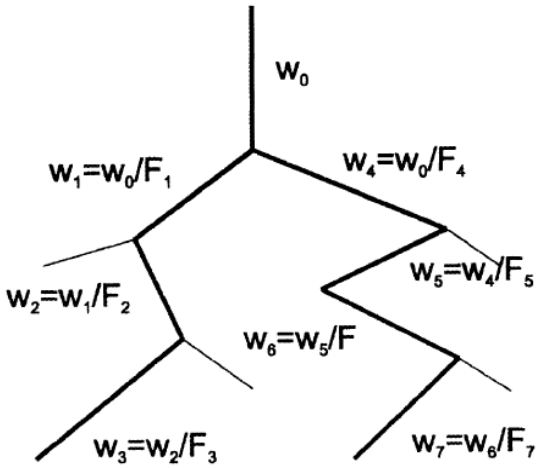


Figure 2.4-1 Schematic illustration of thinning procedures during shower simulations[88]. F_i is the probability of the particle being kept and w_i is the weight carried by the particle. The primary particle carries weight $w_0 = 1$ and both particles produced from the first interaction have energies higher than the energy threshold. Thin lines represent particles that are not tracked. Values of F_i are shown in equation 36.

At the ground level since the number of particles is significantly smaller than if no thinning was applied. The sampling area is increased and shown in Figure 2.4-2. The black dot is the detector, and the fan in grey is the increased sampling area. A particle carries heavy weight is injected to the

station multiple times with the time it enters the detector following a log-normal function, which brings artificial fluctuations and distorts the FADC trace. The procedure of un-weighting particles into the sampling area is often called de-thinning. These de-thinned particles from the same particle which carries high weight would have the same energy. As a result both the energy and the timing information of the arriving particles are lost through thinning and de-thinning process.

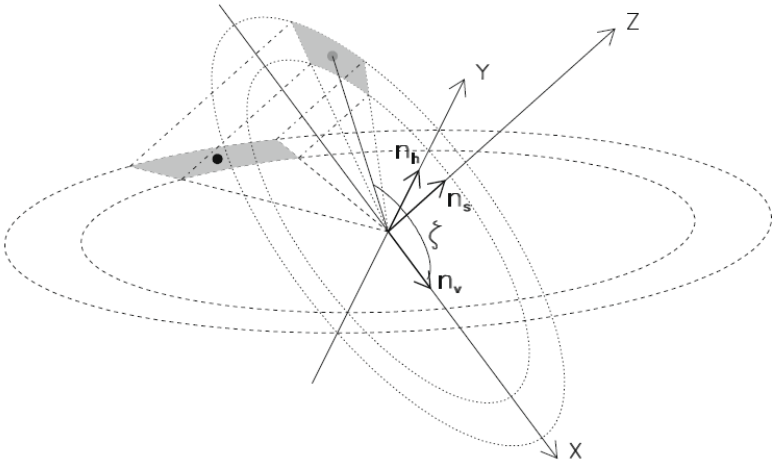


Figure 2.4-2 Sampling area for collecting particles with weight from thinning process[89]. The number of particles on the ground is decreased dramatically due to the thinning procedures. To increase the probabilities of receiving particles, the sampling area is artificially increased. The centre of the ring is the core position. The black dot is a surface detector with core distance r. Z axis is the shower axis and the x-y plane is perpendicular to the shower plane. The grey area on the ring of the x-y plane is projected to the ground and forms a fan shape, which has the outer and inner radius of r ± 10% r and the polar angle with the opening window of $\Delta\zeta = \pm 8.6^{\circ}$. Particles that arrive in the fan are considered that reached the detector.

It is important to understand how thinning and de-thinning procedures create artificial defects in simulations. To study this, unthinned showers were produced in Leeds[90] using the arc1 cluster. Ten showers of proton, iron and photon primary have been shared in the HPSS (High Performance Storage System) in Lyon ($HPSS_PAUGER/lulu/Unthinned/ and $HPSS_ PAUGER/bruijn). By running sub-showers on the parallel CPUs, it took 7.5 days[87] instead of 1.5 years to simulate an unthinned proton shower of 10^{19} eV. Showers were simulated using CORSIKA and the QGSJetII-03 model. The unthinned showers are used in section 6.5.2.2 and the effects of thinning and de-thinning are discussed in section 9.4.

In this chapter, 'tools' that are needed for the analysis of air-showers have been introduced. In the next chapter, signatures of photon-induced air showers are presented. A review on the recent results about the upper limit

of photon flux from the Pierre Auger Observatory and Telescope Array is included.

Chapter 3. Photon-hadron separation methods and results

Air showers induced by a photon primary develop slower in the atmosphere and contain fewer muons compared to showers induced by hadrons. The hybrid design (FD and SD) of the Pierre Auger Observatory allows one to measure both the longitudinal development of the shower and the footprint of signals from muons, electrons and photons on the ground. Observables based on data from FD and SD have been explored to separate photon-induced showers from hadron-induced showers, which are introduced in this chapter. Recent results on the upper limit to the photon flux from the Pierre Auger Observatory and the Telescope Array are summarised. Although at the moment no experiment is designed to measure the muon component of the air shower directly, the role of the muon in photon-hadron separation is discussed in section 3.6. In the absence of direct muon measurements, the entity method was proposed and described in Chapter 4.

3.1 Photon search using FD data from Auger

The first report from the Pierre Auger Collaboration[91] of a search for UHE photons was based on measurements of the depth of shower maximum, X_{max} using the fluorescence detectors. In section 1.3 it was demonstrated that X_{max} is a robust parameter for mass composition study due to protons on average having a larger X_{max} than iron showers. Figure 3.1-1 shows the mean of the shower profiles from different primaries.

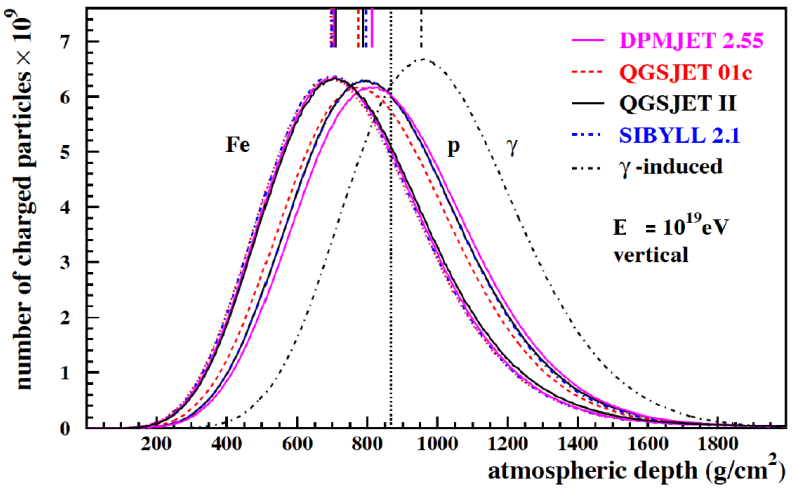


Figure 3.1-1 Mean of shower profiles in the atmosphere for iron, protons and photons[92]. Various hadronic interaction models are used for iron and proton simulations

Unlike simulations of proton and iron showers that are sensitive to uncertainties from hadronic interaction models, photon simulations are based on well-known electromagnetic processes. X_{max} of photon showers is on average ~ 950 g cm^{-2} at 10^{19} eV, which is more than 200 g cm^{-2} larger than the mean of proton showers. The reason photons develop slower in the atmosphere is mostly due to the smaller multiplicity per interaction. Furthermore, the Landau-Pomeranchun-Midal (LPM) effect[93,94] suppresses the cross-section of pair-production, which makes the shower penetrates even deeper at energies above 10^{19} eV.

The search requires at least one detector on the ground to be triggered, to improve reconstruction as explained in section 2.3. The shower of maximum X_{max} is required to be seen in the field of view (Figure 3.1-2).

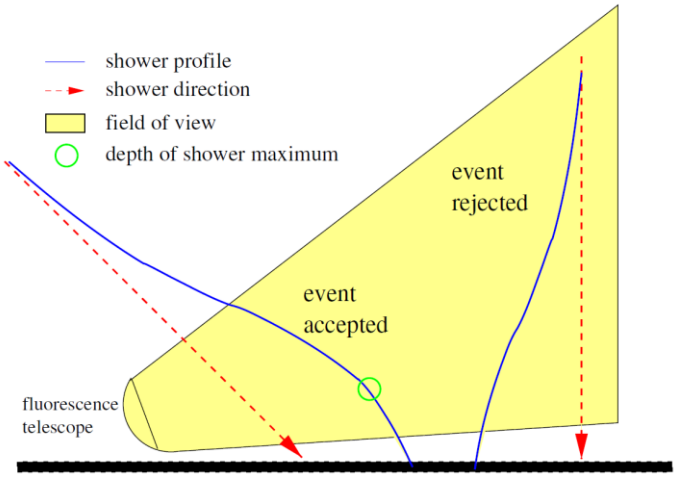


Figure 3.1-2 Field of view[91] cut for X_{max} analysis. Only events with X_{max} within the yellow region are selected.

The field of view from the telescope (30° × 30°) is marked in the yellow area. The green circle indicates the position of the X_{max} of the shower on the left. The shower on the right has X_{max} below the ground and is rejected. This cut is mostly sensitive to showers with nearly vertical zenith angles and showers with high energies. The fact that photons have deeper X_{max} than protons means they are less likely to survive this cut, which lead to a cut on the zenith angle that vertical showers are not included in the analysis. There are also other cuts implemented such as the quality of the profile fit, the amount of cloud at the time of the data-taking, the distance of the telescope from the core and the number of phototubes in the FD triggered by the shower.

The depth of shower maximum is measured by finding the atmospheric depth at which the maximum amount energy is deposited by the air shower.

An example of a shower profile is shown in Figure 3.1-3 (left). The value of X_{max} corresponds to the peak of the profile (measured at the Auger Observatory with a resolution of ~ 20 g cm^{-2}) and is at 780 g cm^{-2}. The energy of the event is reconstructed as 1.6 × 10^{19} eV. More about procedures of the shower reconstruction were introduced in section 2.3. The X_{max}-value is then compared to the distribution obtained from photon simulations for the same energy and zenith angle as shown in the right of Figure 3.1-3.

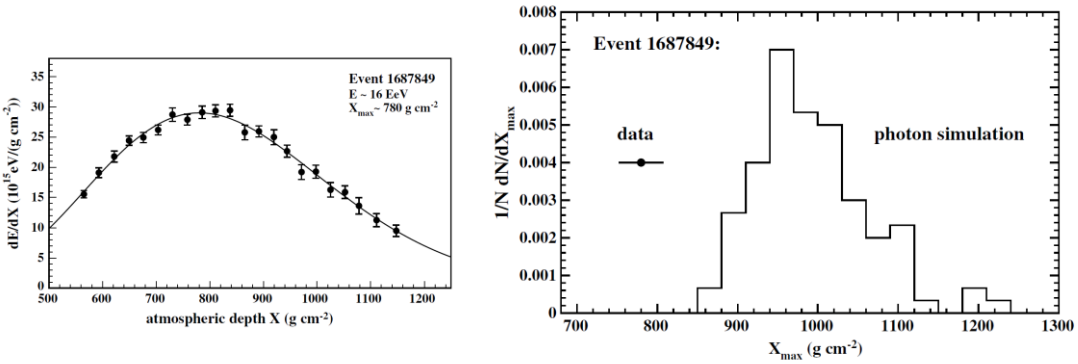


Figure 3.1-3 Example[91] of a shower profile from data (left) and the measured X_{max} compared to photon simulations (right)

The expected mean X_{max} from distributions of photon simulations is ~ 1000 g cm^{-2}. The example event has X_{max} = 780 g cm^{-2} and is ~ 2.9 standard deviation from the mean of photon simulations. An upper limit to the photon fraction in UHECR were calculated based on the 29 events with measured X_{max} and is shown in Figure 3.1-4.

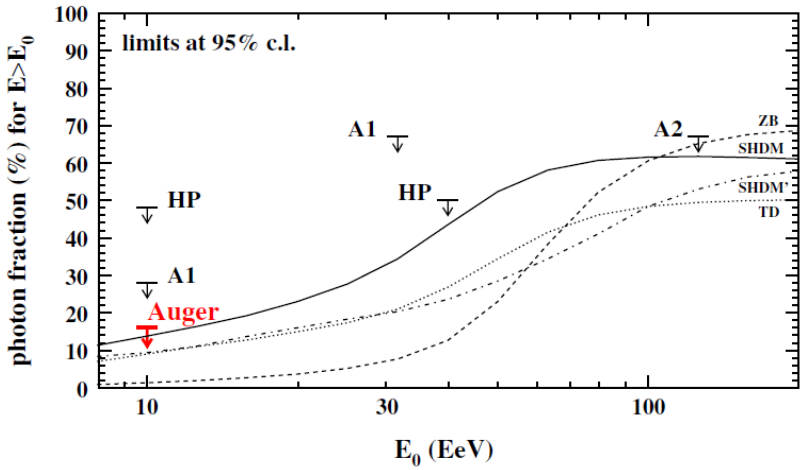


Figure 3.1-4 The first result from the Pierre Auger Collaboration for the upper limit to the photon fraction using X_{max}[91]. Details of results from Haverah Park (HP) and AGASA (A1, A2) have been discussed in section 1.4.3

It was the first time to search for UHE photons using X_{max}. The result showed an upper limit of 16% of photon fraction for UHECR with energy

above 10^{19} eV. The upper limit was updated in 2009[95]. The median of the X_{max} distribution of photon simulations is used for selecting photon-like events from data. This means only 50% of the photons that survived all previous cuts would be identified as photon candidates. Selection efficiencies are accounted by assuming an E^{-2} spectrum of the photon flux. 8, 1, 0 and 0 photon candidates were found for energy above 2, 3, 5 and 10 EeV but all treated as hadronic backgrounds. Upper limits of 3.8%, 2.4%, 3.5% and 11.7% (at 95% c.l.) are found on the photon fraction for $E > 2 \times 10^{18}$ eV, $> 3 \times 10^{18}$ eV, $> 5 \times 10^{18}$ eV and $> 10^{19}$ eV respectively.

X_{max} is the only direct measurement for estimating the mass composition of the cosmic ray. The dependence of the mean of X_{max} on the energy of iron, proton and photon are shown in Figure 3.1-5. Data from various experiments are plotted for comparison. The sudden decrease of the elongation rate (introduced on page 13) at $E \sim 10^{19.6}$ eV is due to photons interact with the earth magnetic field before entering the atmosphere, as

$$\gamma + B \rightarrow B + e^+ + e^- \qquad 37.$$

This interaction creates pre-showers and makes the photon shower develop more rapidly. Showers arriving from south are more likely to be affected by the earth's magnetic field than showers from the north due to the orientation of the magnetic field lines. Photons make such interactions are often named 'converted-photons' and have X_{max} more proton-like.

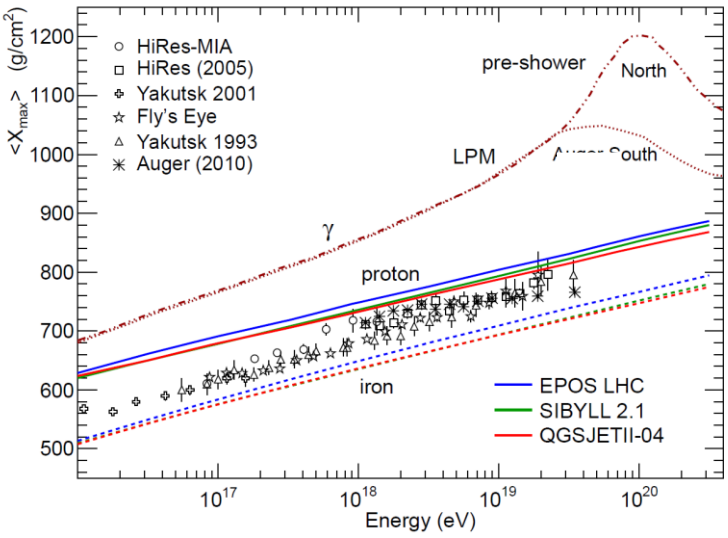


Figure 3.1-5 Mean X_{max} and its dependence on the energy for p, Fe and γ simulations[96]. Pre-showers start to occur when the energy of the cosmic ray is larger than $10^{19.6}$ eV. The probability of being pre-showers is related to the direction of the cosmic rays.

The energy of the pre-shower is too high for the sensitivity of FD. However, the converted photons are still mixed in the simulation sample to represent a reliable library of photons with E^{-2} spectrum.

It is only feasible to use SD data to search for photons with energy ~ 10^{19} eV due to the limitation of the exposure for fluorescence telescopes (duty cycle ~ 13%). In the next section, SD parameters that are sensitive to photon-hadron separations are introduced.

3.2 SD parameters for hadron-photon separation

The flux limits reached with the florescence detectors for UHECR of energy above 10^{19} eV can be improved using the surface array providing appropriate parameters can be discovered, as there are many more events. In this section, relevant parameters that can be measured using the water-Cherenkov detectors are described. Their use in several photon searches will be discussed.

3.2.1 Risetime

Muons travel nearly in straight lines while the electromagnetic component of the shower scatters and on average arrive after muons. The spread of the arrival time is related to the geometry of the shower. An illustration of particles arriving at the detector 1 on the ground via path P_1 and $L + P_2$ is shown in Figure 3.2-1 (left).

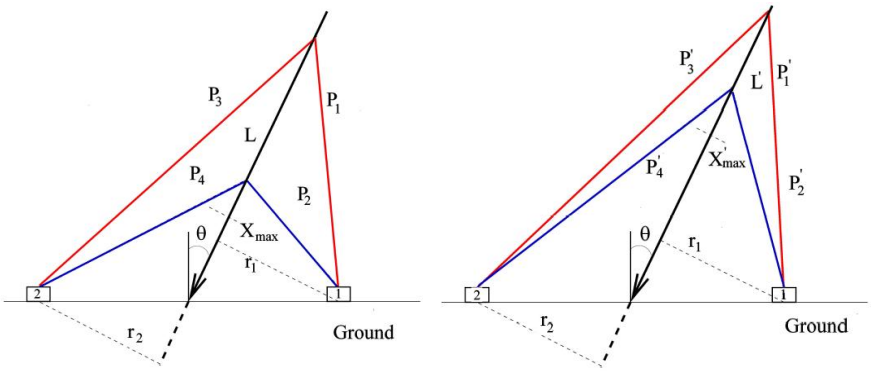


Figure 3.2-1 Schematic diagrams show the relationship between the shower geometry and the delay of particles arriving on the ground[97]. The shower on the left penetrates deeper in the atmosphere than the shower on the right.

The intersection of the red lines indicates a region of shower development where muons are continuously produced. The earliest particle arrives at detector 1 travels along P_1. The shower develops in the direction of the shower axis. For particles travel along L and reach detector 1 via P_2, the delay of arrival time compared to particles arrive via P_1 is $(L + P_2 - P_1) / c$, if

one assumes particles travel in the speed of light. The shower shown on the right develops faster in the atmosphere, for a station also located at distance r_1 from the shower axis, the time of delay for the particle from L′ compared to particles from the early production point is L′ + P_2′ – P_1′. From simple geometry it is clear that the delay of the left case is larger than the right one. As a consequence for air shower with the same zenith angle, the more-slowly developed shower has larger delays of arrival time between the early and late particles at a station with the same distance away from the shower axis.

To describe the spread of arrival time of particles that reach ground detectors, the risetime parameter, $t_{1/2}$, was introduced during the operation of the Haverah Park array[98,99]. Water–Cherenkov detectors of 34 m² and 1.2 m deep (the depth is the same as the Auger detectors) were used. Using this technique, it was shown for the first time that shower-to-shower fluctuations were detectable. It was impossible to deduce the mass composition using $t_{1/2}$ due to limitations of shower simulations at the time. However, today the risetime parameter plays an important role for the standard Auger photon search.

As introduced in section 2.1, FADC traces with sampling rate of 25 ns are used to measure the signal responses from SD. Shown on the left of Figure 3.2-2 is an illustration of the traces from a hadron shower (red) and a photon shower (blue). The time spread of the trace from the hadron shower is smaller than that from the photon shower because the hadron shower develops more rapidly. If the time it takes for the signal to rise to 10% of the total signal is t_{10} and the time it takes for the signal to rise to 50% of the total signal is t_{50}, risetime $t_{1/2}$ is defined as $t_{50} - t_{10}$.

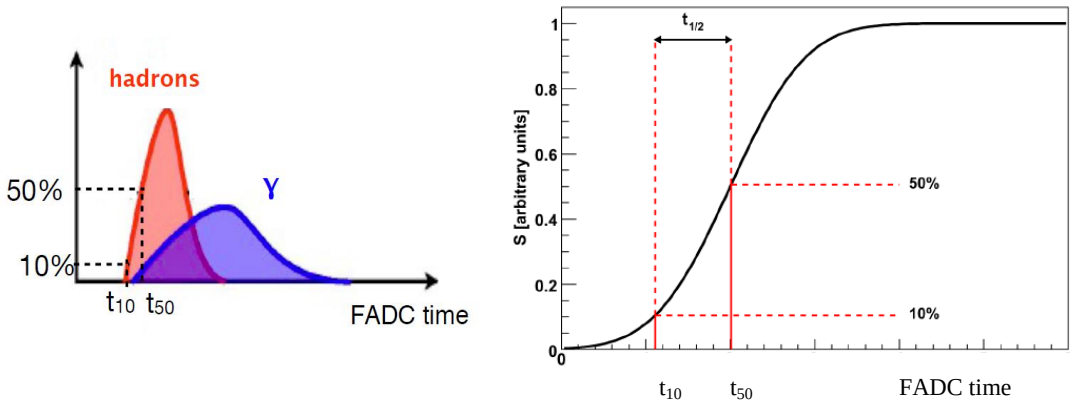


Figure 3.2-2 Cartoon to illustrate that the time trace of detector triggered by hadron shower is shorter than from a photon shower[97] (left) and the definition of $t_{1/2}$ described using integrated FADC trace

From the shower geometry, it is clear that the risetime increases with the core distance and decreases with respect to increasing zenith angle. The mean of risetime from data can be fitted with respect to the core distance using a second-order-polynomial function and shown in Figure 3.2-3[100].

The risetime at 1000 m from the core position of vertical showers for proton primary is on average ~ 300 ns and for photon showers is ~ 450 ns. The measurement uncertainty was obtained by using pair stations, which are detectors with similar distance from the core[101]. It is typically ~ 50 ns at 1000 m from the core position for data, which is consistent with results[98] from the Haverah Park experiment. More details of differences of FADC traces between hadron shower and photon showers using Monte Carlo simulations are shown in section 4.3.

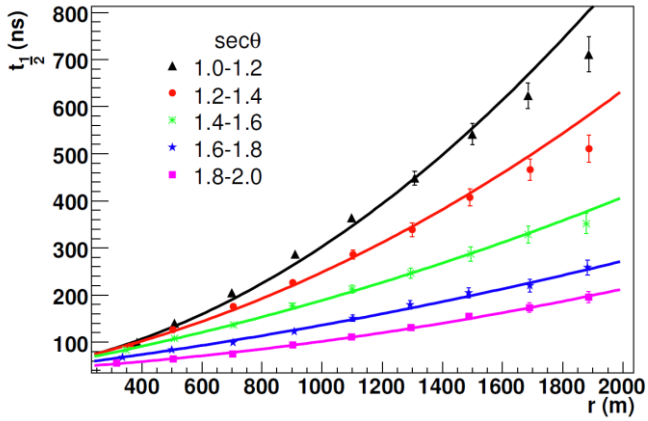


Figure 3.2-3 Mean of risetime for detectors binned in r and sec θ from data[100]

The minimum value of risetime is 40 ns due to the mean of the single particle response in the water-Cherenkov detector. The function fitted to the relationship between is $t_{1/2}$ and distance is

$$t_{1/2} = (40 + a\,r + b\,r^2) \qquad 38,$$

where a and b have dependence on the zenith angle. For photon searches two approaches using $t_{1/2}$ have been adopted.

3.2.1.1 Risetime delta

Equation 38 is also known as the 'benchmark' function of the risetime. The <Δ> method, which has been explored exclusively by the Leeds group[97,101,102] makes use of the fact that photon showers have larger risetime and the measured values should be above the benchmark function obtained from data. Figure 3.2-5 shows how to calculate Δ of one station.

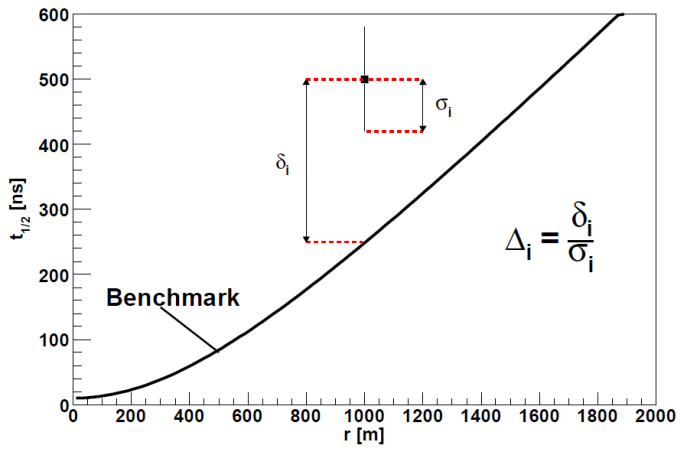


Figure 3.2-4 An example of the measured risetime compared to the benchmark[97] obtained from data.

Δ is the ratio of the deviation (δ) of the measured $t_{1/2}$ from the benchmark over the uncertainty (σ) that parameterised from data. For the i^{th} station, Δ_i is calculated. In the end $<\Delta>$[97,101] is the event parameter by averaging over all Δ_i. Upper limits (95% c. l.) of 1.5%, 3.2% and 2.9% were obtained on the fraction of photons above 10, 20, and 30 EeV[102].

3.2.1.2 Risetime 1000

Alternatively the risetime at 1000 m can be used as an event parameter. Equation 38 is fitted to risetime measured in each station that with non-saturated signal > 10 VEM. At least four stations are needed for such fit and the typical value for a and b are 50 ns km^{-1} and 100 ns km^{-2} respectively. The risetime at 1000 m, $t_{1/2}(1000)$ then can be estimated using the fitted curve.

$<\Delta>$ and $t_{1/2}(1000)$ are both based on the risetime information but use different methods to generate the discriminant on the event level. The former method does not request as many as stations as the latter method, which increases the selection efficiency. However, for the same reason, the fluctuations of $<\Delta>$ are larger than the spread of $t_{1/2}(1000)$. There are further studies regarding which parameter is the best for photon-hadron separation that are on-going within the Wuppertal group. The result published by the Pierre Auger Collaboration in 2008[103] used $t_{1/2}(1000)$ and combined it with the radius of curvature, which is introduced in the next sub-sections.

The dependence of $t_{1/2}(1000)$ on the zenith angle (sec θ) is shown in Figure 3.2-5. Photon simulations and data are with $E > 2 \times 10^{19}$ eV.

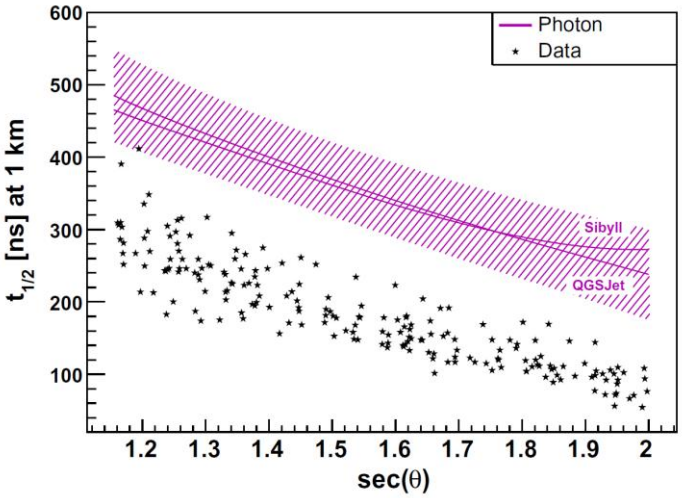


Figure 3.2-5 The dependence of $t_{1/2}(1000)$[103] on sec θ for photon simulations and data

It is difficult to set only one cut for selecting photons using $t_{1/2}(1000)$ since it has dependence on the zenith angle. The solution is to introduce a final parameter $\Delta t_{1/2}(1000)$, which describes the deviation of the measured $t_{1/2}(1000)$ from the expected mean of photon:

$$\Delta t_{1/2}(1000) = (t_{1/2}(1000) - \langle t_{1/2}(1000) \rangle_\gamma) \,/\, \sigma_\gamma(S(1000), \theta) \qquad 39.$$

$\sigma_\gamma(S(1000), \theta)$ is 1 σ of the photon distribution and parameterised as a function of S(1000) and sec θ. This treatment makes the median of photon sit at ~ 0, which allowing photons of all zenith angle to be plotted in one histogram (Figure 3.2-6).

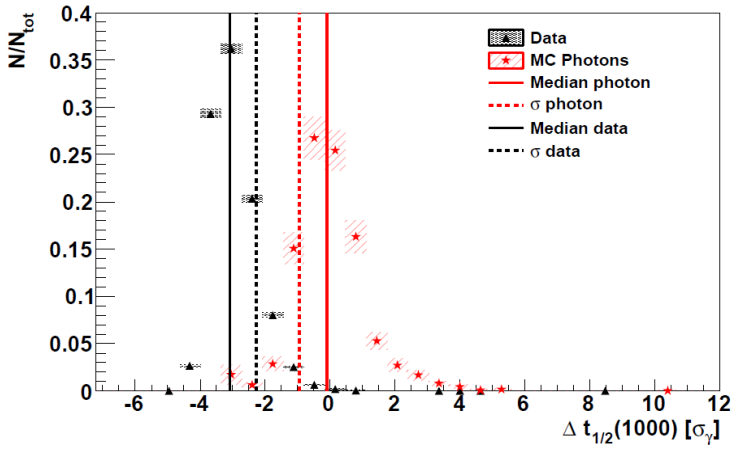


Figure 3.2-6 Distribution of $\Delta t_{1/2}(1000)$ of photon simulations and data[104] (E > 10 EeV). Pre-showers are not included in this plot, the zenith angle range is 30 < θ < 60°.

Solid lines show the position of median (μ_1 and μ_2) of two distributions and dashed lines are 1 σ (σ_1 and σ_2) away from the median.

The Merit factor is used to qualify the separation between two distributions, it is defined as

$$\eta = (\mid \mu_1 - \mu_2 \mid) / \sqrt{\sigma_1^2 + \sigma_2^2} \qquad 40.$$

The larger the Merit factor, the larger separations are between two distributions. The Merit factor is 2.6 and for $E > 10^{19}$ eV and $30 < \theta < 60°$. The plot does not include pre-showers, which makes η larger than the relativistic scenario (that there are pre-showers).

As a summary, the risetime parameter describes the spread of the arrival time of particles. It is related to the geometry of the shower development especially the first interaction and the mean-free-path of pions. Moreover, the rich muon content in the hadron showers makes the difference between hadron and photon even larger.

3.2.2 Radius of curvature

The particles that arrive earliest at the ground in hadron showers are muons that were from charged pions, which were produced in the first interaction[105]. Photon showers develop slower in the atmosphere and with fewer muons. The first particles arrive in photon showers are mostly electromagnetic component, which was created at the altitude much lower than the first interaction point of the hadronic interaction and scatters before reaching the ground. The shape of the shower front (formed by the particles arrive first) carries information of the geometry of the shower development and the mass composition of the primary. The idea of measuring the radius of curvature was firstly reported by Linsley and Scarsi in 1962[105]. The radius of curvature is still used for shower reconstructions as introduced in section 2.2 and also used for photon-hadron separation (Figure 3.2-7).

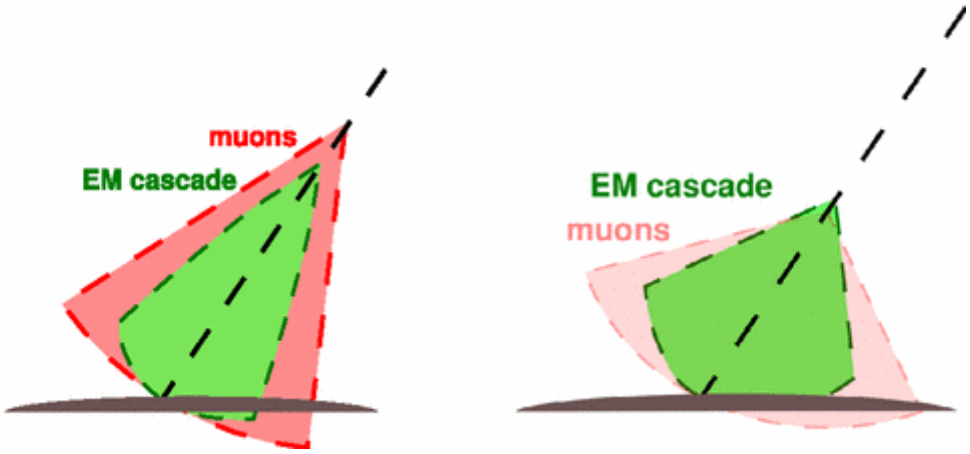


Figure 3.2-7 Cartoon[106] showing the shower front of a hadron shower (left) is less curved than the shower front of a photon shower (right)

The radius of curvature is related to the delays of the trigger time of each detector with respect to the plane that was fitted in the reconstruction process (Figure 2.2-1). Photons have more curved shower front with smaller radius of curvature than hadrons. Figure 3.2-8 shows the dependence of radius of curvature on the zenith angle (sec θ) for photon simulations and data with $E > 2 \times 10^{19}$ eV.

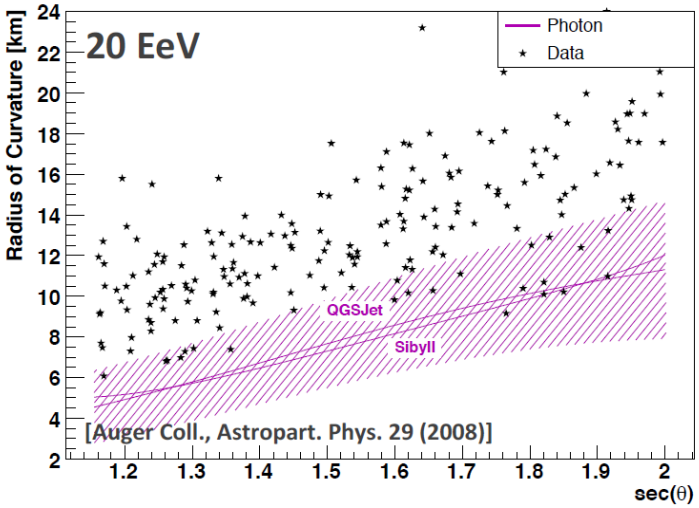


Figure 3.2-8 Radius of curvature[103] of data and photon simulations at different sec θ. The band shows the rms of the QGSJet01 model.

The larger the zenith angle, the longer the shower develops in the atmosphere and therefore the distance from the ground to the depth of the first interaction point is further away. This leads to radius of curvature increases with the increasing zenith angle. To construct a final parameter that does not have dependence on the zenith angle, the same technique used in $t_{1/2}(1000)$ has been applied. ΔR_c is the relative radius of curvature with the reference of mean of photon simulations at zenith angle θ and signal at 1000 m S(1000).

$$\Delta R_c = (R_c - \langle R_c \rangle_\gamma) / \sigma_\gamma(S(1000), \theta) \qquad 41.$$

$\sigma_\gamma(S(1000), \theta)$ is 1 σ of the distribution of radius of curvature from photon simulations. ΔR_c is ~ 0 for photons and > 0 for data as shown in Figure 3.2-9[104]. At least 5 stations per event are needed for the calculation of the radius of curvature. Only events with $30 < \theta < 60°$ and $E > 10^{19}$ eV are selected. It was decided that instead of introducing uncertainties from hadronic interaction models via using proton simulations, only photon simulations and data are used. Pre-showers are not included in the photon simulations.

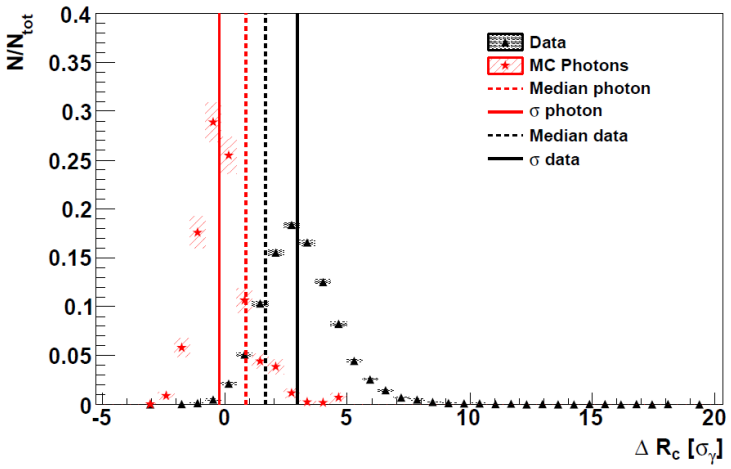


Figure 3.2-9 Distribution of ΔR_c of photon simulations and data[104]

The Merit factor is 1.9, which is poorer than for $\Delta t_{1/2}(1000)$. The radius of curvature and risetime are combined to formulate an event parameter, which is introduced in section 3.3.

3.2.3 Photon-hadron separation using lateral distribution function

The lateral distribution function (LDF) describes the relationship between the total signal of the detector and the distance of the detector away from the shower axis (core). It has been discussed in section 2.2 and an example is shown in Figure 2.2-3. The fall-off of the signal with distance is related to various physics interactions and the transverse momentum, p_t, of the products. The opening angle of each pair-production is typically ~ $mc^2/h\nu$, where $h\nu$ is the energy for the primary photon. Electrons and positrons then produce photons through bremsstrahlung. The energy spectrum of the photons is relatively flat and the low energy photons are emitted at large angles. Charged electromagnetic components with low energy lose energy quickly through ionisation during the propagation. The root mean square scattering of electrons is the key reason for the lateral spread of charged particles in the air shower. At the critical energy of ~ 84 MeV, an electron travelling 1 radiation length on average scatters 14°. The formula to calculate the scattering angle of an electron was introduced by Molière

$$\theta = (21\ \text{MeV}/p\ \beta)\ \sqrt{(x/X_0)} \qquad 42.$$

X_0 is the radiation length and p is the initial momentum of the electron. The spread of electrons (Molière radius) on the ground level of the Pierre Auger Observatory is ~ 70 m. The scattering of muons on the other hand can be ignored as they carry much larger mass. Often they are considered to travel in nearly straight lines.

For each hadronic interaction, the distribution of the transverse momentum of secondary particles is described by

$$g(p_t)dp_t \propto p_t e^{-\frac{p_t}{\langle p_t \rangle}} \qquad 43,$$

where $\langle p_t \rangle$ is the mean of the products with a typical value of 350 MeV/c. The deflection due to the hadronic interaction is greater than that due to the electromagnetic scattering and is responsible for the lateral distribution of muons.

As a consequence the signal falls faster from the shower axis for the electromagnetic component than for the muon component. A plot to help with visualise such feature is shown in Figure 1.3-3 (left). Since photon showers are mostly composed of electromagnetic component while hadron showers contain larger fraction of muons, the LDF of two primaries are therefore different and an illustration is shown in Figure 3.2-10.

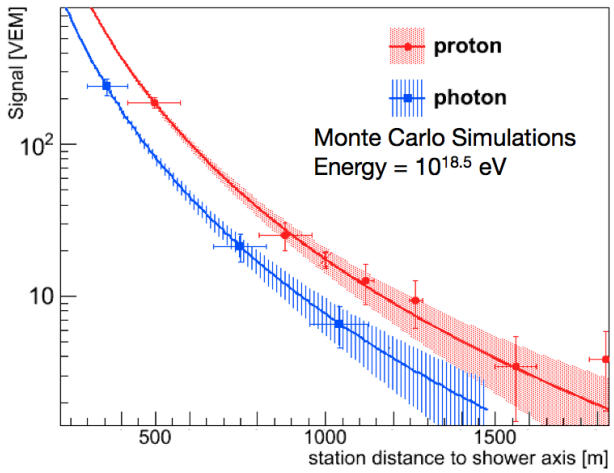


Figure 3.2-10 Lateral distribution function[107] for proton and photon showers with the same Monte Carlo energy

Proton and photon simulations are with the same Monte Carlo energy $10^{18.5}$ eV and vertical. The muon signals from hadronic showers make the energy deposit of each station larger than detectors in the photon shower. In principle one could fit the slope of LDF and use it as the mass-sensitive parameter. However, it is clear from the demonstration that the number of triggered detectors for photon showers with the same Monte Carlo energy and zenith angle is smaller than that of hadron showers. The minimum number of stations needed to fit the slope parameters is 5, which is a strong cut for photon simulations at vertical angles. The alternative is to use the so-called S_b parameter. It does not have a minimum number of stations cut and defined as

$$S_b = \sum_i S_i \, (r_i/1000 \text{ m})^b \qquad 44.$$

S_i is the signal measured at the i_{th} station, r_i is the core distance and b is the factor that should be optimised to achieve the best photon-hadron separation. The optimum separation between data and photon is achieved by setting b = 4 and shown in Figure 3.2-11.

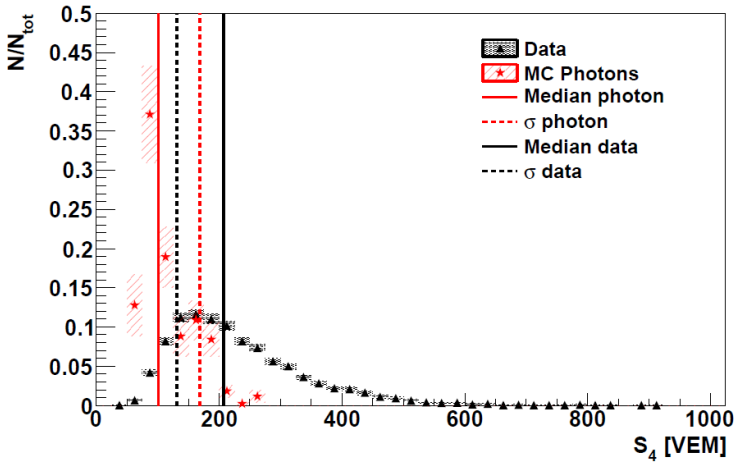


Figure 3.2-11 S_4 distribution[104] for photon simulations and data of E > 10^{19} eV. S_4 is a parameter based on the lateral distribution function.

It should be noted that since both distributions are asymmetric and non-Gaussian, one standard deviation is defined as the position that the area enclosed from the median (the solid lines) to 1 σ (the dashed lines) is 34% of the total area. The Merit factor is < 1.3, which is even worse than that of the radius of curvature.

Although the separation power from S_b is not desirable, the fact that it does not request high number of triggered stations in the event makes it perfect to be combined with X_{max} for the hybrid photon search. Furthermore, in section 9.1, more exploration based on the lateral distribution function are described and it is a part of the future work to construct an event likelihood parameter that uses both information of signal distribution and particle arrival times for the photon-hadron separation for SD photon-search.

3.3 Flux limits using the radius of curvature and the risetime

To avoid uncertainties from hadronic interactions models, it was decided that instead of using proton simulations, 5% of data from 2004 to 2006 were randomly selected for optimizing separations of photon and data (because it had been shown almost all data are hadrons). Events with 30 < θ < 60° and E > 10^{19} eV are selected. Only non-converted photons are in the photon simulations. However, this is not a realistic scenario and work done in this

thesis always includes converted-photons. The photons are assumed to follow E^{-2} spectrum. The left of Figure 3.3-1 shows the two-dimensional separation achieved using two parameters. The deviation of R is ΔR_c and the deviation of $t_{1/2}$ is $\Delta t_{1/2}(1000)$.

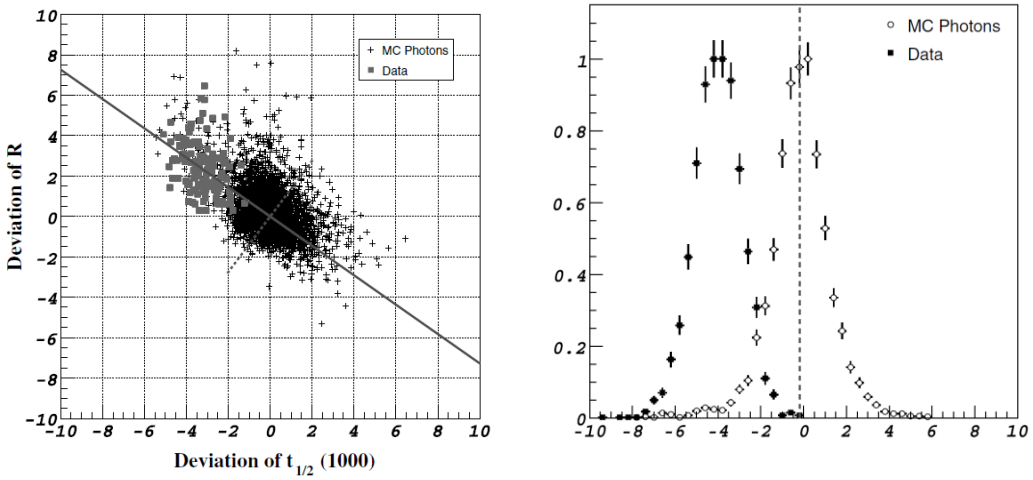


Figure 3.3-1 Photon-hadron separation by combined analysis of $\Delta t_{1/2}(1000)$ and ΔR_c[103]
On the left is a 2D plot to show the separation. 5% of data were used. The solid line is the principle component axis and the dashed line is perpendicular to it. The plot on the right is the projection of all points from the 2D plot on the component axis but used the rest 95% of data. The dashed line shows the median of the photon distribution.

The principle component axis (PCA), defined as the axis with the maximum variance of two clusters projecting on the axis, was derived based on 5% of the data. Then the rest of 95% data were projected on the PCA and the response is shown in the right of Figure 3.3-1. The median of the photon distribution is set as the cut for photon candidates.

No photon candidates were found. An upper limit of photon fraction 2%, 5.1% and 31% were found for UHECR with energy $E > 10^{19}$ eV, $E > 2 \times 10^{19}$ eV and $E > 4 \times 10^{19}$ respectively[103].

Soon after this paper, photon-hadron separation based on the lateral distribution function and Xmax was published.

3.4 Flux limit using hybrid data

The most recent report from the Pierre Auger Collaboration on photon search was presented in ICRC, Beijing in 2011. It is a hybrid analysis that uses X_{max} and S_b. The same cuts have been applied as introduced in section 3.1. Figure 3.4-1 shows the two-dimensional separation of proton and photon[107]. Simulations of energy 10^{18} – $10^{18.5}$ eV and zenith angle $\theta < 60$ °

were selected. The Fisher axis, which is defined as the axis that provides the best separation (e.g., quantified using Merit factor) between two clusters by projecting all points to such axis, was determined using photon and proton simulations.

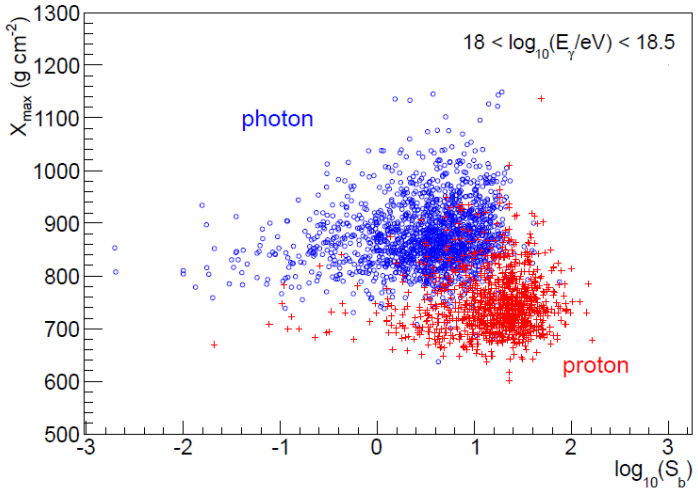


Figure 3.4-1 X_{max} and S_b for the hybrid photon-search[107] of $0 < \theta < 60°$

Points are projected on the Fisher axis to construct the final 1D parameter (Fisher response). The median of the photon distribution is taken as the cut to select photon candidates (Figure 3.4-2).

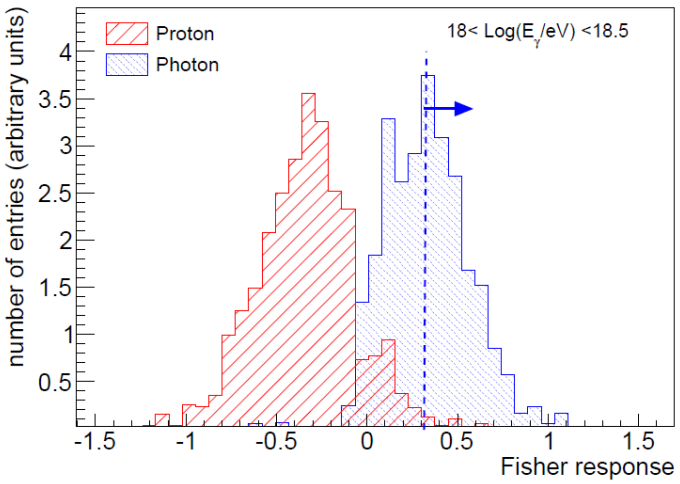


Figure 3.4-2 Fisher response of X_{max} and S_b for proton and photon simulations[107]

The same procedure was applied to data from January 2005 to September 2010. 6, 0, 0, 0, and 0 photon candidates were found for energies above 10^{18}, 2×10^{18}, 3×10^{18}, 5×10^{18} and 10^{19} eV respectively. The number of candidates is consistent with the expectations from proton simulations and therefore all of them are considered as hadronic backgrounds. As the result at 95% of confidence level, the upper limit of photon fraction in the five energy bins stated are 0.4%, 0.5%, 1.0%, 2.6% and 8.9%. This is yet the lowest limit ever

set for this energy range below 5×10^{18} eV. However, photons that produced from the GZK interactions are around 10^{19} eV, which is where the hybrid analysis is least sensitive.

So far all results based on analysis from the Pierre Auger Collaboration on diffuse photon search have been introduced. The next section is about the photon search from the largest UHECR observatory in the North hemisphere, the Telescope Array (TA) in States.

3.5 Photon search using data from Telescope Array

The Telescope Array is located in central Utah and covers an area of 700 km^2, which is less than ¼ of area of the Auger Observatory. It also uses a hybrid technique but the surface array is built of scintillators rather than water-Cherenkov detectors. Each station contains two layers of plastic scintillators with each of thickness 1.2 cm and area 3 m^2. The spacing between detectors is 1200 m on average and in total there are 507 stations. Unlike water-Cherenkov detectors, scintillators respond equally to electric and muon components of air shower. This makes the surface array of TA less sensitive to muon signals.

The analysis[106] carried out was based on SD data that were taken from May 2008 to May 2011. The photon-hadron separation parameter was derived by fitting the time delay of the shower front (Figure 3.5-1) with respect to a plane.

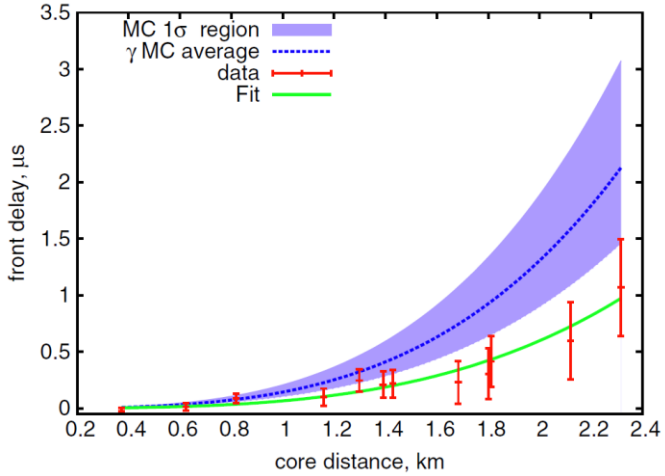


Figure 3.5-1 Example of shower front fit from data compared to the mean of photon simulations from the TA collaboration[106]

The lateral distribution is fitted at the same time with the shower front; at least 7 stations are needed in this process. The equations are summarised as

$$S(r)=S(800)\times LDF(r) \quad 45.$$

$$t_0(r)=t_0+t_{plane}+a\times 0.67\left(1+\frac{r}{30}\right)^{1.5} LDF^{-0.5}(r) \quad 46.$$

S(800) is the signal at 800 m, which is equivalent as S(1000) in Auger analysis. LDF(r) is the lateral distribution function, t_0 is time of the plane and t_{plane} is the delay of the shower front with respect to the plane. The variable 'a' is named Linsley curvature parameter and is used for photon-hadron separation. Photon showers are expected to have larger 'a' than data because of the more curved shower front (Figure 3.5-2, left).

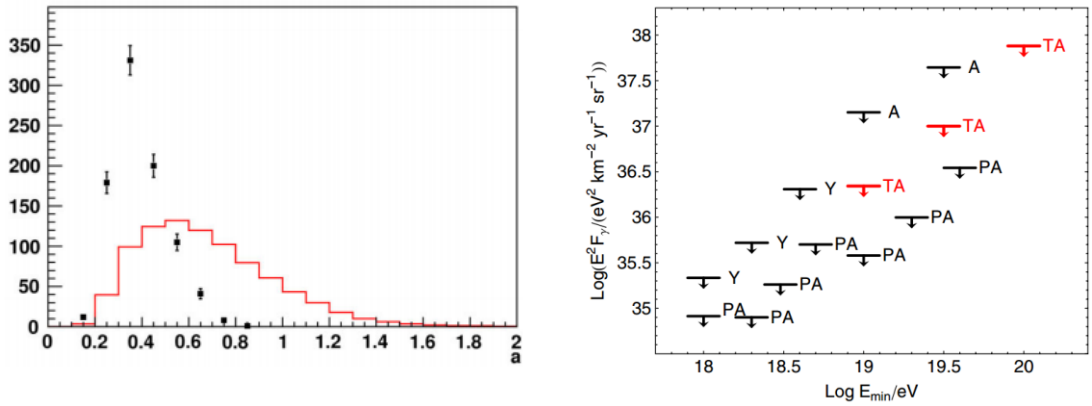


Figure 3.5-2 Photon-hadron separation based on the Linsley curvature parameter (left) and results of upper limit of photon flux from the Telescope Array Collaboration (right). The histogram[106] in red on the left plot shows photon simulations, black points are from data. Events with zenith angle 45 < θ < 60° and E > 10^{19} eV are used.

It was shown that the separation for photon and data at 0 < θ < 45° are too poor to be used for this analysis. The upper limit was obtained by selecting data with 45 < θ < 60°, which decreases amount of exposure that can be used.

TA has presented the lowest upper limit on the photon flux and photon fraction among all observatories in the North hemisphere. An upper limit of 6.2% and 28.5% of photon fraction for UHECR with energy larger than 10^9 and $10^{19.5}$ eV[106] were found respectively. However, compared to results from the Pierre Auger Collaboration (SD analysis introduced in section 3.3), the TA results are several orders of magnitude higher. The fact that TA is not sensitive to muon number and small area of detectors makes it harder to develop a photon-hadron separation parameter from SD with a cheerful Merit factor. The relatively small exposure also makes it impossible to reach the photon limit achieved through the Auger analysis.

3.6 Importance of muon number

It has been stated several times that the number of muons carries important information for photon-hadron separation. In this section, we show quantitatively the role of muon numbers compared to photon-hadron separation based on X_{max}.

Simulations have been generated using the EPOS-LHC model via CONEX programme. Proton, photon and iron showers are simulated with θ = 60° and energy 10^{18} – $10^{19.5}$ eV (E^{-1} spectrum). The reason of choosing such high zenith angle is to have a full grasp of the longitudinal profile (because at 60° the depth of atmosphere the shower goes through is twice as long as it travels at 0°). There are ~ 440,000 proton, 220,000 photon and 220,000 iron simulations. The same library of simulations is also used in Chapter 8 for the estimation of the possible photon background from proton showers. In addition QGSJetII-04 is used for the comparison between hadronic interaction models but is not used in this section. The simulations were produced using the *Pleiades* cluster from Wuppertal.

From Figure 3.6-1 to Figure 3.6-3, X_{max} and the maximum number of muons N_μ^{max} are compared between proton, photon and iron at different energy bins.

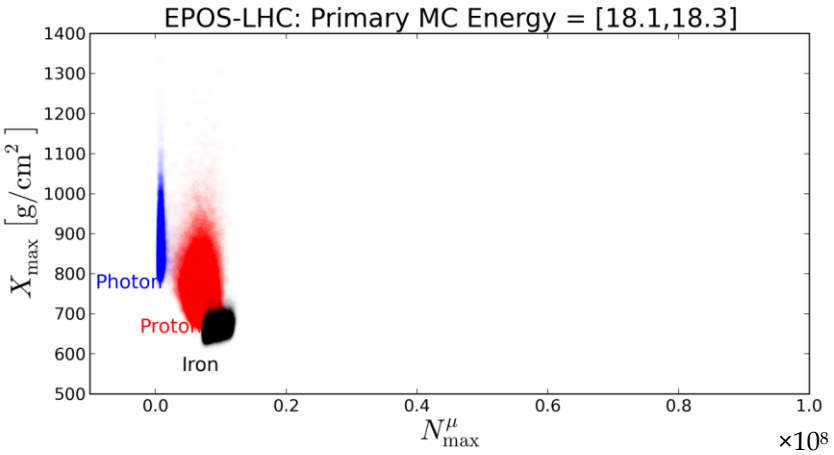


Figure 3.6-1 X_{max} and N_μ^{max} for energy range $10^{18.1} < E_{MC} < 10^{18.3}$ eV using EPOS-LHC model for photon, proton and iron simulations. There is no means of measuring the maximum number of muons using current experiments. It is important to find parameters that can indirectly indicate the number of muons for the mass-composition study.

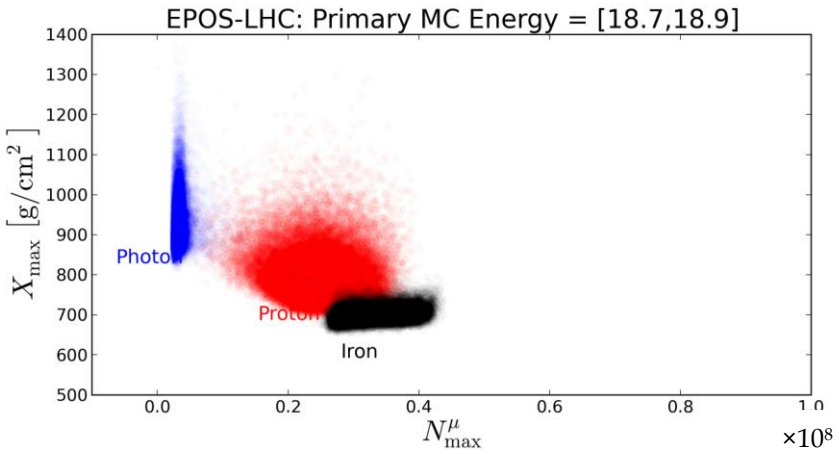


Figure 3.6-2 X_{max} and N_{μ}^{max} for energy range $10^{18.7} < E_{MC} < 10^{18.9}$ eV using EPOS-LHC model for photon, proton and iron simulations

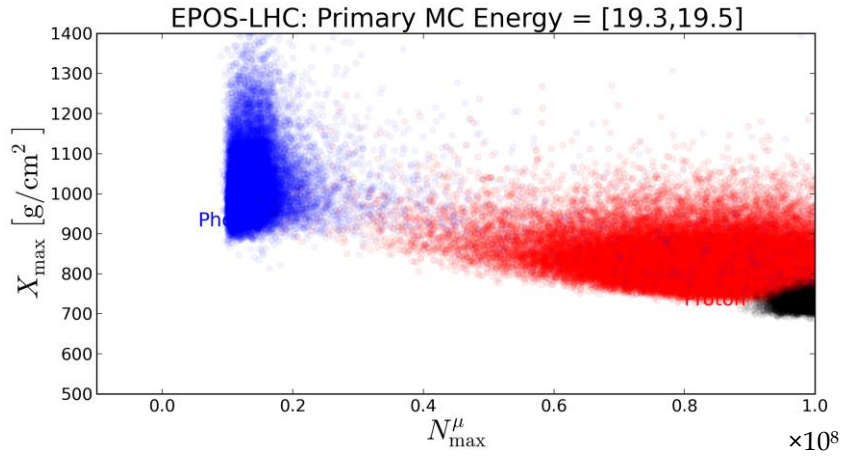


Figure 3.6-3 X_{max} and N_{μ}^{max} for energy range $10^{19.3} < E_{MC} < 10^{19.5}$ eV using EPOS-LHC model for photon, proton and iron simulations

Photon showers carry on average ~ 15% of muons compared to proton simulations. The sources of these muons are discussed in section 6.5.2.1. It is evident that the separation achieved between proton and photon is stronger using muon number than X_{max}. It is not realistic to build an array of muon detectors that covers the area sensitive to UHECR detections based on current technology due to the high cost. Therefore, it is important to find a parameter that sensitive to the muon component of shower.

Through this chapter, photon-searches based on SD and FD data from the Pierre Auger Observatory are presented. Separation powers of various parameters are compared. It was shown that the difference between photon and proton is more evident using muon numbers than X_{max}. In the next chapter, the entity method is proposed as the solution to no direct measurement of pure muon component in the era of experiments.

Chapter 4. Physics behind the entity method

The main aim of the thesis is to explore the possibility of using information from the spread of the arrival times of particles in an air shower for photon-hadron separations by applying a new technique. The method we propose is called the entity method, which is described in section 4.1. To demonstrate the physics behind the entity method, evidence from past experiments and Monte Carlo simulations are discussed showing that mouns arrive earlier at the ground compared to the electromagnetic components of a shower. The dependence of the muon fraction on the shower development is also introduced.

4.1 The birth of the entity method

Water-Cherenkov detectors measure the combined signal from the muonic and electro-magnetic components. Each vertically-going muon contributes 240 MeV to the energy deposit while nearly all the energy of electrons and photons is lost in the water. In Auger the signal is measured in 25 ns time intervals. The different entities ($\mu^{\pm}$, $e^{\pm}$, γ) contribute to the signal and create different fluctuations in each time bin. Muons arrive earlier than the electro-magnetic components in the shower. Evidence from experimental measurements and Monte Carlo calculations are shown in section 4.2 and 4.3 respectively. The entity method is designed to make use of the shape, signal fluctuations and 'muon fraction' of the FADC traces.

The idea of the entity method was first proposed by A. A. Watson: since the muon component is the key for photon-hadron separation, one expects a more slowly growing signal at the beginning of the FADC trace for photons than that of a hadronic primary. To quantify the difference of the trace shape and the signal fraction at the beginning of the trace over the total signal, χ^2/ndf is calculated by comparing time intervals in each measured trace to those of an averaged trace derived from data. The type of the 'entities' can therefore be estimated indirectly. The entity method has been developed further and the procedure for obtaining the entity parameter is as the following:

1). Parameterisations of the mean FADC trace and signal uncertainties of each 25 ns bin for stations located at all possible positions with respect to the geometry of the event.

2). Compare the measured FADC trace to the averaged time trace from data for each station. Calculate the logarithm of the reduced χ^2 value: $\log_{10}\chi^2/\text{ndf}$.

3). Calculate the response of $\log_{10}\chi^2/\text{ndf}$ and parameterise the probability density functions for both photon simulations and data.

4). Calculate the likelihood that a detector with a measured $\log_{10}\chi^2/\text{ndf}$ belongs to data or photon respectively. Calculate the ratio of the two probabilities and take the logarithm.

5). Repeat procedure 2 and 4 for all stations selected in the event. Convert all probability ratios to a logarithmic scale.

6). Average over all results from procedure 5 to get the final parameter which is named as the 'entity likelihood ratio'.

Step 1 only needs to be done once. To describe the mean of the FADC trace of data, the shape (normalised for the selected period of time) and the scaling (signal fraction of the selected time period over the total signal) are parameterised separately. In section 5.1 the goodness of fit of several functions to the shape is tested and discussed. Determining the shape of the FADC trace is similar to measuring the rise time but the entity method contains more precise information. The signal fractions, which are used to scale the normalised shape functions, describe the amount of energy deposited at the beginning of the trace. It is a rough approximation of the muon fraction and therefore assists with the photon-hadron separation (section 4.4 and 5.1).

The parameterisation of the uncertainties in each 25 ns bin is performed using twin stations and is introduced in section 5.2. The probability density functions of photon and data expectations on $\log_{10}\chi^2/\text{ndf}$ are described in section 5.3. The flow diagram to summarise the key procedures (2, 4, 5 and 6) to calculate the entity parameter for an event is shown in Figure 4.1-1.

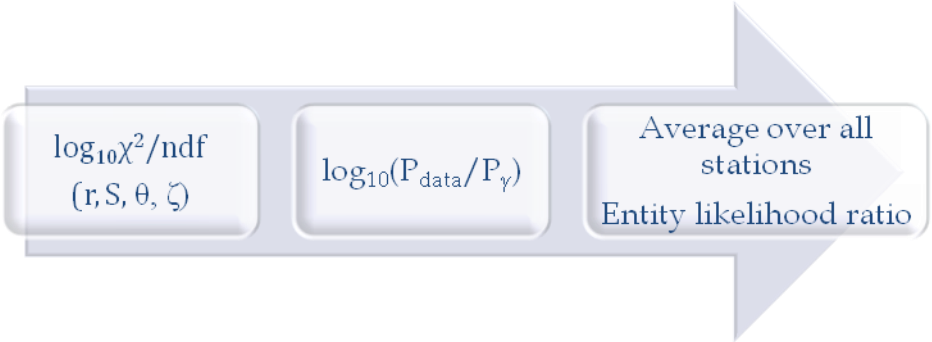


Figure 4.1-1 Flow-diagram of the entity method. The first two steps are on the station-level and the final step is to give an event parameter. An example of how to get the entity likelihood ratio is demonstrated in section 6.1.

The entity method has a strong physics motivation for photon-hadron separation. Moreover, the method does not require a fit to quantities of stations from the event, which means no hard cuts on the number of stations need to be made so that the selection efficiency is higher than most SD variables (i.e., $t_{1/2}(1000)$ and R_c) that are sensitive to photon-hadron separations.

4.2 The time structure of the shower as seen in previous work

The word 'entity' is used to describe the components that contribute to the signal of the FADC trace. There are two components, the muonic component and the electromagnetic component, which consists of electrons, positrons and photons. The first study of the time spread distributions of shower particles was reported by Bassi, Clark and Rossi more than 60 years ago[108]. By shielding one of the three liquid scintillator detectors they were able to separate the muonic component of the shower. They found for cosmic ray showers with energy ~ 10^{15} eV, electrons form a thin disk of thickness between 1 and 2 m at arrival and reach the ground before the muonic component by about 3 m (corresponds to ~ 10 ns). These electrons were probably produced just a few radiation lengths above the detector and generated by leading nucleons from the earlier collisions. This is the pioneering work that illustrated that it is possible to study the shower development by understanding the time structure of the shower front.

The Volcano Ranch array, which was composed of 20 scintillators covering an area of 8 km^2, was sensitive to cosmic rays with much higher energies (> 10^{17} eV). Linsley and Scarsi published their results[105] and showed that beyond 260 m muons arrive earlier than the electromagnetic components and that the spread of the arrival time increases with core distance. Most of the electromagnetic component comes from X_{max} and then scatter around before reaching the ground while muons travel in straight lines. At 950 m the median delay of the electrons is ~ 530 ns. The direct measurement of the arrival time is shown in Figure 4.2-1. Their work offers great inspiration even today when the shower Monte Carlo simulations are widely used.

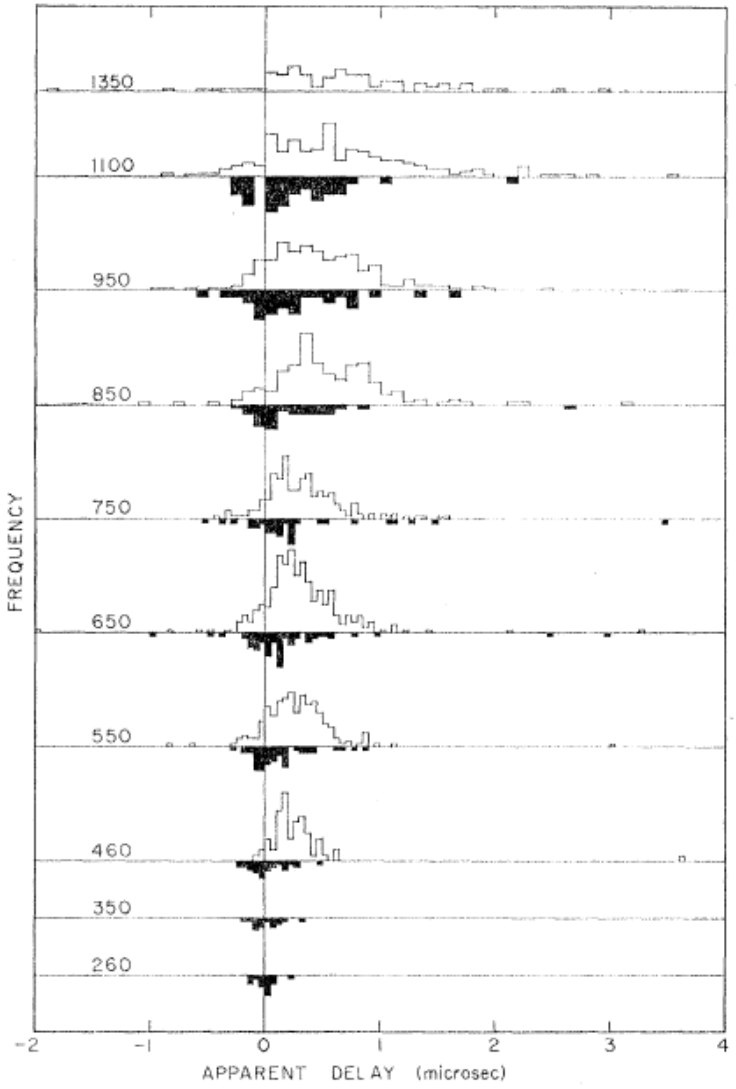


Figure 4.2-1 Single-particle delay histogram[105] for scintillators at various core distances. Distances are shown on the left: the shaded histogram corresponds to the shielded detectors (with 10 cm of lead) that measure muon signals. The unshielded detectors measure the combination of electric and muonic components of the shower and are shown as the upper histograms. The area of each detector is 3.26 m^2 which makes it impossible to study the phenomenon on an event-by-event basis.

In the next section, the differences of arrival times of two components for proton and photon showers are shown using Monte Carlo simulations.

4.3 The 'entities' of the FADC traces from proton and photon simulations

Monte Carlo simulations give unique access to visualising the details of component traces for different primaries of the cosmic rays. The averaged component traces were calculated for stations at various core distances ranging from 500 m to 1200 m. Figure 4.3-1 shows the showers with zenith

angle θ ~ 22˚ and Figure 4.3-2 shows θ ~ 58˚. The EM component includes photons, electrons and positrons.

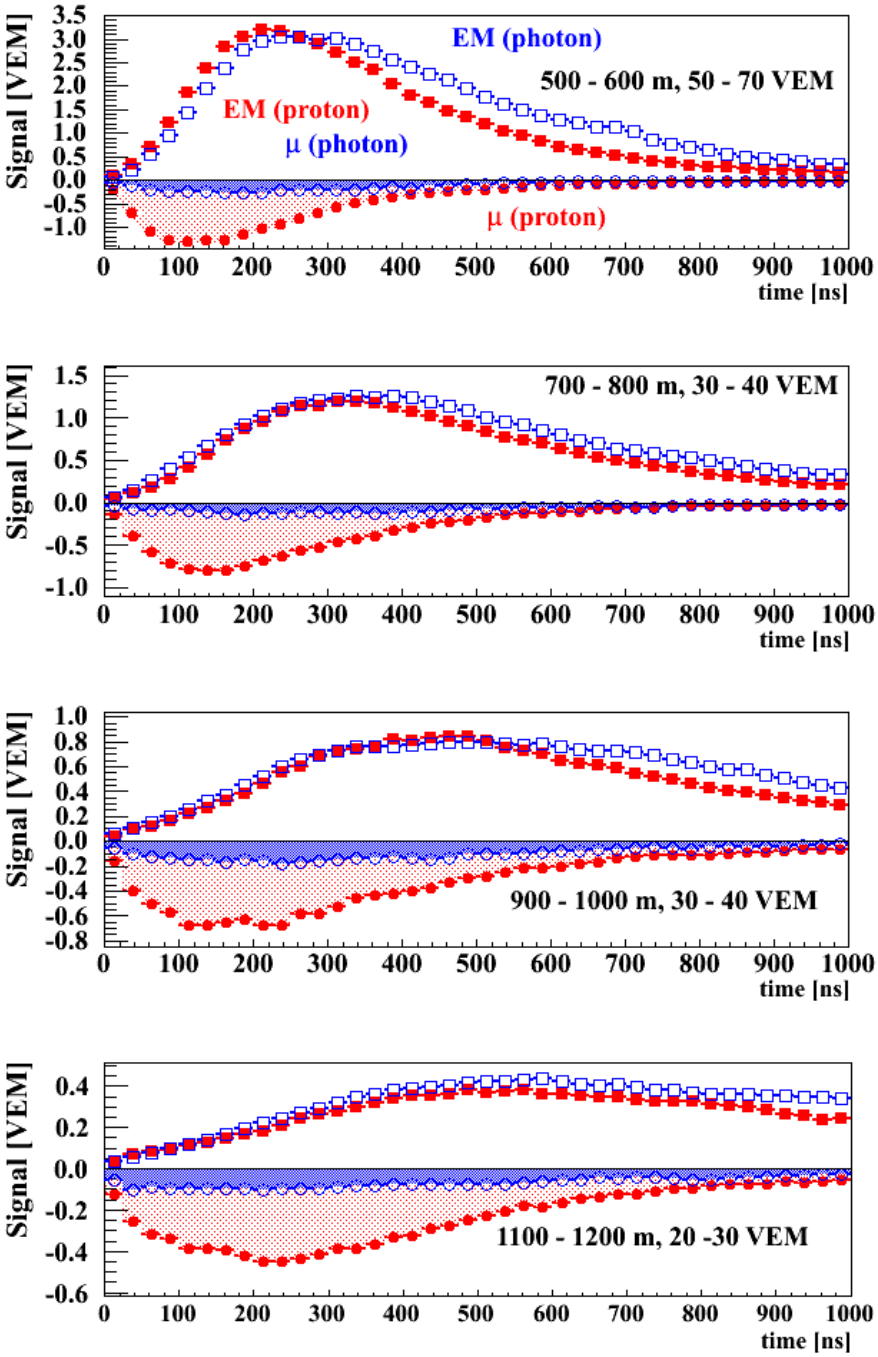


Figure 4.3-1 Component traces of the proton and photon simulations (vertical, θ ~ 22˚). Red solid points are from proton simulations and blue open-circles are from photon simulations. The EM component is shown in the positive axis in squares and the muon component is shown in the negative axis in circles. The later arrival of the EM component is consistent with the experimental results shown in Figure 4.2-1.

It is evident that the muon component arrives earlier than the EM component. The difference in the EM traces is relatively small compared to the differences in the muon traces.

At more inclined angles, the signal contributed by the muon component grows as showers go through more air. For instance at θ = 60° the depth is ~ twice the depth of air comparing to if it goes through the air vertically.

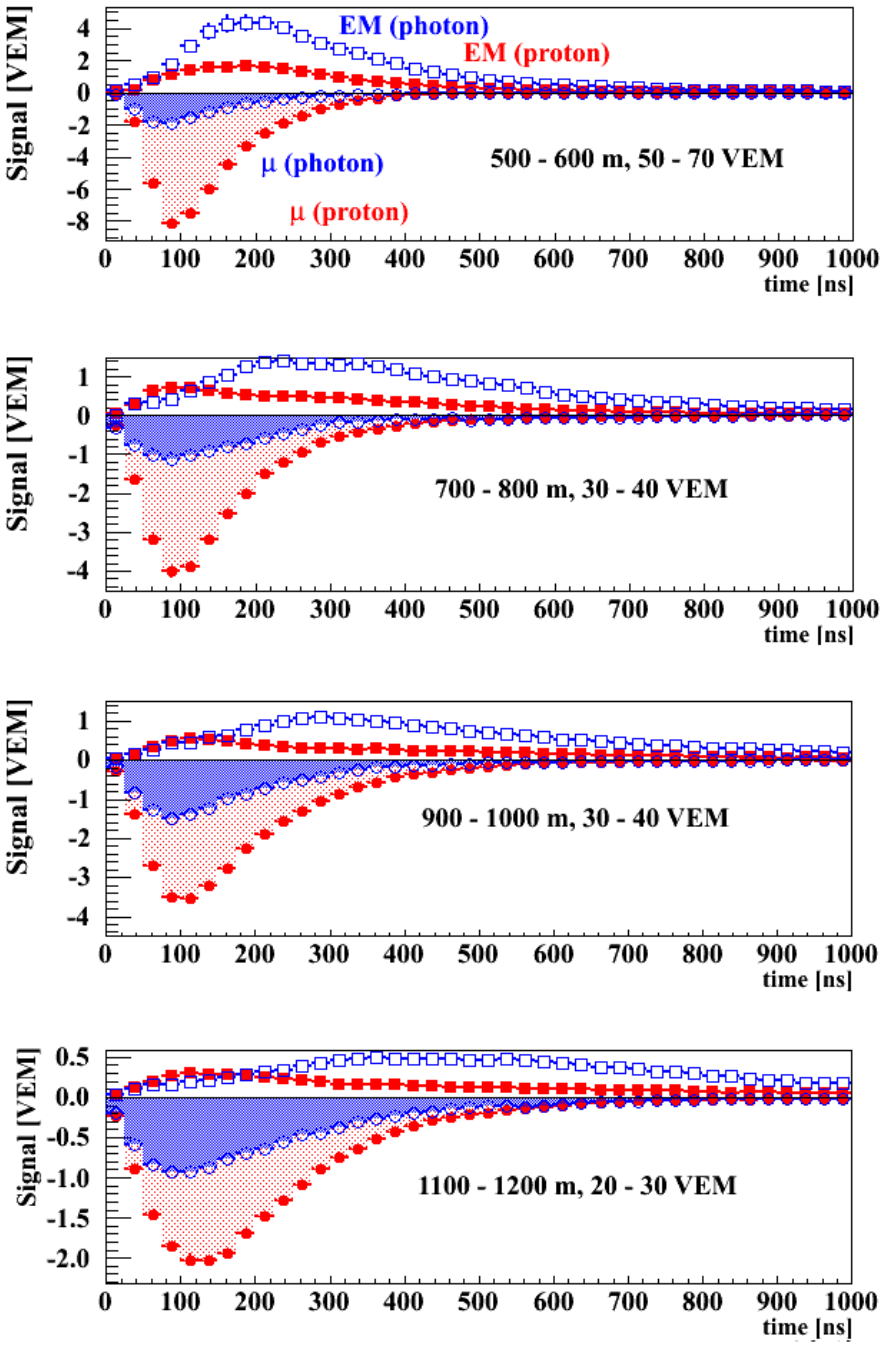


Figure 4.3-2 Component traces of the proton and photon simulations (inclined, θ ~ 58°). Same notations are used as in Figure 4.3-1. The time traces are shorter and the contribution from the muon component is more significant compared to vertical showers.

All the above cases show that the content of the 'entities', the shower components, are intrinsically different for proton and photon. The difference is most evident at the beginning of the time trace. Experimentally the

combined FADC trace – as measured in the water–Cherenkov detectors – provides access to quantify this difference. The results of simulations are shown in Figure 4.3-3 where blue points are from photon showers and black are from proton showers. Detectors selected are with r ~ 675 m and S ~ 35 VEM.

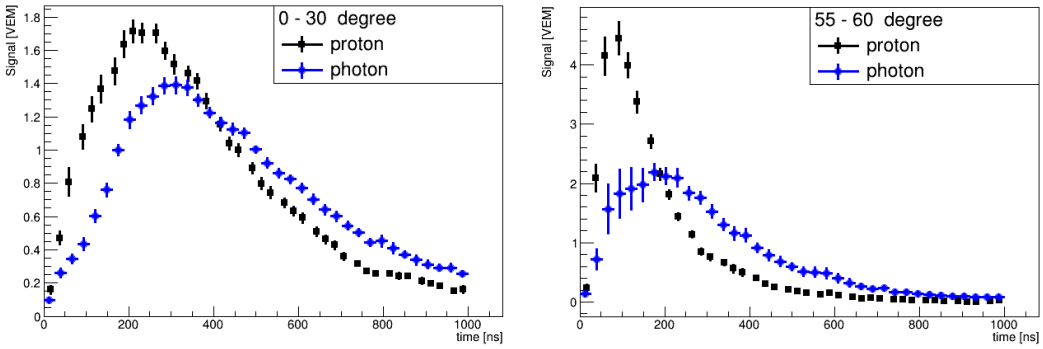


Figure 4.3-3 Simulated FADC trace for proton and photon showers with vertical (left) and inclined (right) zenith angles

To distinguish the differences between the two curves, one could use parameters such as the position of the peak, the width of the trace, the slope of the increase (risetime) / drop (falltime) and the signal fraction deposited at a certain range of time period over the total signal. The entity method tries to combine four of the mentioned elements, the peak position, the slope for the signal rise, the signal fraction and the width of the trace. The idea is that for each detector with core distance r, total signal S and zenith angle θ, and the polar angle ζ, the trace obtained from the mean of many events with identified r, S, θ and ζ can be compared to a trace from a selected shower. By calculating χ^2/ndf, the combined information of the 'likeness' of the trace can be determined by the comparison to the parameterised mean traces.

The signal fraction μ/EM (the ratio of the signal contributed by muons to the signal contributed by EM component for an entire FADC trace) is indicative of the development of the shower. It changes in a complex way in the multidimensional space of core distance r, total signal in the station S, zenith angle of the shower θ and the polar angle of the detector ζ (introduced in the next section). The muon fraction is closely related to the shape of the time trace and therefore understanding its behaviour will help understand the optimised configuration of the entity method for the photon-hadron separation.

4.4 Muon fraction and its dependence on shower development

The core distance is defined as the distance of the detector from the shower axis, which means that for two stations triggered by an inclined shower, the same core distance could correspond to a different depth of development of the shower. The part of the shower that hits a 'late' detector has travelled through more atmosphere than the part of the shower that triggered an 'early' detector. As shown in Figure 4.4-1, the shower arrives at the station on the right at time t while the station on the left is triggered later at time t + δ t. Therefore the station in the right is the 'early station' and the one on the left is the 'late station'.

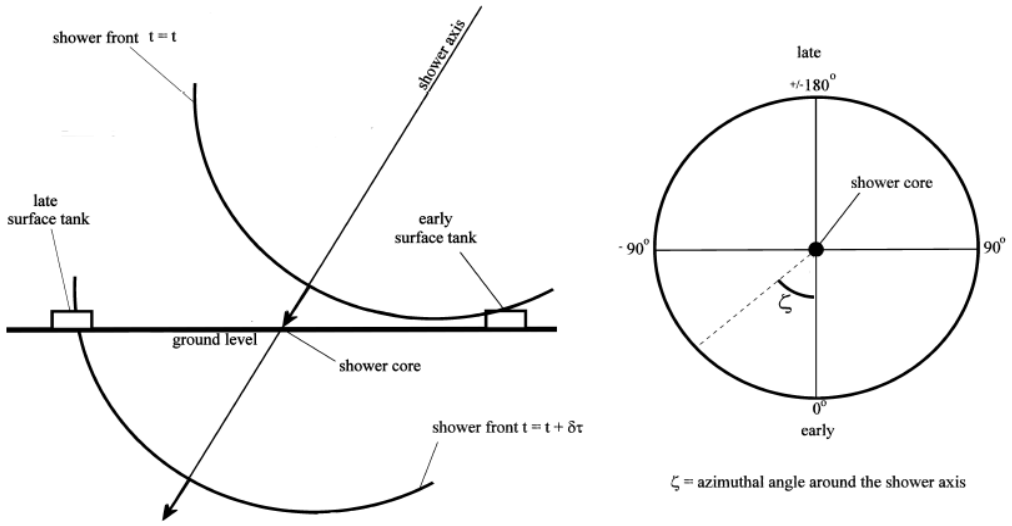


Figure 4.4-1 Definition of the early / late station and the polar angle (ζ) of the stations[97]. Two stations are with the same core distance and the one triggered the first is marked as the 'early' surface detector, which corresponds to a younger 'age' of the air shower. The figure on the right is the projection of the shower to the plane that is perpendicular to the shower axis.

The polar angle ζ is defined in the plane perpendicular to the shower axis as shown on the right of Figure 4.4-1. The core is in the centre and all stations distributed covering 360° isotropically. The station that was triggered earliest has ζ = 0° and the station that was triggered the latest has ζ = 180°.

Since the shower front has gone through more atmosphere in the case of the late station, the signal deposited in the detector is more dominated by the muon component than for the early stations. This is shown in Figure 4.4-2 for both proton and photon simulations at two zenith angles. The selected stations are at a core distance 1000 m and the reconstructed hadronic energies are larger than 10^{19} eV.

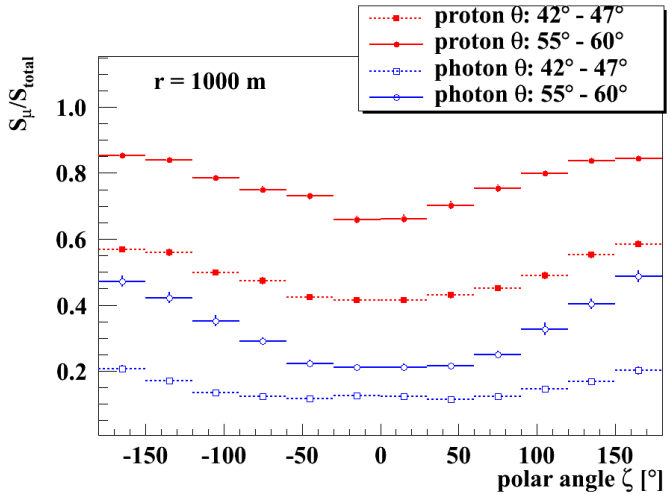


Figure 4.4-2 Dependence of the muon fraction on the polar angle for p and γ simulation. Muon fraction is the ratio of the muon signal over the total signal of the FADC trace. The blue are photon and the red are proton. The asymmetry of the muon fraction indicates that there are more EM components contained in the early stations.

It is obvious from the shower geometry that the more inclined the shower is, the more atmosphere the shower has to travel through. The muon signal fraction should increase with respect to the zenith angle (more precisely sec θ). This dependence is shown in Figure 4.4-3.

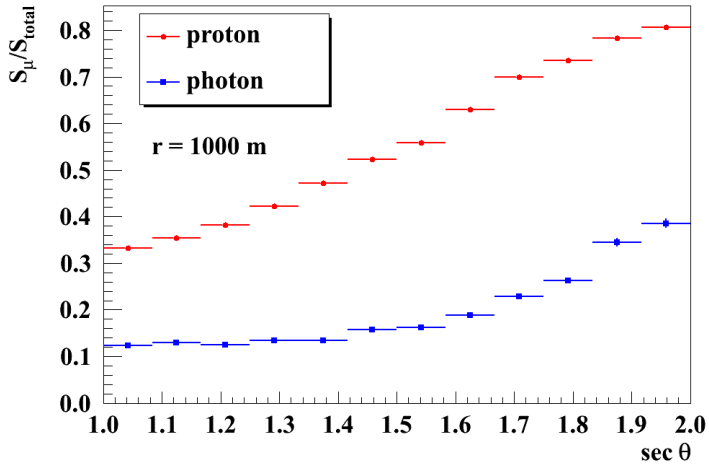


Figure 4.4-3 Dependence of the muon fraction on the sec θ for MC protons and photons. Stations are located at 1000 m and the reconstructed energy of the shower are larger than 10^{19} eV.

The core distance is another important element that the muon fraction depends on. For a station triggered by a vertical shower, the further away it is from the core position, the longer distance the particles travelled to reach the station from the shower axis. The lateral distribution function describes the change of the signal in the detector with respect to the core distance

(section 3.2.3). With the help of simulations, it is possible to study the lateral distribution functions for different components in the shower.

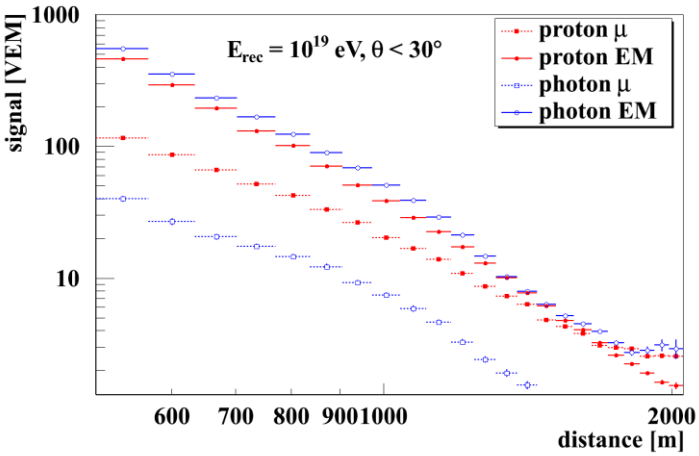


Figure 4.4-4 Lateral distribution functions of μ and EM components of p and γ showers. The showers selected are with reconstructed energy (on the hadronic scale) ~ 10^{19} eV and at vertical angles. Red are proton and blue are photon. The slopes of the lateral distribution of muons are flatter compared to the EM component for both proton and photons.

The muon signal fraction as a function of distance is shown in Figure 4.4-5.

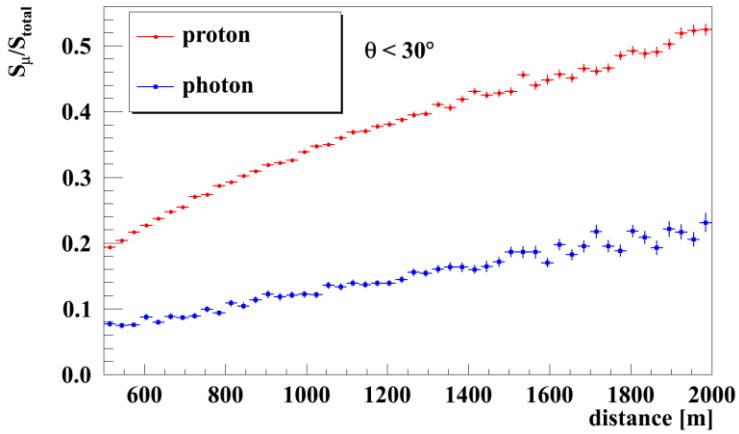


Figure 4.4-5 Dependence of the muon fraction on the distance for p and γ simulations. Stations selected are with total signal ~ 35 VEM and events with $E_{rec} > 10^{19}$ eV.

The difference of the muon fraction increases with respect to the core distance. It seems for the photon–hadron separation it makes more sense to go to larger distances. However as the distance increases, the total signal in the station drops and gradually the small number of particles makes the measurements suffer larger uncertainties due to the sampling fluctuations.

The last dependence we discuss in this work is the dependence on the total signal of the station. For illustration vertical showers are selected and stations at core distance of 1000 m are used for the calculation. The signal at 1000 m is normally known as S(1000), which is used for the energy calibration. In this sense one could relate the signal dependence to the energy dependence. The result is shown in Figure 4.4-6.

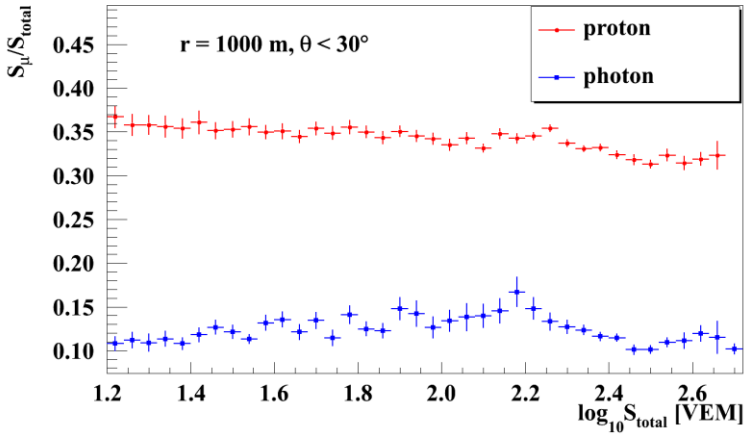


Figure 4.4-6 Dependence of the muon fraction on the total signal of p and γ simulations. The signal dependence is correlated to the energy dependence. Stations are located at 1000 m. Only vertical showers are selected.

It appears the dependence of the muon ratio on the signal is the smallest one among the four mentioned dependences. For proton simulations the muon fraction decreases as the signal increases. This is probably because a larger signal corresponds to a larger energy, therefore a deeper X_{max} has a higher contamination of the EM component.

As introduced in this chapter, it is vital to find a parameter that is sensitive to muon numbers for photon-hadron separation. The water-Cherenkov detectors respond to the combination of both electromagnetic and muonic components of the shower. Without shielding the detector, one could try to de-convolve the FADC traces and then count peaks to estimate the muon number or use multivariate analysis to constrain muon numbers with respect to simulations[109]. Both indirect approaches depend on assumptions from hadronic simulations. The entity method is designed to tackle the difference of the muon component of hadronic and photon showers without relying on hadronic simulations by setting the parameterisations based on data. Details are described in the next chapter.

Chapter 5. Parameterisations of the entity method

As discussed in section 4.1, to compare a measured FADC trace to a benchmark trace and calculate the χ^2/ndf of the station, the mean time-traces (section 5.1) and signal fluctuations in each 25 ns time bin (section 5.2) must be parameterised using data. The entity likelihood parameter has been proposed and will be introduced in section 5.3. The data reconstruction was carried out using offline, which is a modular framework designed by the Pierre Auger Collaboration.

5.1 Parameterisations of mean time-traces from data

As demonstrated in section 4.4, the muon fraction of stations have dependences on core distance r, total signal of the station S, zenith angle of the event θ and the polar angle of the station ζ. Therefore it is reasonable to parameterise the mean time-traces according to these four quantities. The θ range is divided into 6 equal bins in sec θ. Core distances are grouped in equal bins of 50 m from 350 m. Signals are binned in logarithmic intervals from 15 VEM with bin width 0.1 (bin number = lg (signal – 15 VEM) / 0.1). The reason to start with 15 VEM is to reduce fluctuations in particle number. Polar angles are divided into 12 equal bins. Figure 5.1-1 lists the binning of the parameter space used in obtaining the mean FADC traces from data.

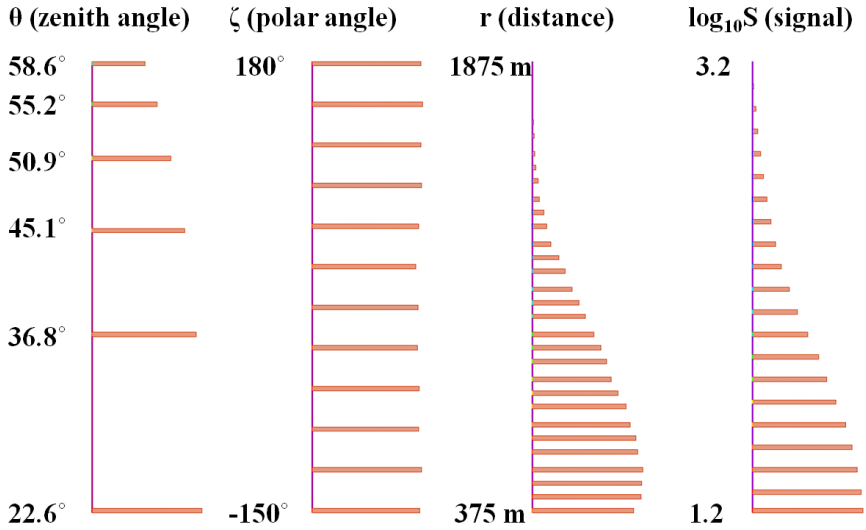


Figure 5.1-1 Binning of zenith angle θ, polar angle ζ, distance r and signal S from data. For fitting mean time-traces, zenith angles are binned in equal sec θ bins, polar angles are with bin width of 30˚, core distance has bin width of 50 m and signals are binned in logarithmic intervals. The length of the horizontal bar indicates the number of entries.

Data used in the parameterisation are from January 2004 to August 2012 using offline ('Shannon' v2r7p8) reconstructions. We have performed tests showing the mean trace parameterisations do not have obvious dependences on the version of offline. No low-gain saturated stations are allowed in the station selection and only bins with at least 5 entries are kept. In total there are more than 10,000 bins in the multidimensional space in r – S – θ – ζ.

A number of functions were tried to fit the shapes of the mean traces of data, including the log-normal function, the Moyal function and the partly-normalised Landau function (the standard Landau function with signal under a period of time normalised to 1). Examples of the trace fitting for data at vertical angle (θ ~ 23°) and inclined angle (θ ~ 59°) are shown in Figure 5.1-2 and Figure 5.1-3 respectively. The mean traces are from data and have been normalised to areas equal to 1 between the time-period of 0 – 600 ns. It can be seen that in both cases the partly-normalised Landau function is a better fit. At low zenith angles, the Moyal function gives reasonable fits while the log-normal function fits the data poorly. However at more inclined angle, when signals rise faster, the Moyal function cannot be used to describe data but the log-normal function fits relatively well.

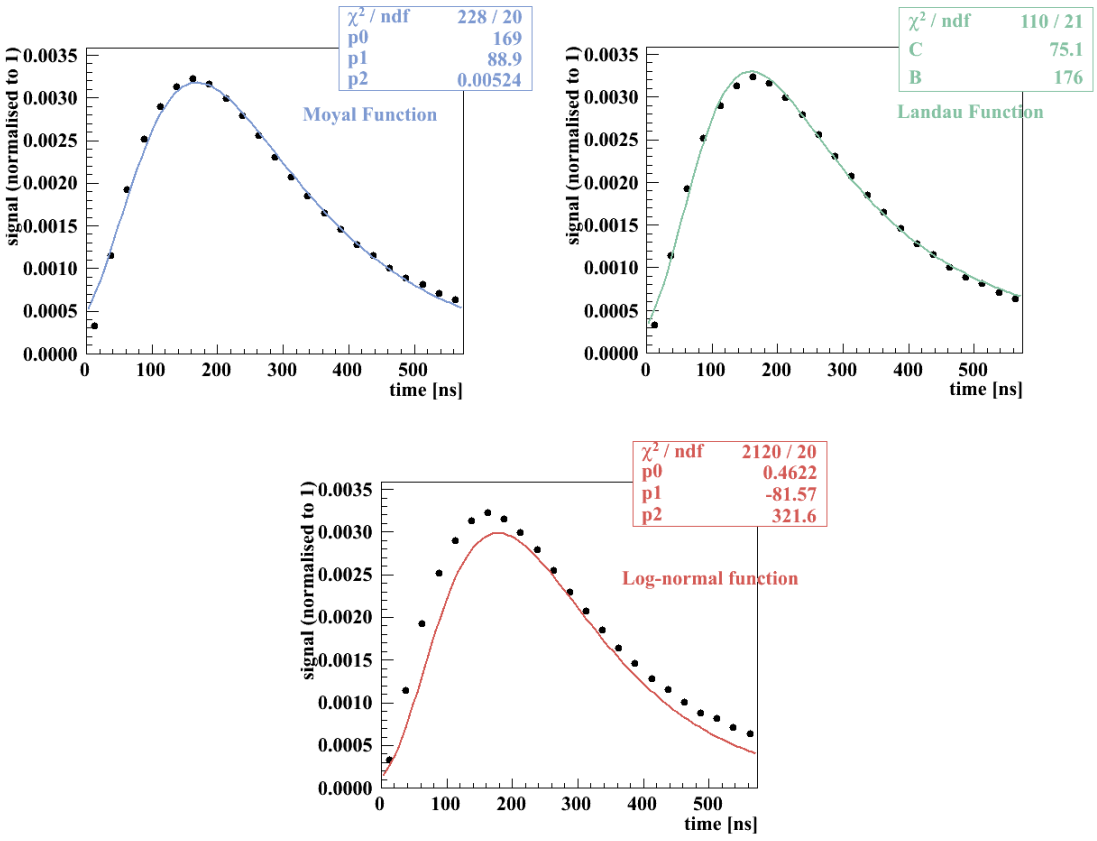


Figure 5.1-2 Fitting mean traces with the Moyal, log-normal and partly-normalised Landau function for vertical showers (θ ~ 23°)

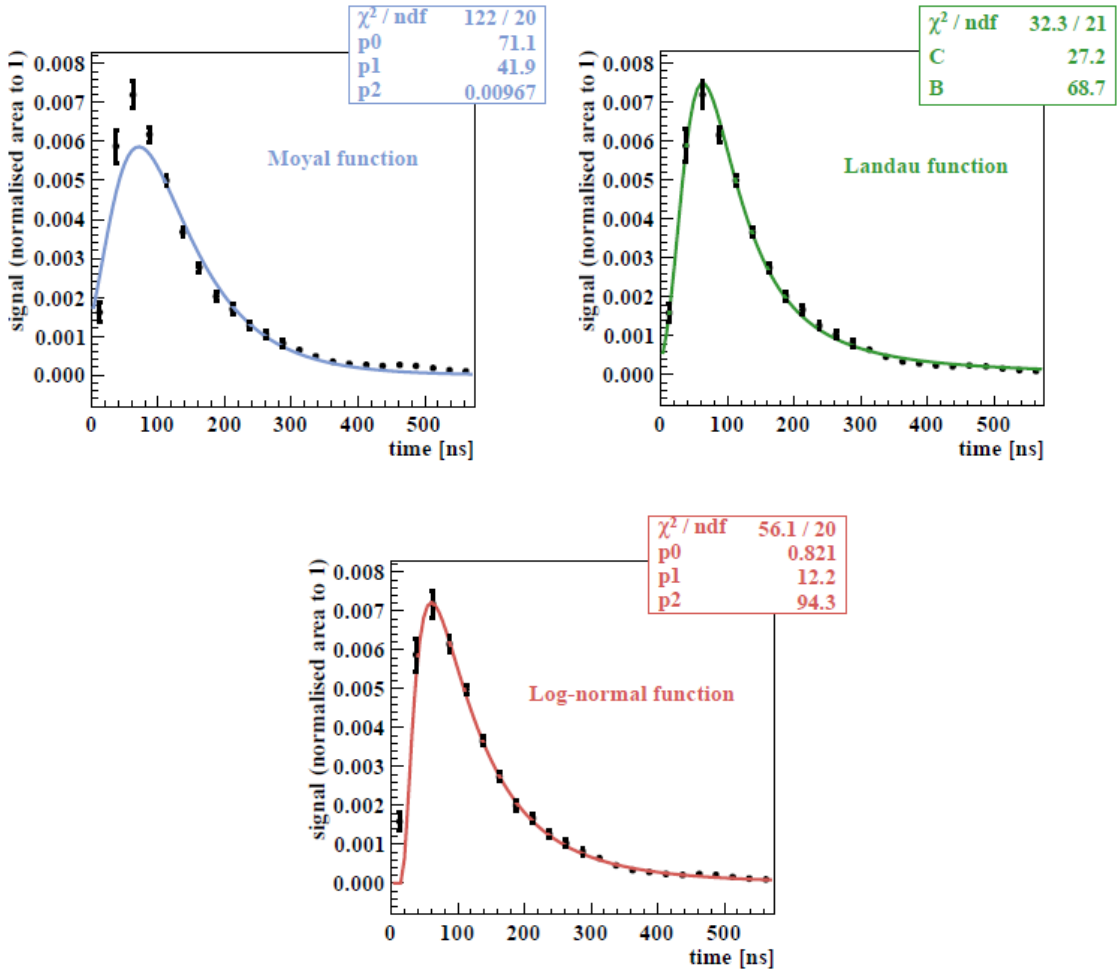


Figure 5.1-3 Fitting mean traces with the Moyal, log-normal and partly normalised Landau function for inclined showers (θ ~ 59°)

The Landau function gives smaller χ^2/ndf values (Figure 5.1-6) and is the best choice for fitting the shape of the mean time-traces. It has a complex form and contains an integral from – i∞ to + i∞. The probability density function is as

$$p(x)=\frac{1}{C}\Phi(\lambda) \qquad 47,$$

$$\lambda=\frac{x-(B+C\times 0.22278298)}{C} \qquad 48,$$

$$\Phi(\lambda)=\frac{1}{2\pi i}\int_{C-i\infty}^{C+i\infty} e^{\lambda s+s\log s}ds \qquad 49,$$

where B is the most probable value (MPV) and C is ¼ of the full width half maximum of the Landau function. Figure 5.1-4 shows the change of the Landau function with varying B and C. The area under the Landau curve has been normalised to 1 from -100 ns to 300 ns.

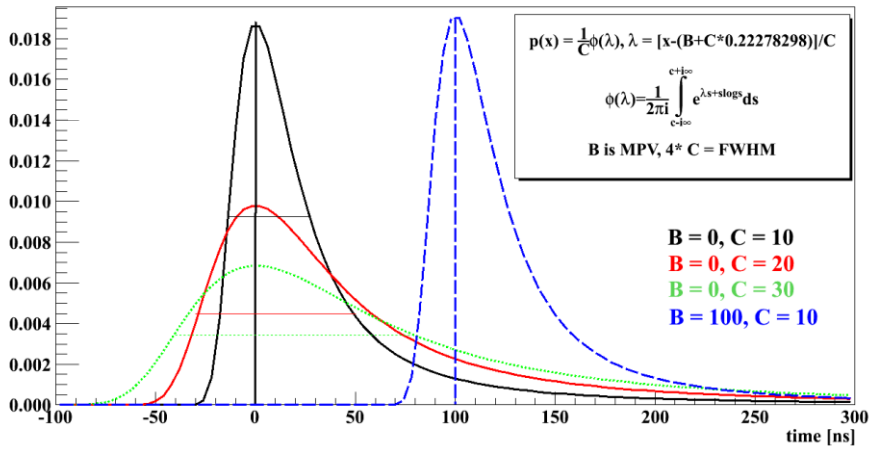


Figure 5.1-4 Parameters in the partly-normalised Landau function used for trace-fitting. B and C as free parameter: B is the maximum probable value and C is the ¼ of the full width half maximum.

The Landau function is 'partly-normalised' by requiring that the function has the same area as the mean trace within a selected time period. The selection of the time period has to be longer than 200 ns because of differences of FADC traces between proton and photon (Figure 4.3-3). However the time range has to be short enough to provide reasonable fits for all 10,000 bins of the mean traces. An example to illustrate the improvement of the fitting result by using a shorter range of time-period is shown in Figure 5.1-5. The mean trace is from data with vertical zenith angle, small core distance and high signal, where the dominant signal is from electromagnetic component.

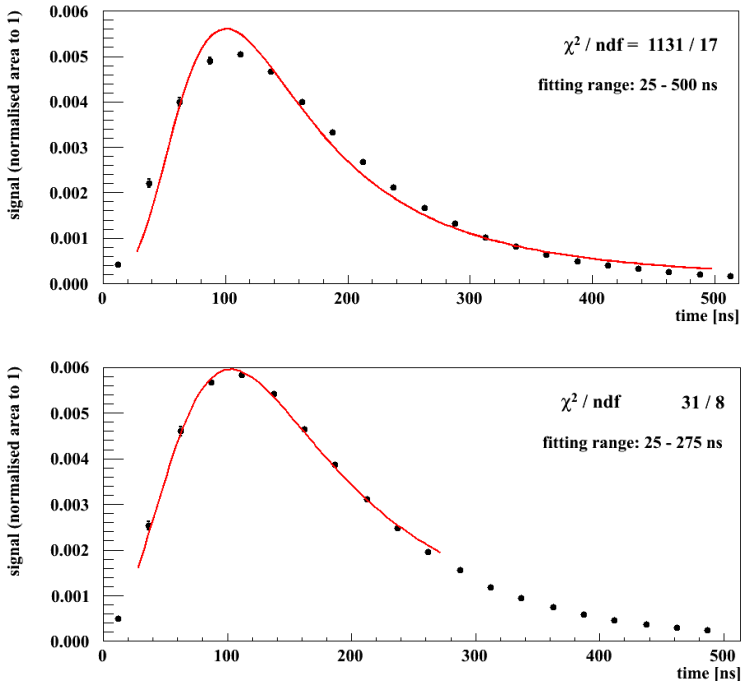


Figure 5.1-5 Comparing fitting results when selecting different time range using partly-normalised Landau functions (θ ~ 23°)

It was decided to set the time-period for trace-fitting as 25 to 275 ns. To compare the fitting quality, distributions of $\log_{10}\chi^2$/ndf for all r – S – θ – ζ bins of data are shown in Figure 5.1-6 (Moyal function and Landau function for fitting range 25 – 500 ns). Both Moyal and Landau functions have a second peak, which is due to detectors with small core distances and large signals (such as shown in Figure 5.1-5). It is obvious that the Landau function describes the data better than the Moyal function.

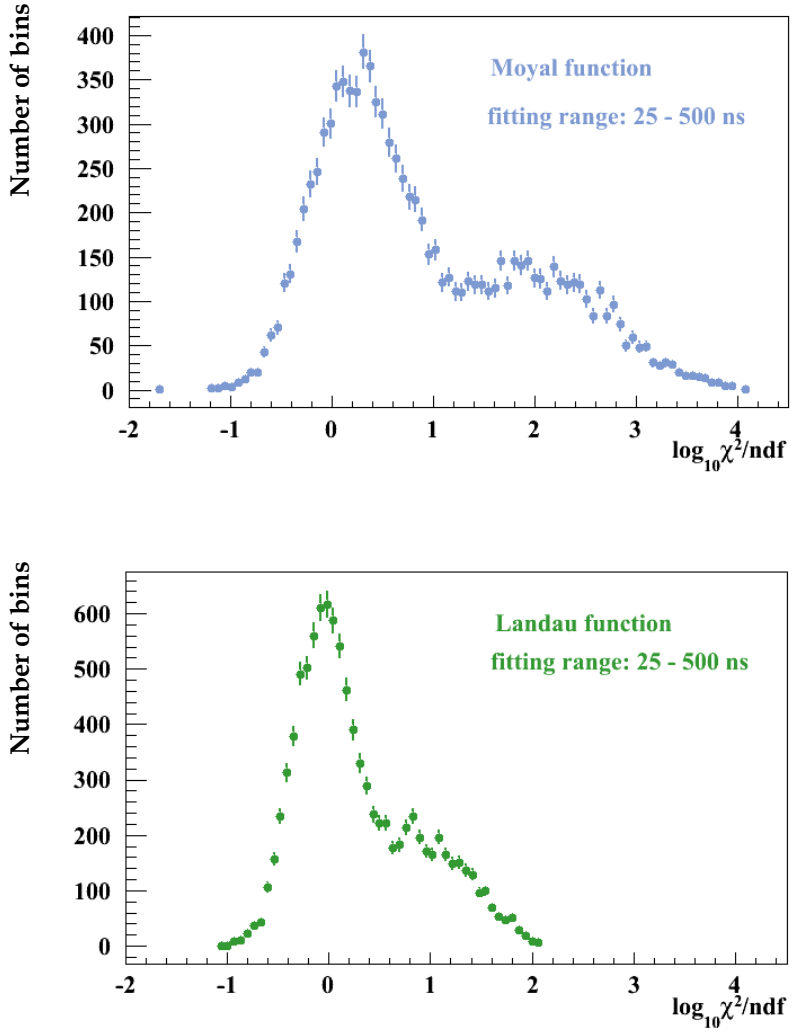


Figure 5.1-6 Distribution of $\log_{10}\chi^2$/ndf when fitting with Moyal or Landau functions within time period of 25 – 500 ns. The Y-axis is the number of bins in the r – S – θ – ζ space. The second peak is due to detectors with small distances and large signals

The second peak of $\log_{10}\chi^2$/ndf disappears when using shorter period of time for fittings due to improvement of the fitting results. Shown in Figure 5.1-7, the distribution of $\log_{10}\chi^2$/ndf from fitting traces with partly-normalised Landau function within the time-period of 25 – 275 ns can be fitted reliably with a Gaussian function.

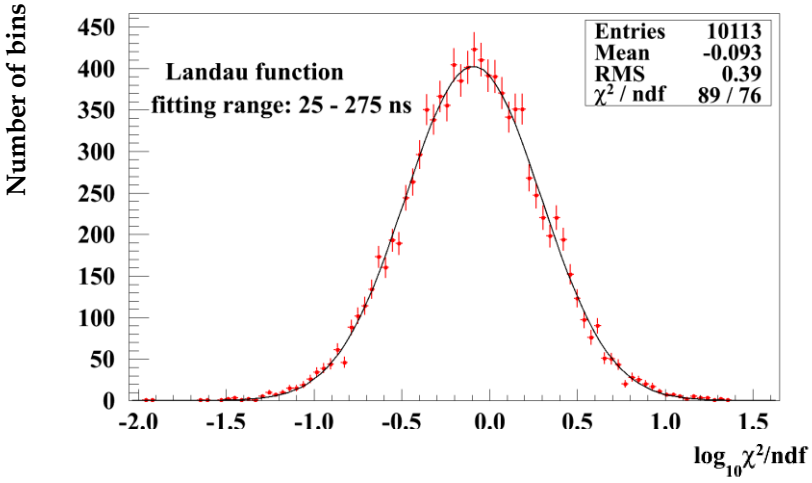


Figure 5.1-7 The $\log_{10}\chi^2$/ndf distribution for all r – S – θ – ζ bins using a partly normalised Landau function within the time range 25 – 275 ns

The dependence of the peak position B and the width-sensitive parameter C (as introduced in Figure 5.1-4) are then parameterised with respect to r, S, θ and ζ. To understand the dependencies, mean traces are compared to each other when only one of the variables is changing while the other three are fixed. All mean traces have an area of 1 for the time period 25 to 275 ns. This dependence is mostly sensitive to the shape of the trace, i.e. the position of the peak, the slope of the rise and the width of the spread.

For detectors with same S, θ, and ζ, a smaller core distance is associated with less delay of the electromagnetic component from the shower. Therefore traces have shorter spreads and an earlier peak value (Figure 5.1-8).

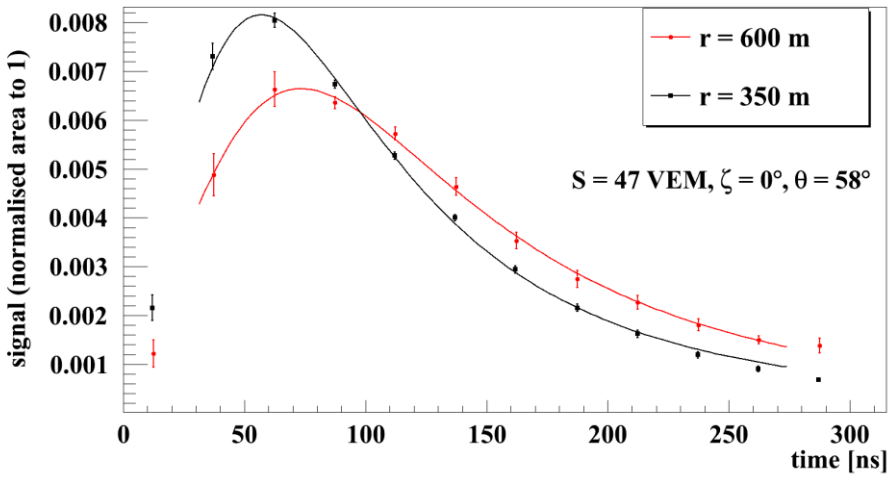


Figure 5.1-8 Mean time-trace (normalised) with varying core distance

For detectors with the same r, θ and ζ, it is likely that stations with larger signals are from events with larger X_{max}. As shown before the muon fraction

decreases slowly with respect to the rising signal (Figure 4.4-6), which means the station with the lower signal is likely to have a faster trace. An example is shown in Figure 5.1-9. However, the signal dependence on the shape of the trace is the least apparent one compared to the other three dependencies.

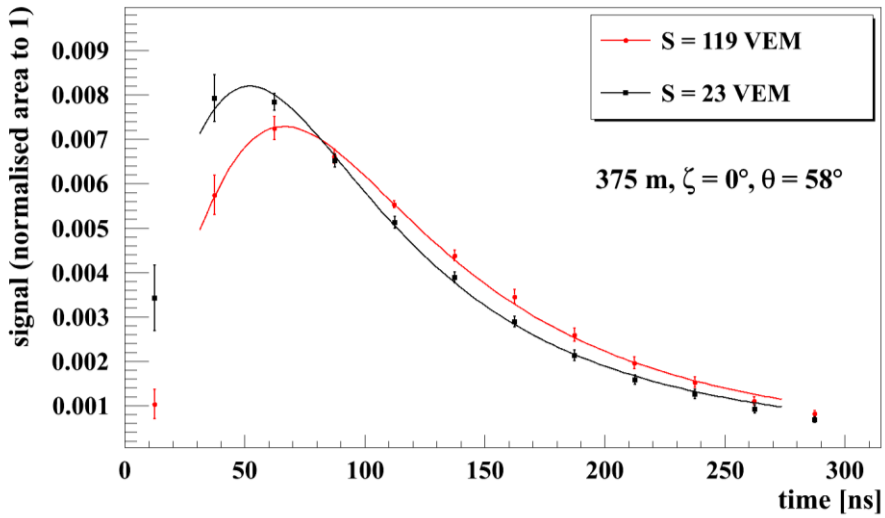


Figure 5.1-9 Mean time-trace (normalised) with varying total signal of the station

Traces from stations with the same r, S and ζ but different θ are shown in Figure 5.1-10. A shower has to travel through more atmosphere if it is more inclined and consequently contains a larger muon fraction. Thus the traces from inclined showers are narrower and faster with early peaks.

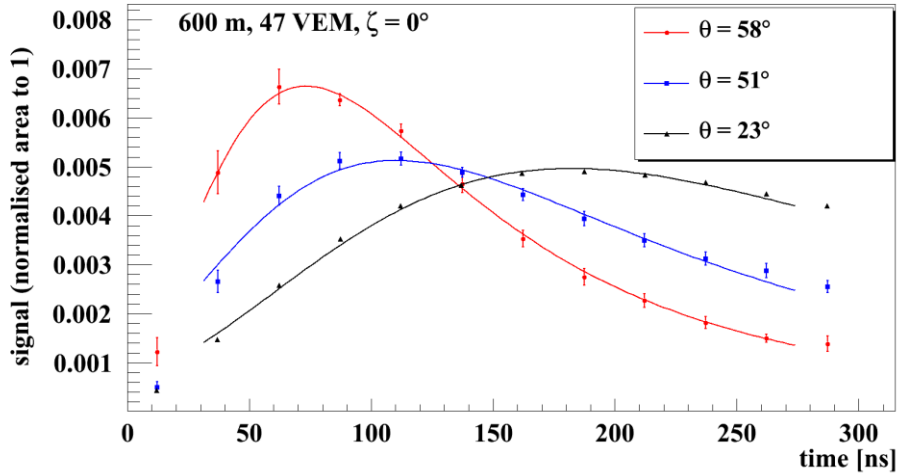


Figure 5.1-10 Mean time-trace (normalised) with varying zenith angle

The last dependence is on the polar angle (Figure 5.1-11). As introduced in Figure 4.4-1, stations with ζ = 0° are early stations and observe particles

closer to X_{max}. As the combination of smaller muon fraction and less delayed electromagnetic component, the early stations have a larger spread of the trace.

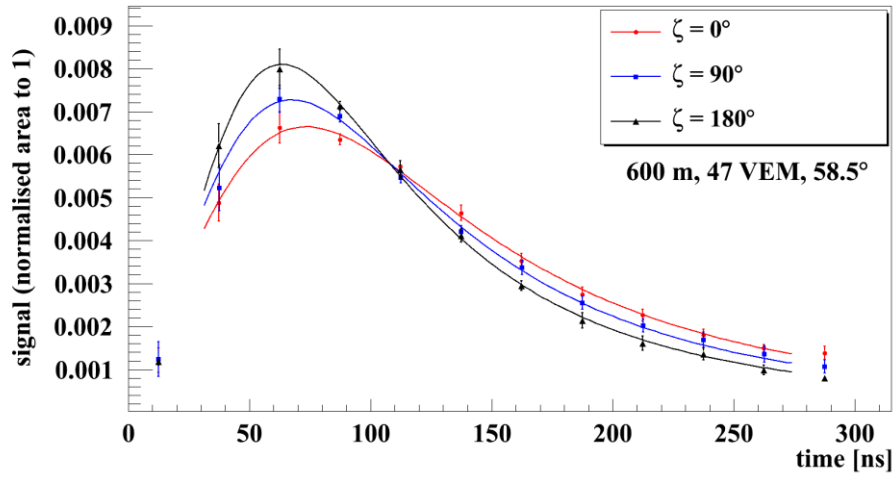


Figure 5.1-11 Mean time-trace (normalised) with varying polar angle ζ

The fitted values of B and C are then parameterised with respect to sec θ using second order polynomial functions. The next 6 plots (Figure 5.1-12 to Figure 5.1-17) are the examples to show some of the fitting results and the dependencies on r, S and ζ.

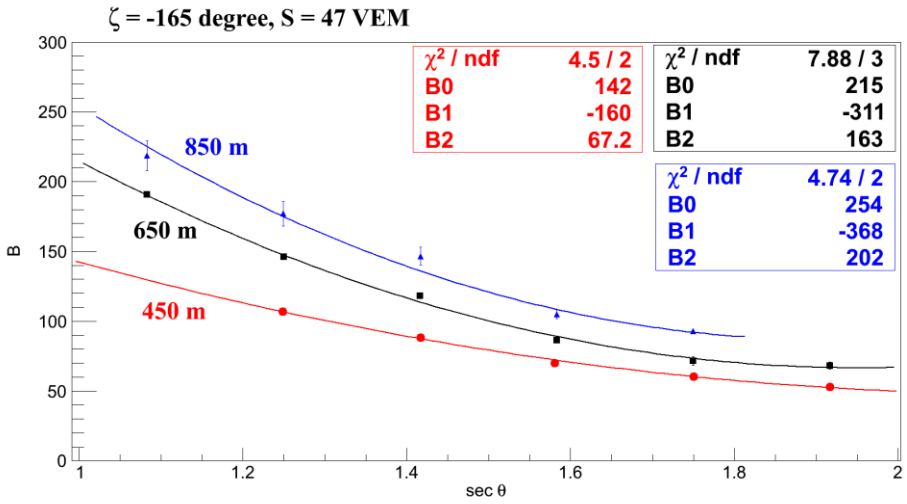


Figure 5.1-12 Peak parameter B and sec θ for various distances

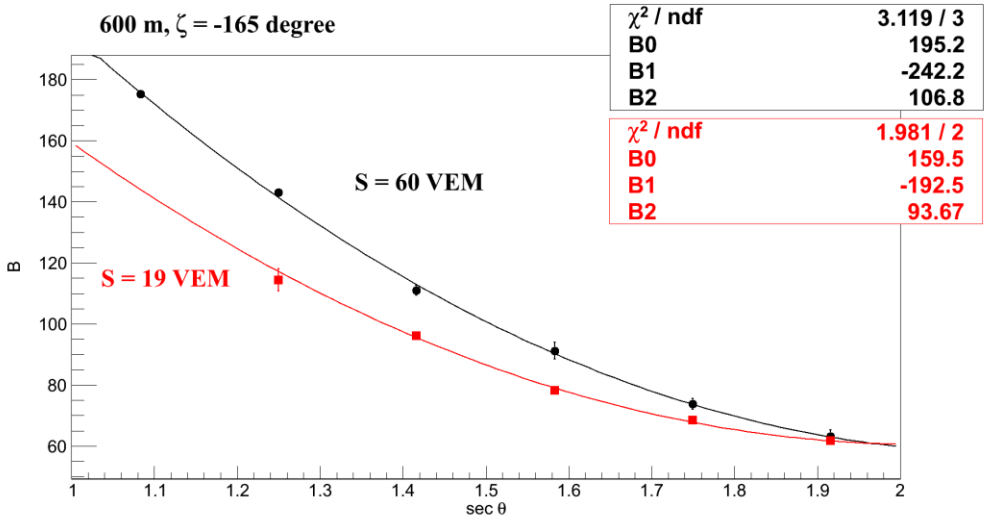


Figure 5.1-13 Peak parameter B and sec θ for various signals

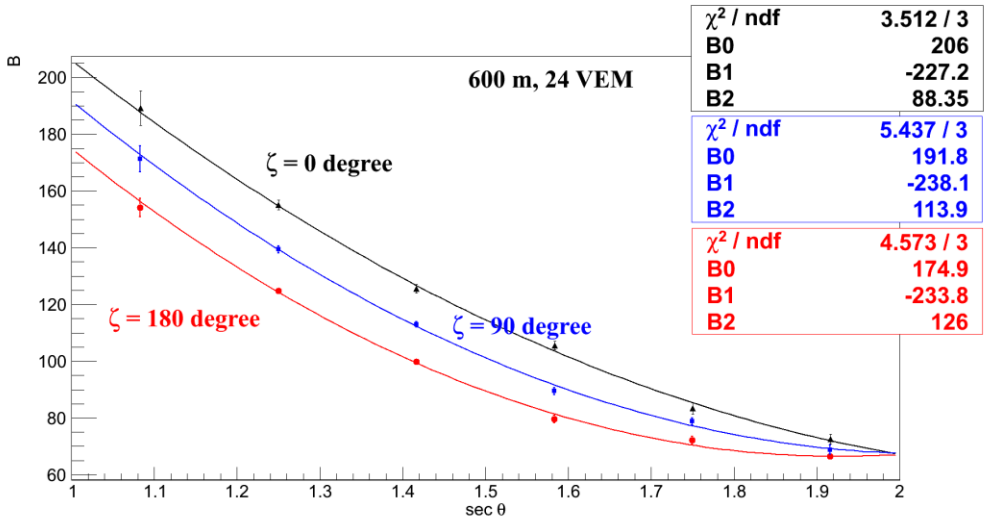


Figure 5.1-14 Peak parameter B and sec θ for various polar angles

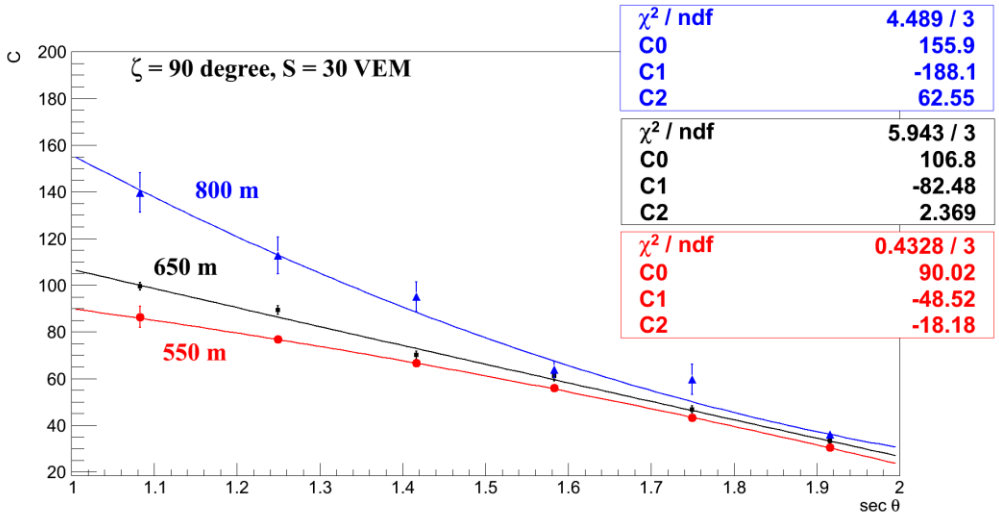


Figure 5.1-15 Width parameter C and sec θ for various distances

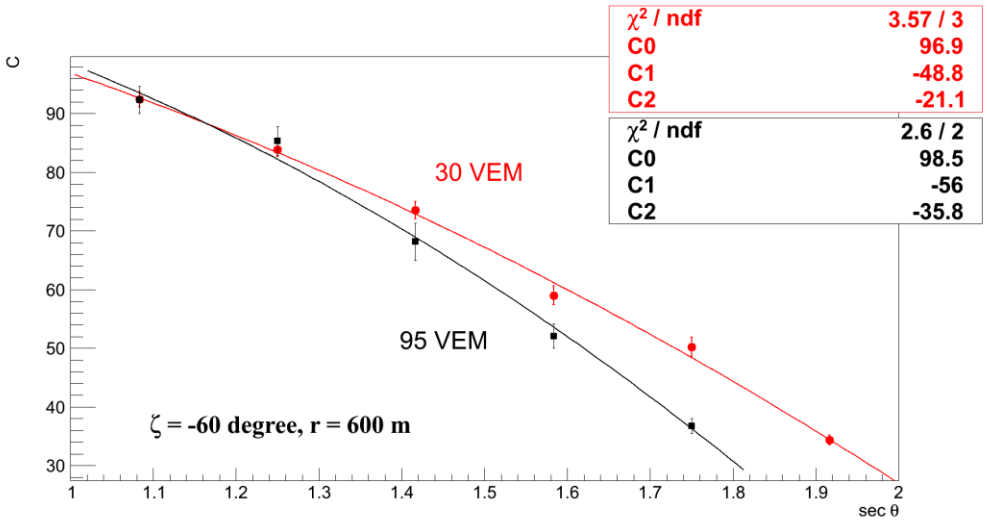


Figure 5.1-16 Width parameter C and sec θ for various signal sizes

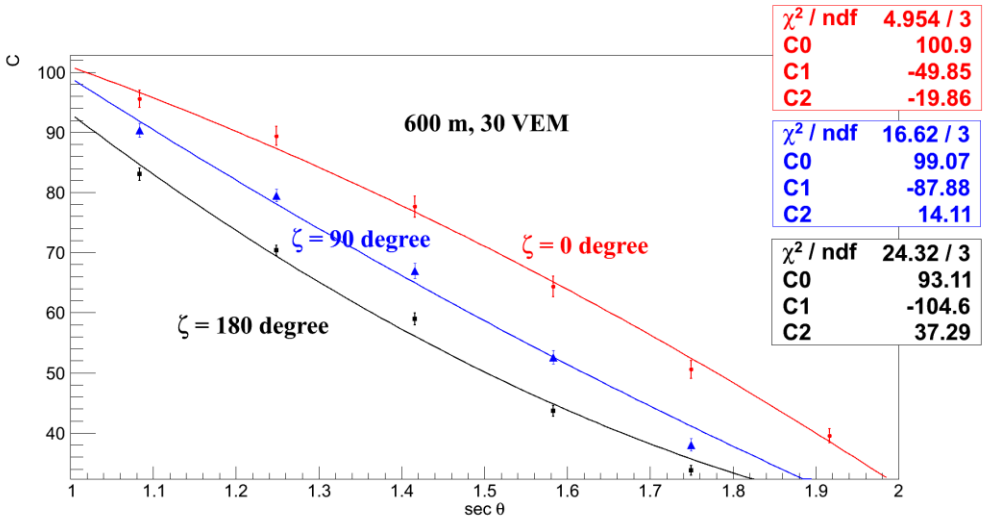


Figure 5.1-17 Width parameter C and sec θ for various polar angles

The next step is to parameterise the scaling of the trace, so that the normalised trace can be scaled to the realistic size. Signal fraction A, which is defined as the signal between 25 to 275 ns over the total signal in the station, is used to do the scaling. The binning of r – S – θ – ζ follows the same rules as shown in Figure 5.1-1. The aim of introducing A is to enhance the separation between photons and protons since the start of the trace is dominated by the muon signal and the parameter A is therefore correlated to the muon fraction. The dependences of A on the four parameters have similar behaviours with the muon fraction, which is introduced in section 4.4. The next three plots (Figure 5.1-18 to Figure 5.1-20) show an example of the parameterisation between A and sec θ with the dependences on r, S and ζ.

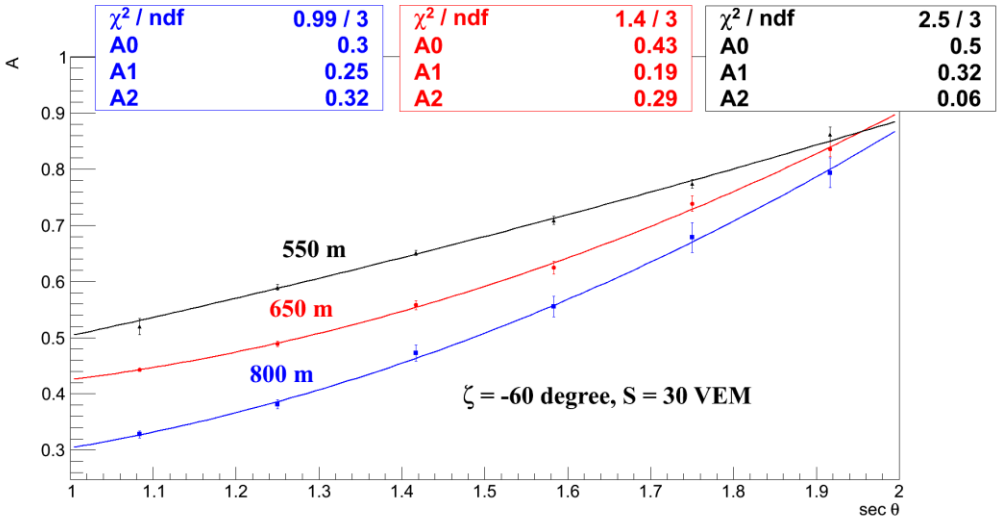


Figure 5.1-18 Signal fraction parameter A and sec θ for various distances

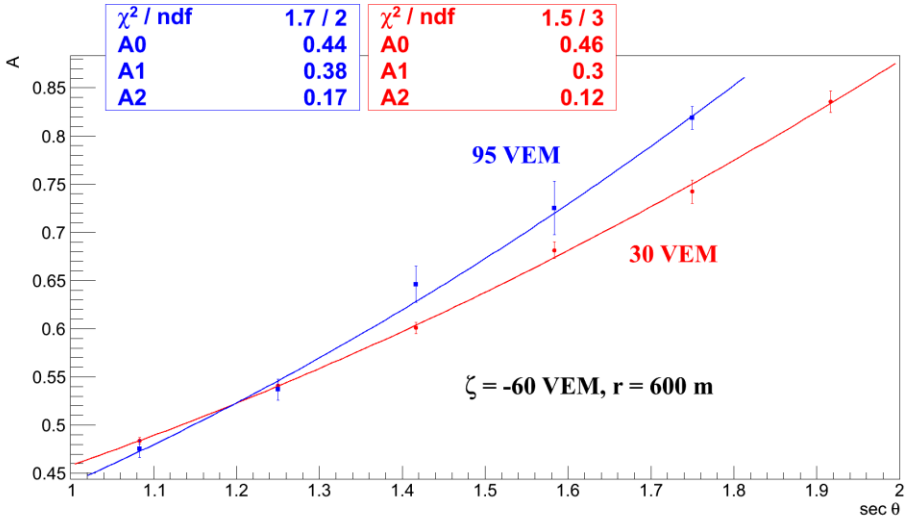


Figure 5.1-19 Signal fraction parameter A and sec θ for various signal heights

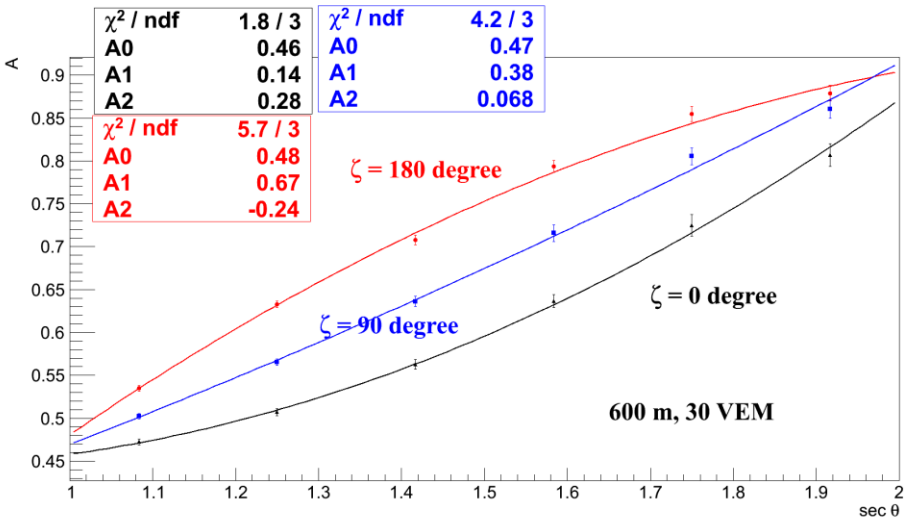


Figure 5.1-20 Signal fraction parameter A and sec θ for various polar angles

The relationship between A, B and C for all bins in the r – S – θ – ζ space is summarised in Figure 5.1-21. As expected B and C are positively-correlated while a larger signal fraction (A) is more likely to come from a trace with smaller peak value B.

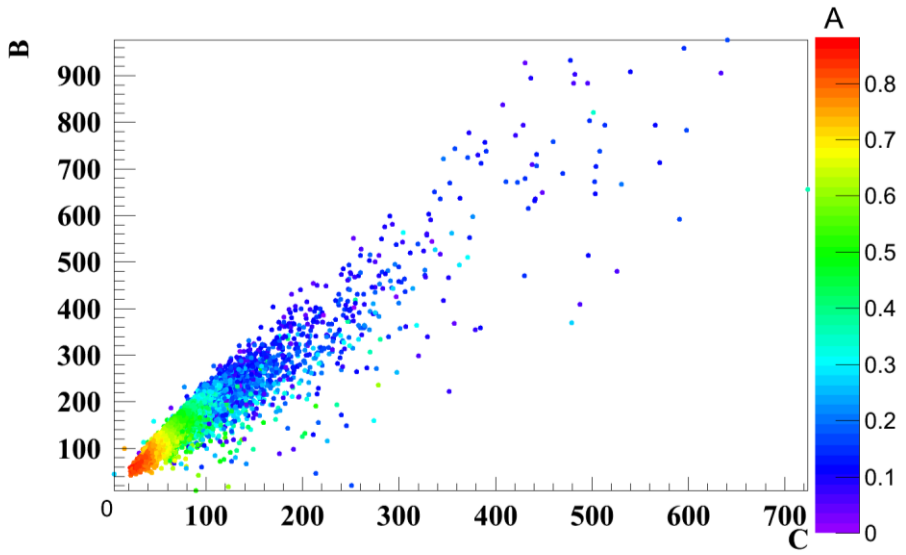


Figure 5.1-21 Correlations between A, B and C from the signal fraction and shape fit. All fittings based on data. Fittings of B and C are independent of fittings of A.

The fitted results A, B and C are parameterised as functions of r, θ, ζ and S. In total 61 parameters are used for the description of the mean trace and stored at http://www.ast.leeds.ac.uk/~phyll/Thesis/Trace_Para.txt (Auger password 2014). Some of the plots can be found in the appendix A.

5.2 Parameterisations of the signal variance for each 25 ns from the FADC trace

Twin stations that are ~ 11 m away from each other are used to calculate the signal variances of each 25 ns interval of the data. There are only a few pairs of twin stations in the observatory, which limits the available statistics. Events with zenith angle 0 - 60˚ and triggering the 1500 m-array are selected. The minimum total signal averaged over two stations is set as 15 VEM and the minimum averaged core distance is 350 m. In total there are 20,177 pairs of twins selected. The differences in core distances and polar angles are shown in Figure 5.2-1.

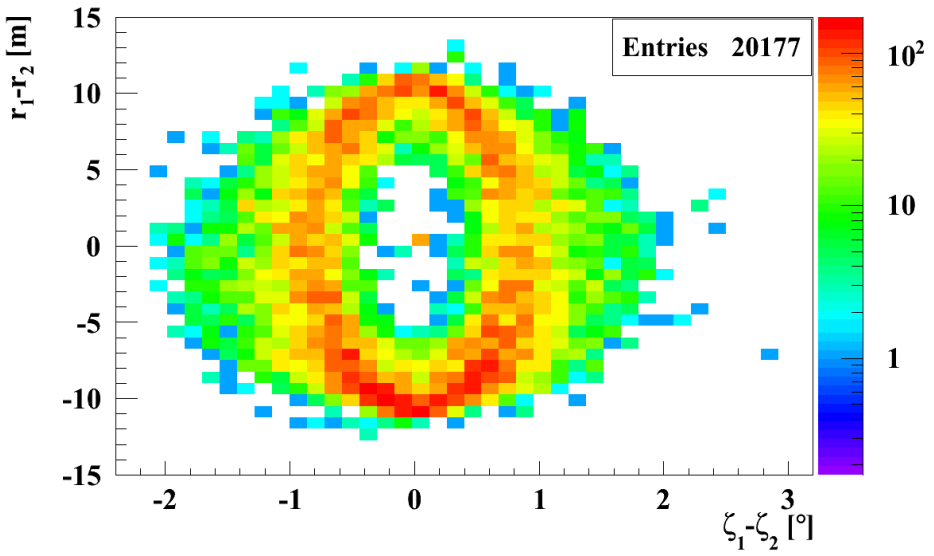


Figure 5.2-1 Difference in core distance and polar angle of selected twin stations

It is important to have similar polar angles for calculating the variance at inclined zenith angles (similar shower depth), which is the advantage of using twin stations as opposed to using the pair stations[98]. The idea of pair stations is to select detectors that have similar distance from the core position in each event, which was firstly introduced during the Haverah Park experiment. The early study of the uncertainties of risetime in Auger also used pair stations due to not enough events of twins[97].

The distribution of averaged core distances and signals of twin stations are shown in Figure 5.2-2.

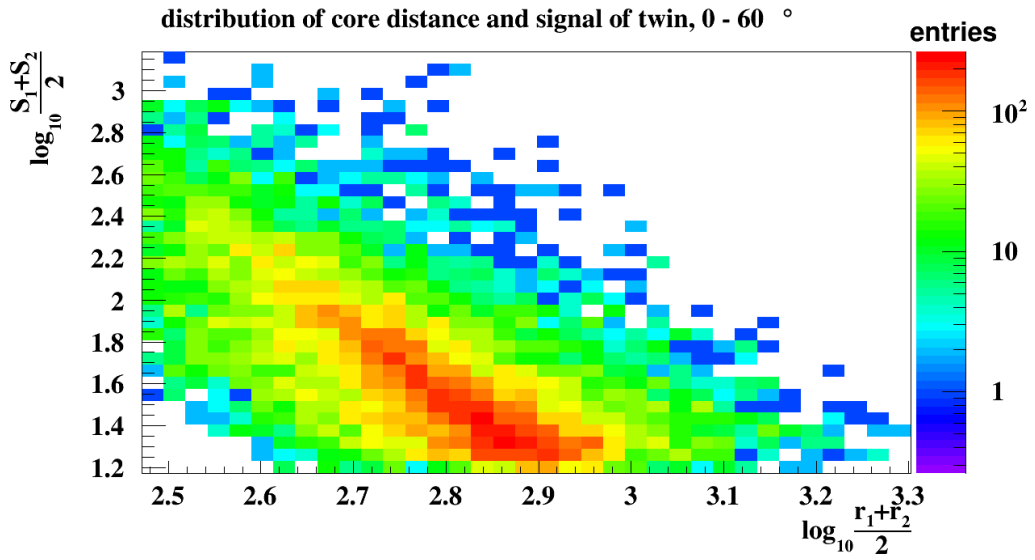


Figure 5.2-2 Distribution of averaged core distance and signal of selected twin stations

The distribution of sec θ is shown in Figure 5.2-3. Most events are with θ < 30˚ (sec θ < 1.5).

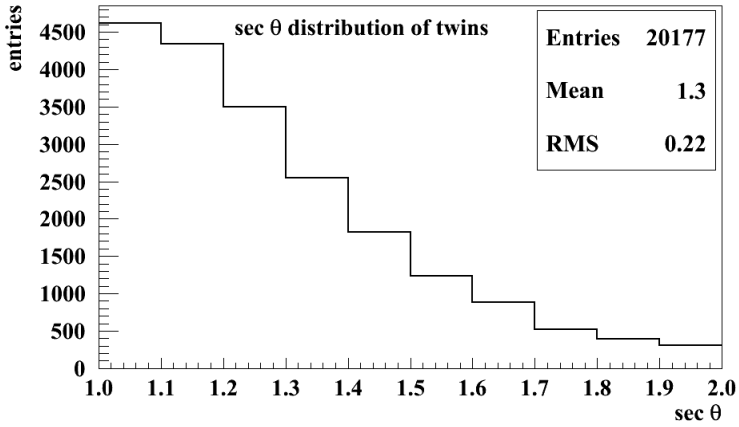


Figure 5.2-3 Distribution of sec θ of selected twins (01/01/2004 – 01/31/2014)

We define the relative fluctuation of the signal height h within a time bin t as L. If h_1^t (from the detector 1 of the twin) and h_2^t (from detector 2 of the twin) are the measured signals from the two selected detectors within the bin t, the unbiased relative fluctuation becomes

$$L = \frac{\sigma(h)}{h} = \frac{\frac{\sqrt{\pi}}{2}<\left|h_1^t - h_2^t\right|>}{<(h_1^t + h_2^t)/2>} \qquad 50.$$

L in each 25 ns bin changes with respect to time. This relationship can be described using a Moyal-like function and an example of fitting is shown in Figure 5.2-4.

$$L = \frac{\sigma(h)}{h} = \exp\left\{0.5 \times \left[\left(\frac{t-p_0}{p_1}\right) + \exp\left(-\frac{t-p_0}{p_1}\right)\right]\right\} \times p_2 \qquad 51.$$

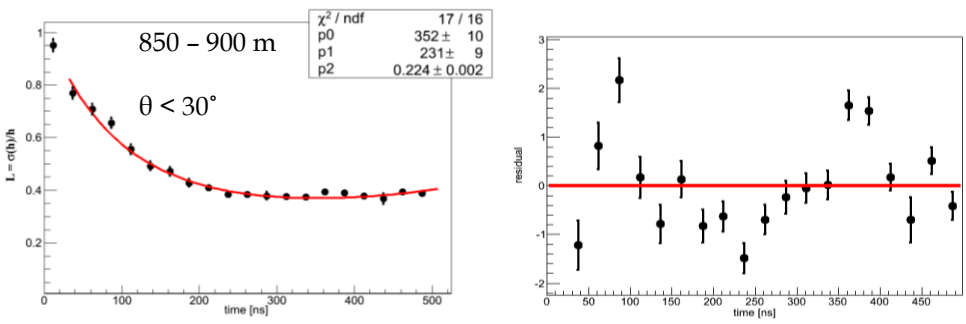


Figure 5.2-4 Fitting L-t curve with a modified Moyal function (left) and the fitted residuals (right). The χ^2/ndf of the L-t fit is 17/16.

The next part of this section is the exploration on the dependencies of L on r, S, sec θ and ζ. The relationship between L and core distance at time 112.5 and 312.5 ns is shown in Figure 5.2-5.

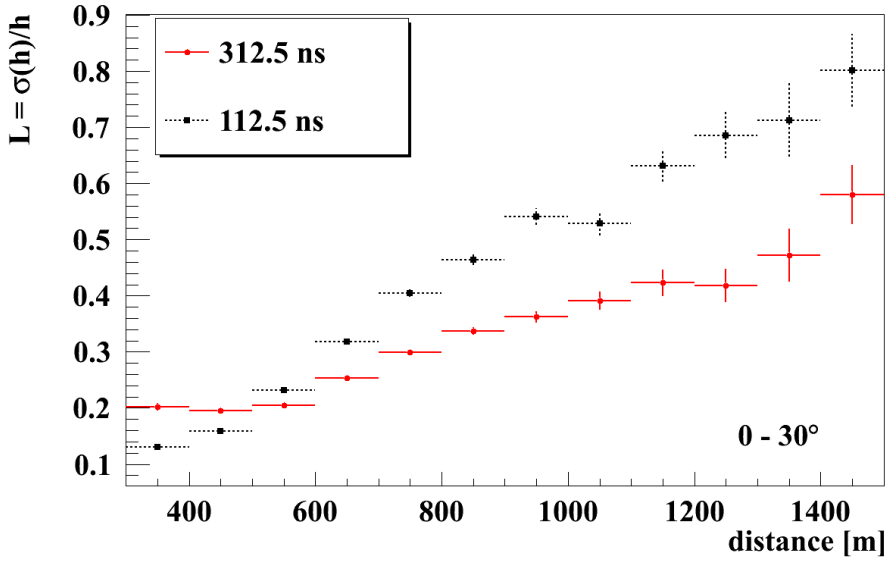


Figure 5.2-5 L and its dependence on core distance for t = 312.5 ns and 112.5 ns

It is reasonable to assume that for FADC traces from vertical showers, at 112.5 ns, the majority of the signal is from muons (Figure 4.3-1) and at 312.5 ns, the dominant signal is from electromagnetic component. In both cases L increases with larger core distances, although the change is slower at 312.5 ns. The L-t plot for showers with same zenith angle and detectors with same total signal but different core distances is shown in Figure 5.2-6. The time-period of the fitting is set the same with the mean time-trace fitting, which is between 25 – 275 ns.

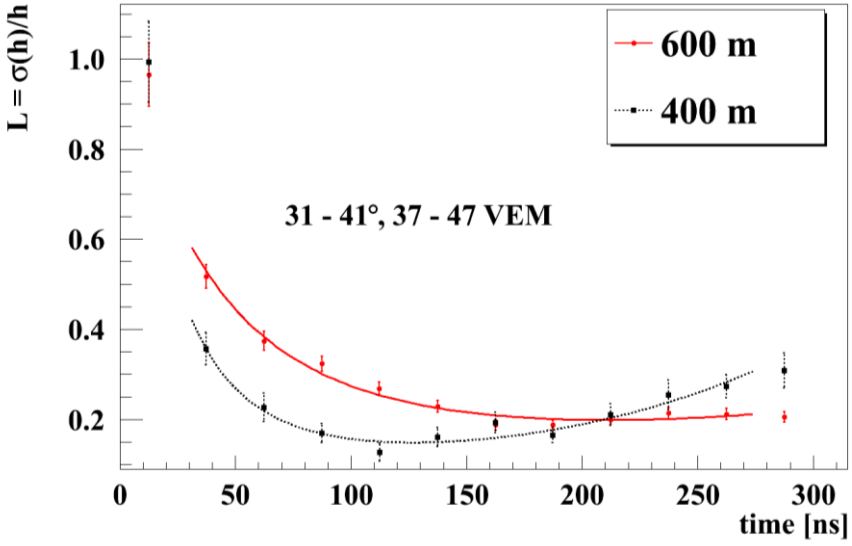


Figure 5.2-6 L-t plot for showers with different core distances

Similarly, the dependence on the total signal of the station is shown in Figure 5.2-7. Detectors from vertical showers and have core distance ~ 550 m are selected. The total signal varies from few tens of VEM to > 10^3 VEM.

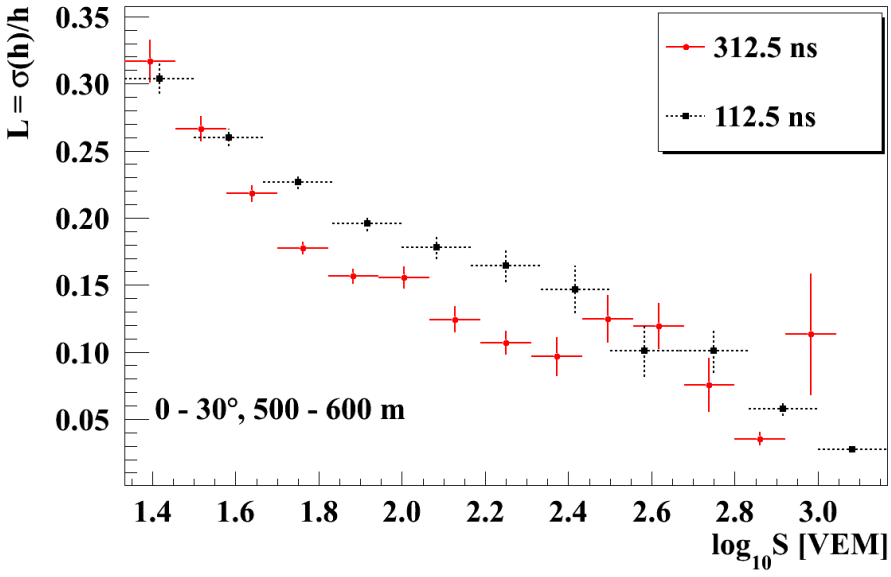


Figure 5.2-7 L and dependent on signal height for t = 112.5 ns and 312.5 ns

There are more particles in the station that has larger signals, which results in a smaller relative fluctuation. The L-t plot for showers with same zenith angle, core distance but different signal heights is shown in Figure 5.2-8.

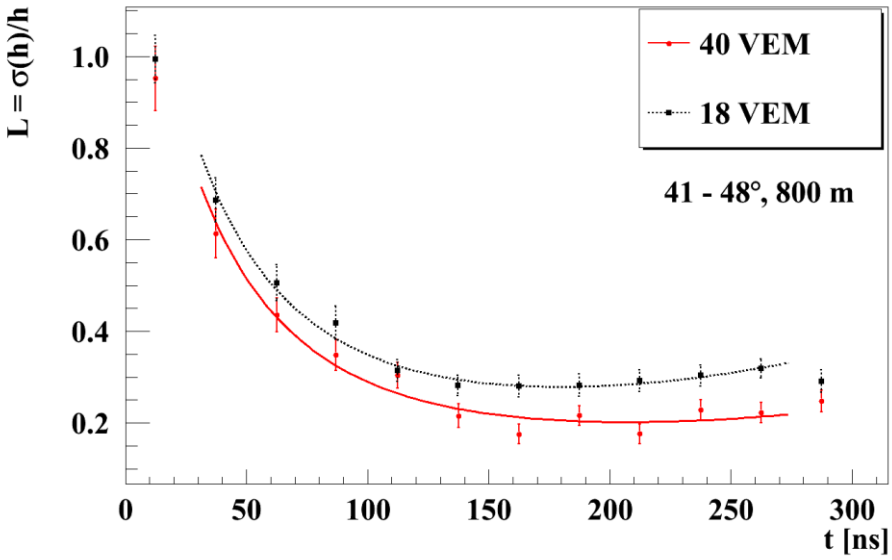


Figure 5.2-8 L-t plot for showers with different signal heights

The results on the signal dependence are consistent. However, because of too few twin-stations with high signals, extrapolations are needed to take account the signal effect.

The dependence on the zenith angle is shown in Figure 5.2-9. Detectors are selected with same core distance and signal heights.

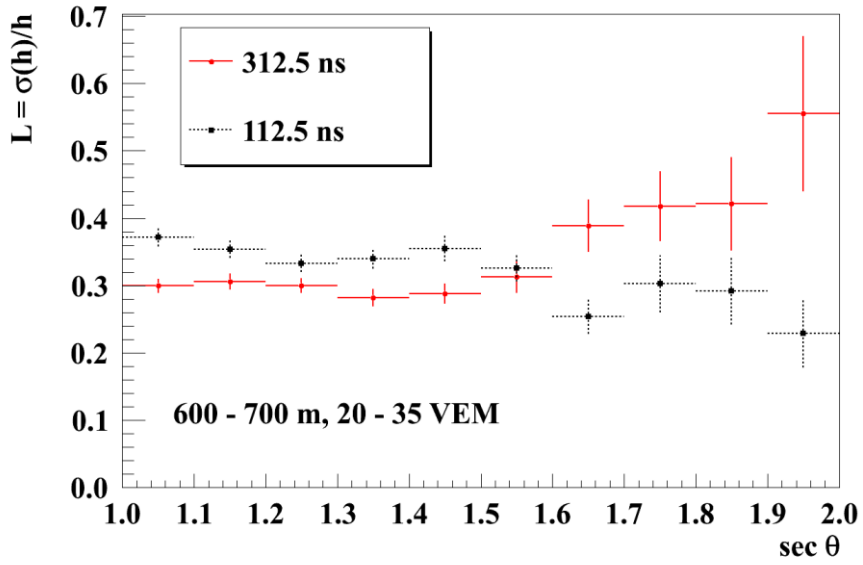


Figure 5.2-9 L and dependent on zenith angle sec θ for t = 112.5 ns and 312.5 ns

At 112.5 ns, the signals are mostly from muons and since the muon fraction increases with respect to the zenith angle, the relative fluctuation decreases. However at 312.5 ns, which is dominated by the electromagnetic signal, the relative fluctuation increases with the zenith angle. The L-t plot for showers with same core distance and signal but different zenith angle is shown in Figure 5.2-10.

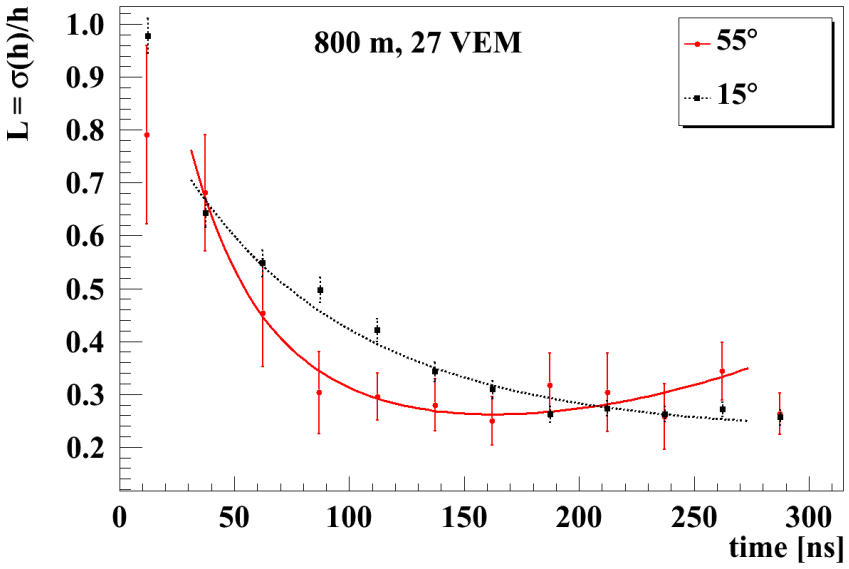


Figure 5.2-10 L-t plot for showers with different zenith angle sec θ

It is evident that L has to be parameterised with respect to zenith angle. The last dependence is on the polar angle (Figure 5.2-11).

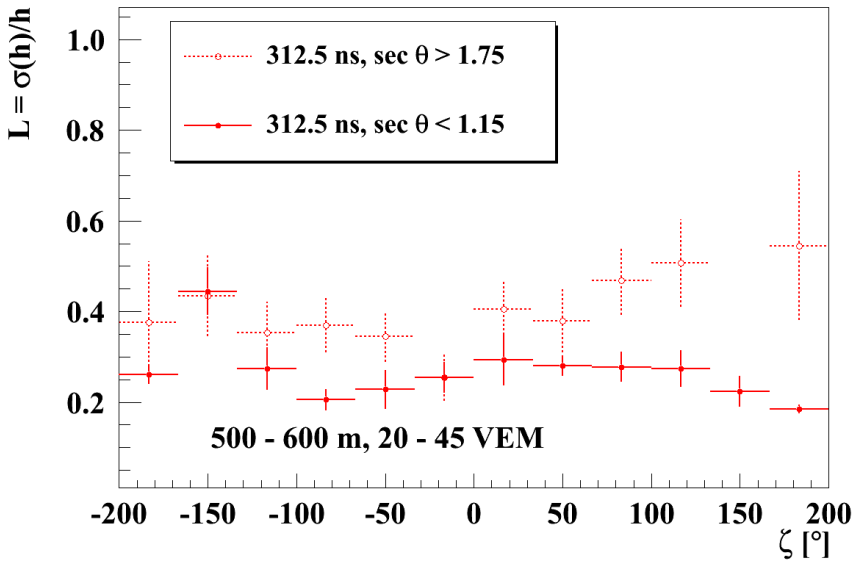


Figure 5.2-11 L and dependent on polar angle ζ for t = 312.5 ns

The dependence on the polar angle is not as apparent as the other three variables. Given the fact that the number of events is limited, it was decided to leave out the polar angle-dependence. Data from twin stations are binned in r (100 m bin width), S and sec θ (same binning with the mean trace-fittings)

The $\log_{10}\chi^2$/ndf of fitted results of L-t curves using the modified Moyal function is shown in Figure 5.2-12. The distribution is fitted with a Gaussian function. The parameterisations for signal variances at each 25 ns used in this thesis are stored at (required Auger 2014 password to access) http://www.ast.leeds.ac.uk/~phyll/Thesis/L_Para.txt.

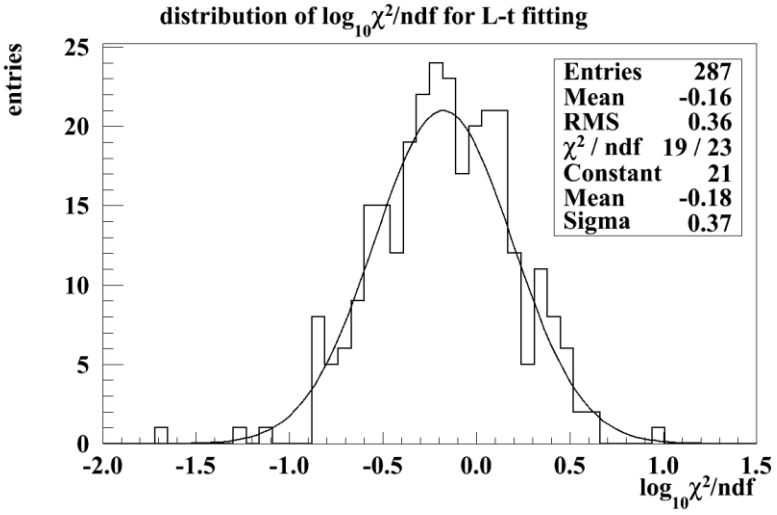


Figure 5.2-12 $\log_{10}\chi^2$/ndf from fitting L-t curves with a modified Moyal function

It is not reliable to simulate twin stations in the standard simulations because of the de-thinning procedures (section 2.4 and section 9.4). However it is possible to simulate ring detectors, which are detectors that are located at the same core distance in a ring. Twelve stations per ring were simulated at 750 m and 1150 m for proton and photon simulations. Using the same definition of L, Figure 5.2-13 and Figure 5.2-14 shows the signal variances from Monte Carlo calculations.

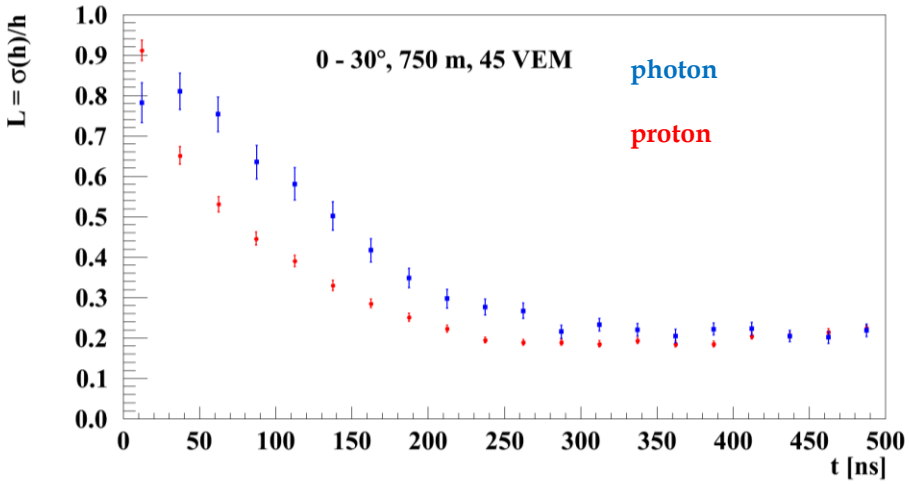


Figure 5.2-13 L determined from proton and photon simulations that are vertical

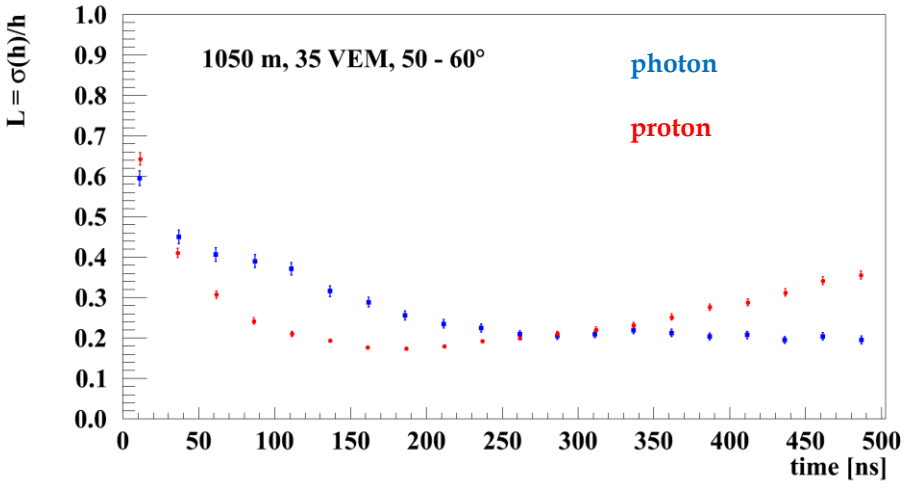


Figure 5.2-14 L determined from proton and photon simulations that are inclined

Considering the common time period of both zenith angle ranges where photons have larger L than that of protons, and to avoid the fluctuations due to the start time, the time period of the entity method was selected to be between 87.5 to 237.5 ns.

A simple check to test if the L measured in data agrees with Monte Carlo simulations was carried out. In this study[110] pair-stations were selected instead of using twins. Same cuts on the energy, zenith angle and the total signal of the station were applied to data, photon, proton and iron simulations. Figure 5.2-15 shows that the data agree with hadronic simulations at the very start of the time-trace but begin to lie above hadronic simulation results from ~200 ns. This could be because in simulations, although detectors are selected using the pairs algorithms[110], they are originally from rings of 12 simulated detectors around the core position in every 50 m, which is much shorter than the distance between pair-stations. Or it might be that the electromagnetic to muon ratio in simulations is higher than found in data[41].

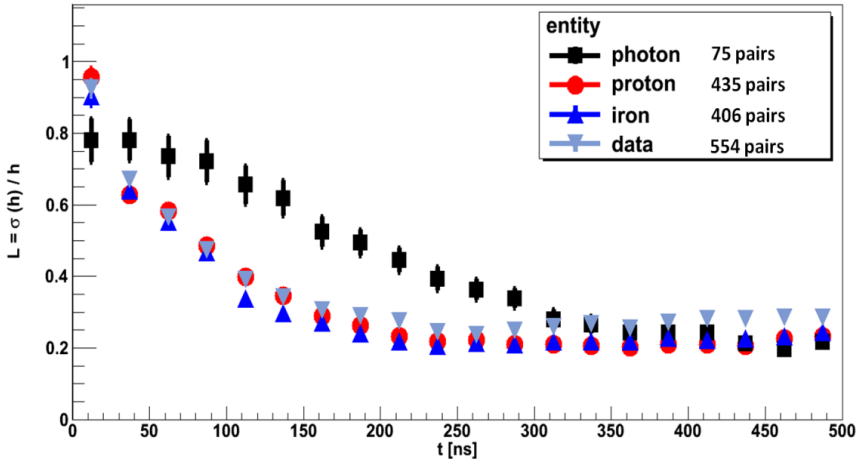


Figure 5.2-15 L from data comparing to L from thinned proton, iron and photon simulations. Pair-stations instead of twin-stations were used for data. It was shown that the value of L using pair-stations agree with values using twin-stations[110].

5.3 Parameterisation of the probability density functions for χ^2/ndf of traces

With the mean time-traces fitted to partly-normalised Landau functions and uncertainties of signal heights in each 25 ns fitted to Moyal-like functions using data, it is now possible to calculate χ^2/ndf of a measured FADC trace by comparing it to the mean trace at distance r, total signal S, zenith angle θ and polar angle ζ. There are in total 812,440 events with $E_{rec} > 10^{18}$ eV (on hadronic scale) and $0 < \theta < 60^{\circ}$ selected from January 2004 to March 2014. All events satisfy the basic data quality cuts namely T4 and 6T5 (section 2.1) and are not taken during bad periods. 649,838 events have at least one detector with S > 15 VEM and 350 < r < 1500 m. In total there are 1,029,611 detectors

selected from data, 48,132 detectors from proton simulations, 48,499 detectors selected from photon simulations. The distributions of χ^2/ndf are shown in Figure 5.3-1.

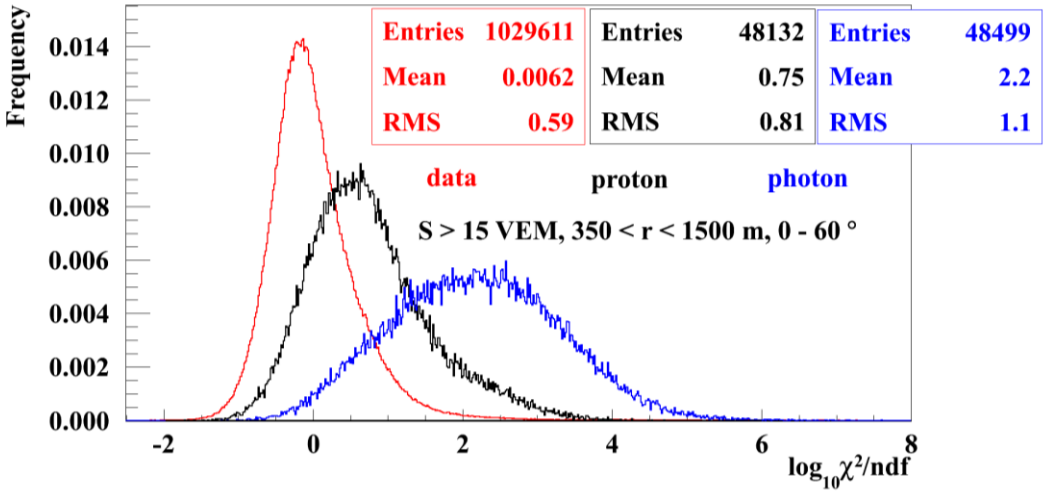


Figure 5.3-1 Distribution of $\log_{10}\chi^2$/ndf for all detectors from proton and photon simulations and data (01/2004 – 03/2014)

The mean of $\log_{10}\chi^2$/ndf for data is ~ 0, which indicates that the parameterisations of mean time-traces and signal fluctuations describe the data on average rather well. The mean from the proton simulations is between data and the photon simulations. This may be because a mixed-composition of light and heavy nuclei, and consequently the mean, is different from the pure-proton scenario. In addition the mismatch could also be due to problems with hadronic interaction models used in the proton simulations. Protons were simulated using QGSJetII-03, which is known to predict smaller number of muons compared to what is found in the data. As a summary, the photon-hadron separation based on $\log_{10}\chi^2$/ndf is evident on the station level. However, one needs an event parameter to quantify how photon-like the event is.

To combine values of $\log_{10}\chi^2$/ndf of each station in an event to a variable that describes the event, one could calculate the mean of $\log_{10}\chi^2$/ndf over all stations, use multivariate analysis to project values of all detectors onto one axis, or calculate the probability of the photon-likeness and combine them in an event probability.

A demonstration of the multivariate approach is shown in Figure 5.3-2 using simulations. For each event from proton and photon simulations, the two detectors with the highest signals are selected and values of $\log_{10}\chi^2$/ndf are drawn on x (station with the highest signal) and y axis. Both stations of the event have 30 < S < 200 VEM and 500 < r < 600 m. It is clear that by

projecting points to an optimised axis, the photon-hadron separation is enhanced more than if just one detector is used. However, there are at least three problems that limit this idea.

a). If the trace gives an exceptionally large $\log_{10}\chi^2$/ndf value (e.g. greater than 6) but behaves neither data-like nor photon-like (as might be caused by a problem with a PMT), it will be considered as photon-like and create pseudo-candidates.

b). Stations are not weighted according to the ability of distinguishing photons although at closer distances to the core position may have a different photon-hadron separation power compared to stations at far distances.

c). For inclined events, the number of stations that survive analysis-cuts is in general larger than 2. Using only two stations does not make use of all information associated with the event. However if one decides to use three or even more stations per event, the number of surviving showers would decrease dramatically, which is not appropriate for the purpose of photon-search.

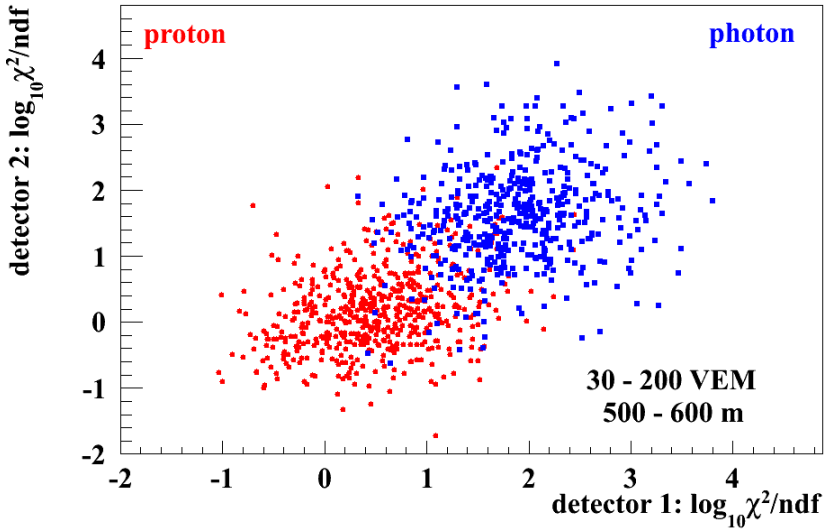


Figure 5.3-2 $\log_{10}\chi^2$/ndf from two detectors of each event from proton and photon simulations

The 'averaged-over-all' treatment faces criticisms similar to those from the multivariate analysis (point a and b). To solve question a, one needs to not just compare the measured trace with the mean of data, but also make the comparison with photon simulations. This solution requires a large number of simulations since there are more than 10,113 bins (Figure 5.1-7) in data and there are at least 5 entries in each bin. Alternatively, as introduced in this thesis, one could calculate the probability that the measured $\log_{10}\chi^2$/ndf

belongs to the expected mean of data over the probability that it belongs to the expected mean of photon simulations. The parameterisation of the mean requires data binned in distance, signal and zenith angle and photon simulations binned in distance and zenith angle (not in signals because of not enough number of simulations). For example in Figure 5.3-3, distributions of $\log_{10}\chi^2$/ndf for data and photon with varying core distances and 1.83 <sec θ < 2 are shown. The total signal of each selected station from data is ~ 45 VEM.

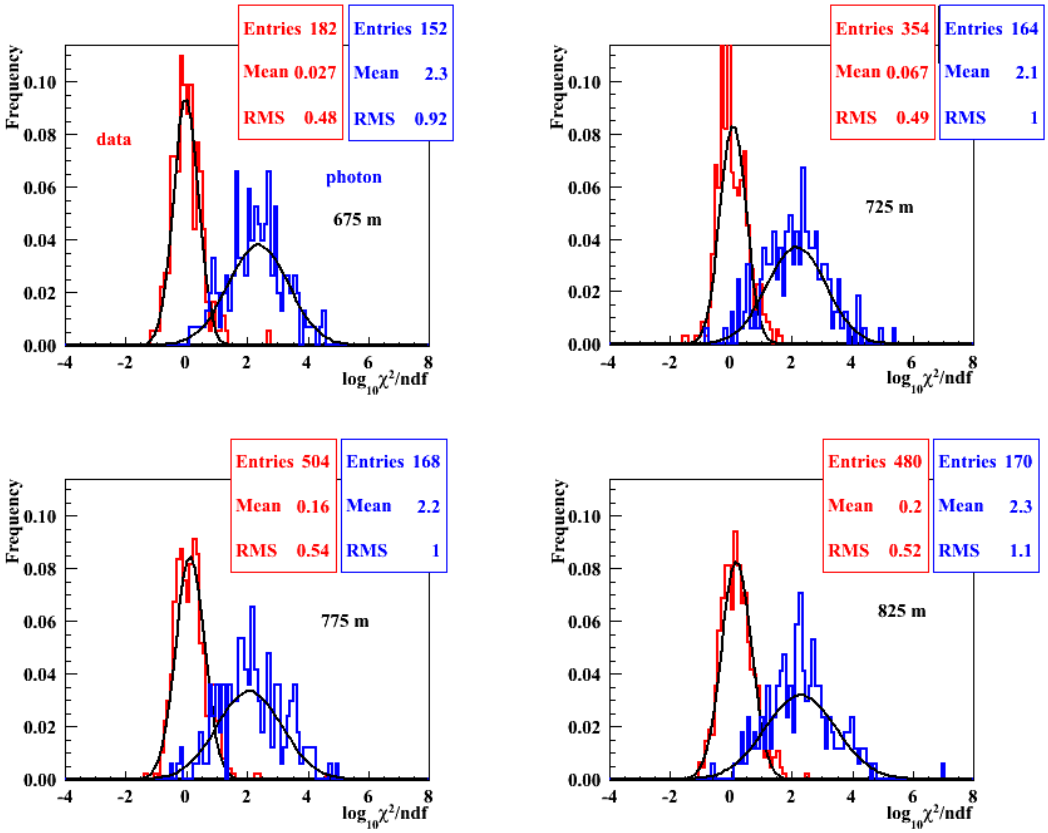


Figure 5.3-3 Distribution of $\log_{10}\chi^2$/ndf for detectors at different distances for data and photon simulations. Each distribution has been normalised and fitted with a Gaussian probability density function. Stations are taken from events that have 1.83 <sec θ < 2.

No dependence on θ has been found from the mean of data. An example of the relationship between the mean of $\log_{10}\chi^2$/ndf and sec θ is shown in Figure 5.3-4.

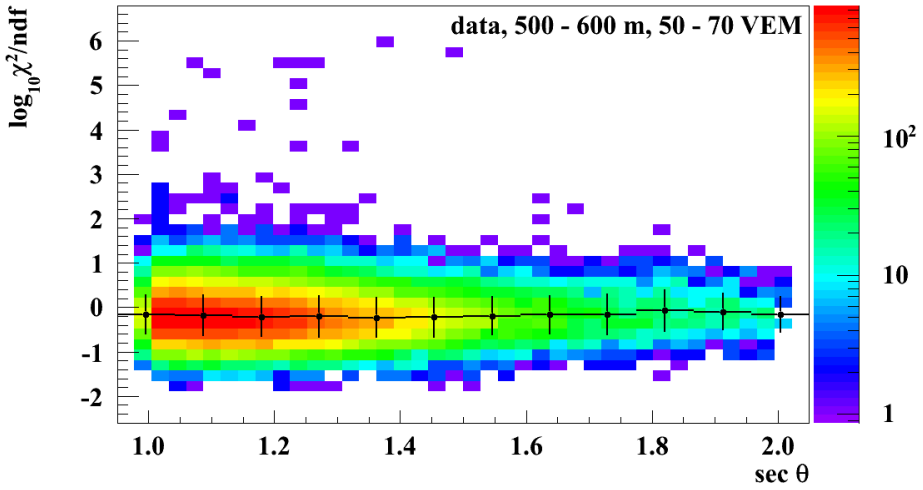


Figure 5.3-4 No dependences found from the mean of $\log_{10}\chi^2$/ndf and on sec θ for data. The error bars show one sigma of the distribution. Detectors are with 500 < r < 600 m and 50 < S < 70 VEM

The dependence on the signal is also not evident. Shown in Figure 5.3-5 is the mean with respect to the core distance for three different signal bands.

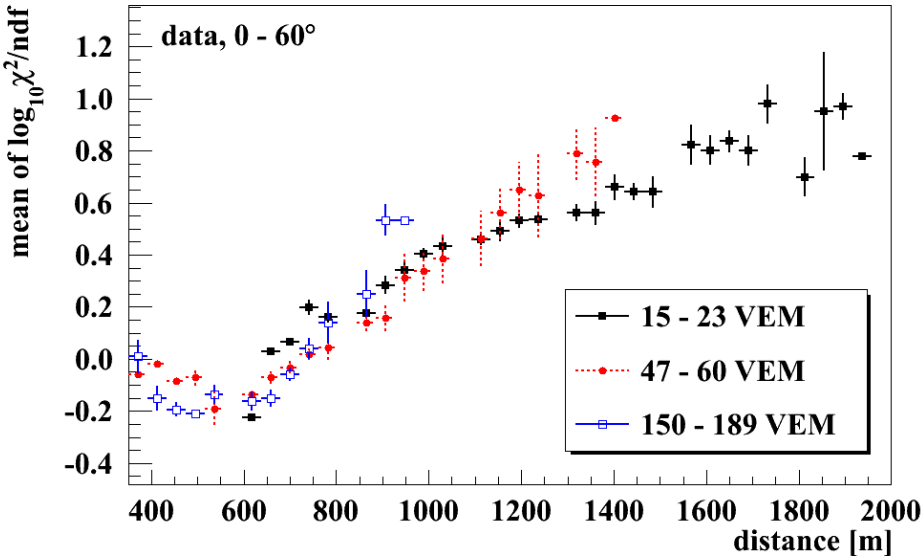


Figure 5.3-5 Distribution of Gaussian mean of $\log_{10}\chi^2$/ndf with respect to core distance for data that have various total signals. No dependence on signal found.

It is thus safe to conclude that the mean only has dependence on the core distance. Using detectors with S > 15 VEM and 0 < θ < 60˚, the mean of data has been fitted with two linear functions (Figure 5.3-6) that covers r < 550 m and r >= 550 m respectively.

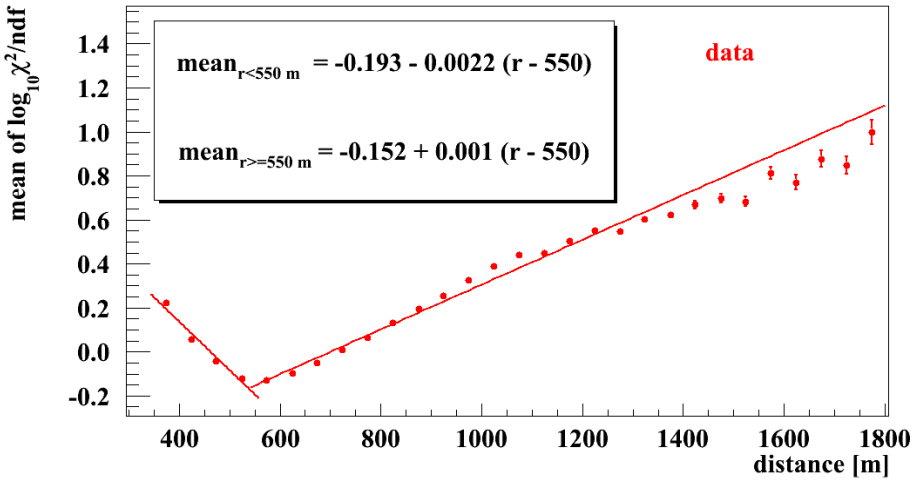


Figure 5.3-6 Relationships between mean of Gaussian distributions of $\log_{10}\chi^2$/ndf and core distance for data. The dependence on the core distance suggests that the parameterisation of the mean trace gives the best description of the data at 500 – 700 m, where most of the data are concentrated. More investigations are needed and the mean-trace parameterisations might be improved as there should be no dependence of $\log_{10}\chi^2$/ndf on the distance. In this thesis we parameterise the dependence on the core distance using two linear fits to the data.

Similarly for photon simulations, the mean of $\log_{10}\chi^2$/ndf is shown as a function of core distance in Figure 5.3-7.

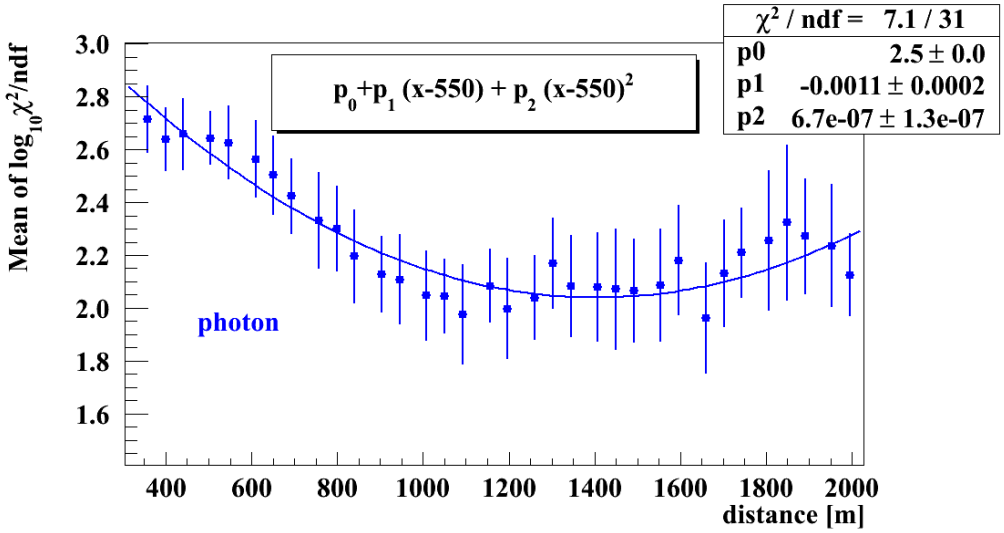


Figure 5.3-7 Relationships between mean of Gaussian and core distance for photons

.

The mean of photon has been fitted with a second-order polynomial function. To calculate the probability, one also needs to know the sigma of the pdf. The distribution of sigma from data is shown in Figure 5.3-8.

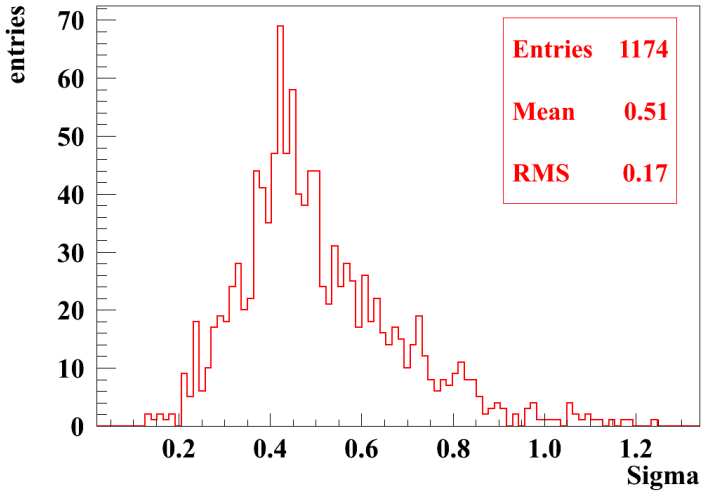


Figure 5.3-8 Distribution of sigma for the Gaussian probability density function of data. It was decided to use the mean of this distribution for the sigma of pdf for both data and photon simulations.

It was found that the dependences of sigma on core distance, signal or zenith angle are not apparent and therefore the sigma is fixed to the mean of the data, 0.51, for all probability density functions. The sigma of photons is set to be the same with the sigma of data. This is because issues of artificial fluctuations in photon simulations brought from thinning processes (section 2.4 and section 9.4).

Thus the probability of the station with $\log_{10}\chi^2/\mathrm{ndf}$ compared to the mean of data and photon can be calculated respectively. The ratio of the two probabilities is commonly used in applications as the likelihood ratio. The relationship between the likelihood ratio and the station distance is shown in Figure 5.3-9.

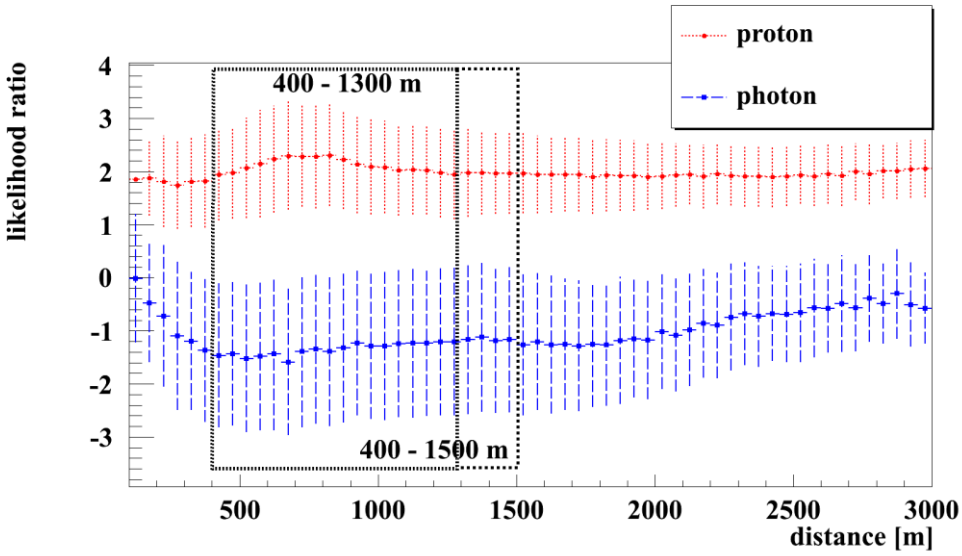


Figure 5.3-9 The likelihood ratio of stations and the dependence on distance of proton and photon simulations. It was decided to select at least one detector with distance 400 – 1300 m and one detector with distance 400 – 1500 m.

Similarly the relationship between the likelihood ratio and the station signal is shown in Figure 5.3-10.

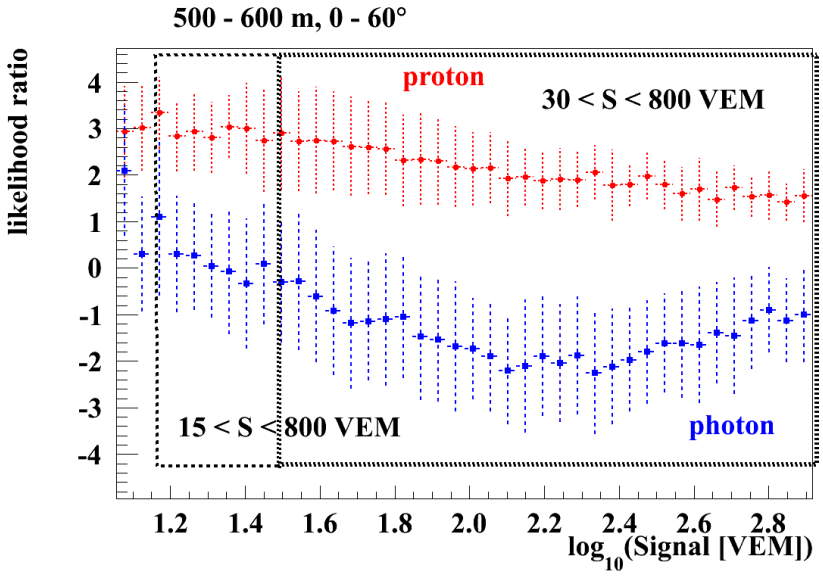


Figure 5.3-10 The likelihood ratio of stations and the dependence on the station signal. It was decided to select at least one detector with signal 30 – 800 VEM and another detector with signal 15 – 800 VEM. No low-gain saturated stations are included.

To calculate the event parameter, at least two stations are needed to be selected. One detector is required to be located within 400 – 1300 m from the core position and have signal of 30 – 800 VEM. The other detector is required to have a core distance 400 – 1500 m and signal 15 – 800 VEM. The event parameter is to combine the likelihood ratio of all detectors, and defined as

$$LL_{normalised} = \lg \frac{\prod_{1}^{n}(P_1, P_2 ...)}{\prod_{1}^{n}(P_1^{\gamma}, P_2^{\gamma} ...)} \div n \qquad 52,$$

where P_n is the probability that the time trace of the n_{th} station is data-like and P_n^{γ} is the probability that the trace is photon-like. The normalisation is to made by calculating the average over n stations.

The value of $LL_{normalised}$ is named the entity likelihood ratio, which is used for the photon search and discussed in the next chapter.

Chapter 6. Applying the entity method to search for ultra-high energy photons

The likelihood-ratio-parameter introduced in Chapter 5 will now be used in the search of UHE photons. In this chapter, the dependences of the entity likelihood ratio on the number of selected stations, zenith angle and energy are first explored. The final parameter – the Δ-entity likelihood – ratio is defined and the numbers of candidates in various energy bands are described. Then photon-like events are categorised and examples are listed. In particular, the possible causes of candidates with a signature of 'pre-pulses' are discussed. A statement of number of photon candidates is made.

6.1 Entity likelihood ratio on the event-to-event level

The idea of how to calculate the entity likelihood ratio has been introduced in section 5.3. To make it clearer an example of the procedure using event 627697 is demonstrated. The event (Figure 6.1-1) has θ = 38.5° and 6 detectors triggered. Detector 143, 135 and 145 passed the selection defined in section 5.3. Detector 143 is located at r = 543 m, ζ = 173° and has S = 242 VEM. The FADC trace is shown in Figure 6.1-2 a.

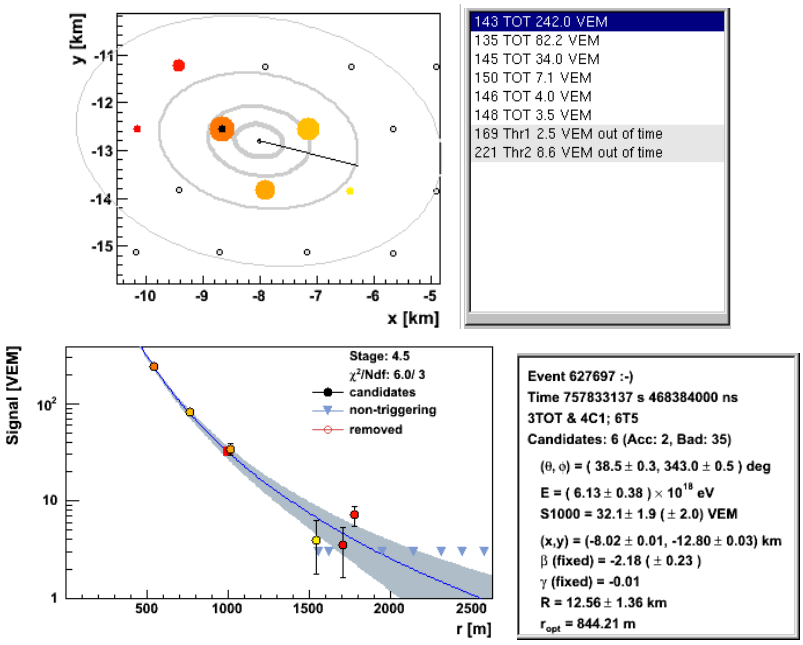


Figure 6.1-1 Event used to show the procedure of calculating the entity likelihood ratio . The top left plot describes the geometry of all detectors triggered in the event. The size of dots is proportional to the size of the signal in the station; the colour shows the triggered time as stations triggered, the latest marked in red. Signals of each station are listed on the right. The lateral distribution and the summary of the event detail are shown at the bottom.

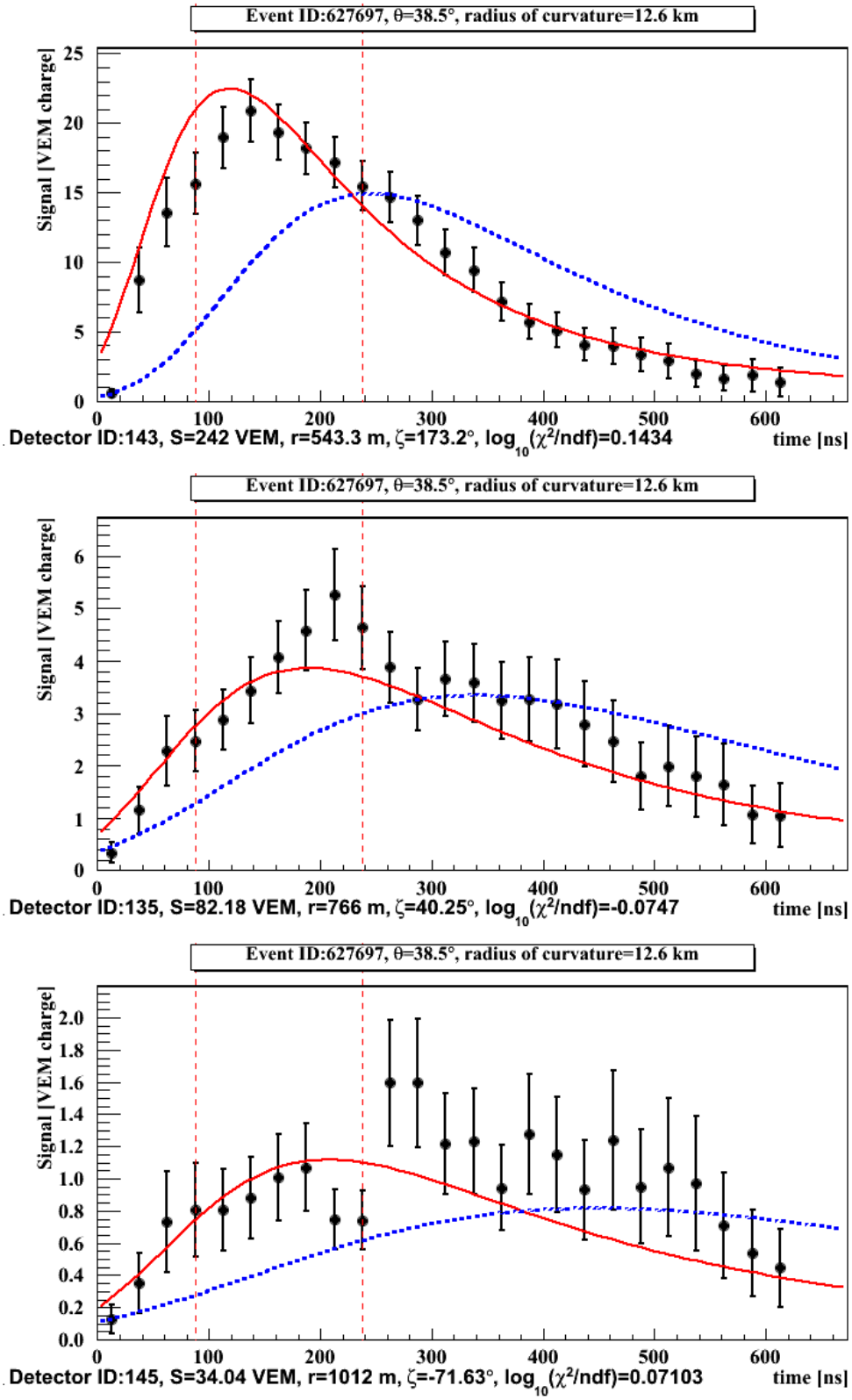


Figure 6.1-2 FADC traces from station 143 (a), 135 (b) and 145 (c) of the event 627697. Black dots are from data and the error bars are from parameterisation of signal variance using twin stations. The red curve is the mean of data, parameterised using a partly-normalised Landau function. The dotted blue curve is the mean time-trace from photon simulations and is only for reference. It is not used in the calculation of the entity parameter. The dashed red lines illustrates the time period used to calculate the $\log_{10}\chi^2/\text{ndf}$, which is between 87.5 – 237.5 ns. All three stations have traces that are data-like.

According to parameterisations introduced in Chapter 4, signal heights and uncertainties in each 25 ns of the FADC trace are estimated.

Table 6.1-1 Measured, expected and sigma of signals of each 25 ns at the beginning of the FADC trace for station 143

Time [ns]	$S_{measured}$ [VEM]	S_{exp} [VEM]	Sigma [VEM]	χ^2/ndf
87.5	15.66	20.85	2.17	5.72
112.5	18.95	22.41	2.22	2.43
137.5	20.87	22.04	2.22	0.28
162.5	19.30	20.50	1.98	0.37
187.5	18.18	18.40	1.88	0.01
212.5	17.17	16.19	1.85	0.28
237.5	15.48	14.08	1.78	0.62

The sum of χ^2/ndf is ~ 9.7, which makes the average as ~ 1.39 ($\log_{10}\chi^2$/ndf = 0.14). The expected value of $\log_{10}\chi^2$/ndf from data and photon simulations were parameterised as functions of the core distance (Figure 5.3-6, Figure 5.3-7) and are found to be -0.18 and 2.54 for data and photon respectively. Therefore the probability of the measured $\log_{10}\chi^2$/ndf to be from a Gaussian distribution with sigma 0.51 and mean of data is 64% whereas the probability to be from the photon distribution is 0.0013%. The ratio of the two probabilities is 5×10^4, which makes the log likelihood ratio 4.7.

Table 6.1-2 Summary of quantities used for calculating the log likelihood ratio of the station 143

Data Gaussian mean	Photon Gaussian mean	Prob data	Prob photon	Likelihood ratio	Log likelihood ratio
-0.18	2.54	0.64	1.3×10^{-5}	5×10^4	4.70

Similarly the log likelihood ratios are obtained for the other two selected detectors and the results are 4.75 and 3.54. The averaged log likelihood ratio is named as the entity likelihood ratio, which is (4.70 + 4.75 + 3.54) / 3 = 4.3.

The entity likelihood ratio does not have dependence on the zenith angle as shown in Figure 6.1-3.

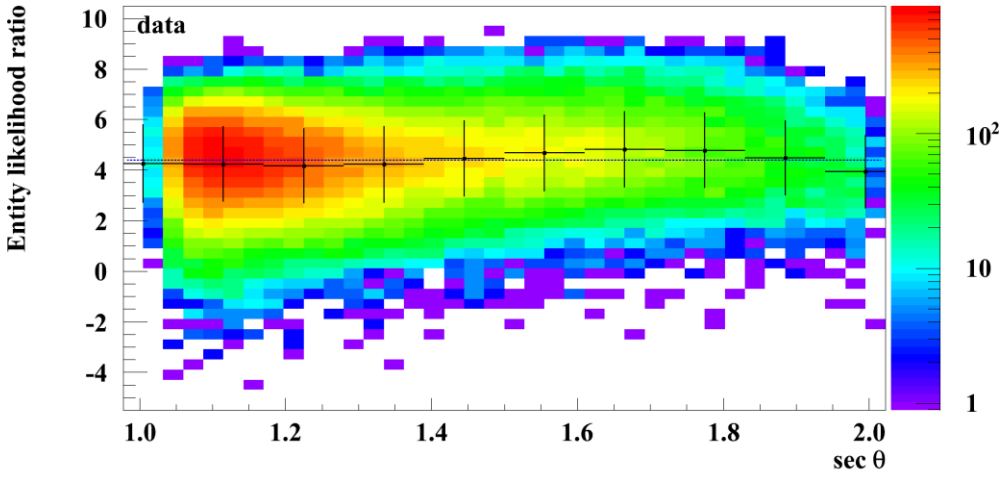


Figure 6.1-3 No dependence found for the entity likelihood ratio from zenith angles for data

The fact that the entity likelihood ratio is independent of zenith angle allows us to combine results of events with $0 < \theta < 60^{\circ}$. Photon simulations and data that have energy on the photon scale (as defined in section 7.1) $E^{\gamma}_{rec} > 10^{19}$ eV are selected and the distribution of entity likelihood ratio is drawn in Figure 6.1-4.

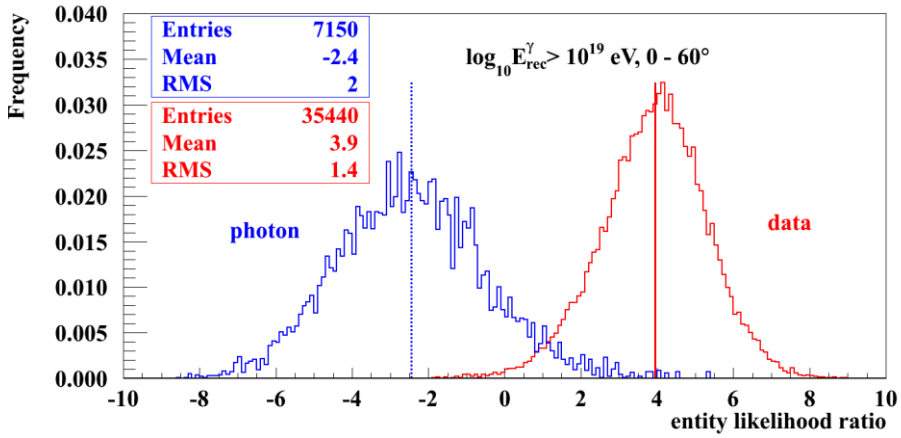


Figure 6.1-4 Entity likelihood ratio of photon simulations and data have $E^{\gamma}_{rec} > 10^{19}$ eV. Full zenith angle range $0 < \theta < 60^{\circ}$, Merit factor (defined in Chapter 3) is ~ 2.6. Straight lines indicate positions of median of each distribution. Data used from 01/04 to 03/14. Photon simulations are with E^{-2} spectrum and include pre-showers.

However, when plotting the entity likelihood ratio of photon simulations and data in narrower energy bands, it is clear that there is a dependence on

the energy. The distribution of data and photon simulations in each energy range is shown in Figure 6.1-5 and the median of each distribution are indicated.

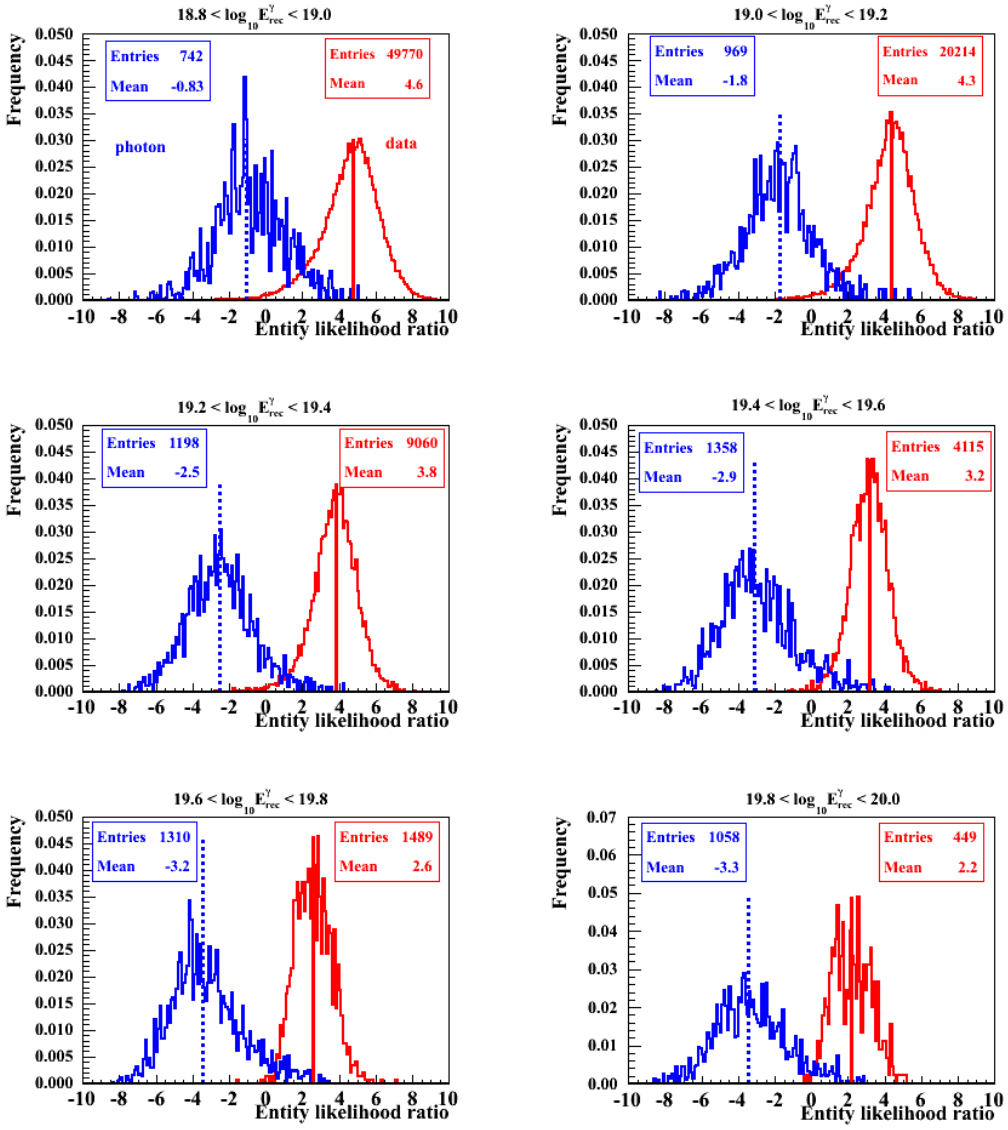


Figure 6.1-5 Entity likelihood ratio (LLR) of data and photon simulations (LLR_γ) for differential energy bins from $E^\gamma_{rec} = 10^{18.8}$ eV to 10^{20} eV and zenith angle $0 < \theta < 60^\circ$.

The dependence is summarised in Figure 6.1-6. All data, proton and photon simulations have the mean of the entity likelihood ratio decrease as the energy becomes larger. To be consistent with other work done by the Auger photon group, the cut to select photon candidates is going to be set at the median of the photon distribution. It is then mandatory to 'correct' the dependence on the energy if one wants to set the median cut on integrated energies.

The solution is to introduce the Δ-entity likelihood ratio (ΔLLR), which is defined as

$$\Delta LLR = (p + LLR - LLR_{\gamma}) / q \qquad 53,$$

where LLR is the entity likelihood ratio of the event, LLR_{γ} is the expected value of the entity likelihood ratio from photon simulations, q is the sigma of the distribution of LLR_{γ} and p is a constant.

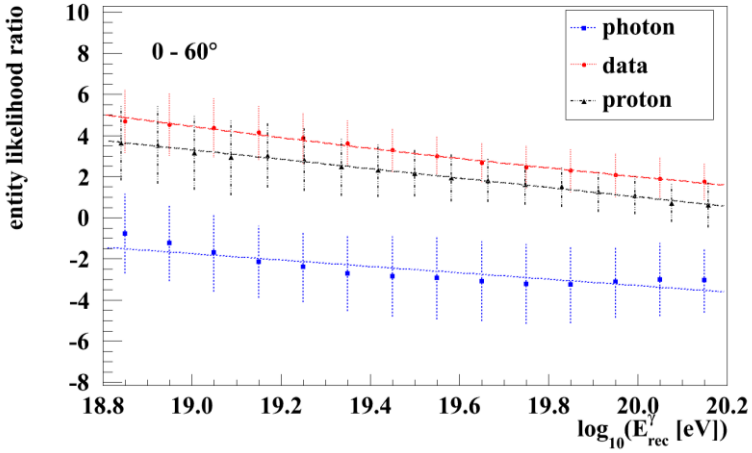


Figure 6.1-6 Energy dependence of entity likelihood ratio for data, protons and photons

The sigma is found and set as 2 for all energy bins, the median values are tabulated in 7 bins. Values of p are adjusted so that the ΔLLR has the median of -0.15 for photon simulations at all energy intervals. For energy intervals of $\log_{10}E^{\gamma}_{rec} \in$ (18.8, 19.0), (19.0, 19.2), (19.2, 19.4), (19.4, 19.6), (19.6, 19.8), (19.8, 20.0), (20.0, 20.2), (20.2, 20.4) and $E^{\gamma}_{rec} > 10^{20.4}$ eV, p + LLR_{γ} are -0.69, -1.5, -2.2, -2.9, -3.2, -3.2, -2.7, -2.6 and -2.2 respectively. The example event 627697 as shown earlier has the entity likelihood ratio 4.3 and energy $E^{\gamma}_{rec} = 10^{19.25}$ eV. The Δ-entity likelihood ratio is therefore (4.3 – (–2.2)) / 2 = 3.3.

In the next section, some basic checks for Δ-entity likelihood ratio are performed. The photon-hadron separations based on the entity method for various energy bins are presented. The separation power is compared to other SD parameters introduced in Chapter 3. The number of photon-like events is summarised.

6.2 Δ-entity likelihood ratio and photon-hadron separation

The Δ-entity likelihood ratio, which describes the deviation of the measured entity likelihood ratio from the expected value from the photon simulation, was introduced to set the cut to select photon candidates from data. The Δ-entity likelihood ratio does not depend upon the zenith angle (Figure 6.2-1),

the number of stations selected in the event (Figure 6.2-2) or the energy of the event (Figure 6.2-3). The median of the photon simulations is -0.15. The error bars show 1 sigma of the distribution.

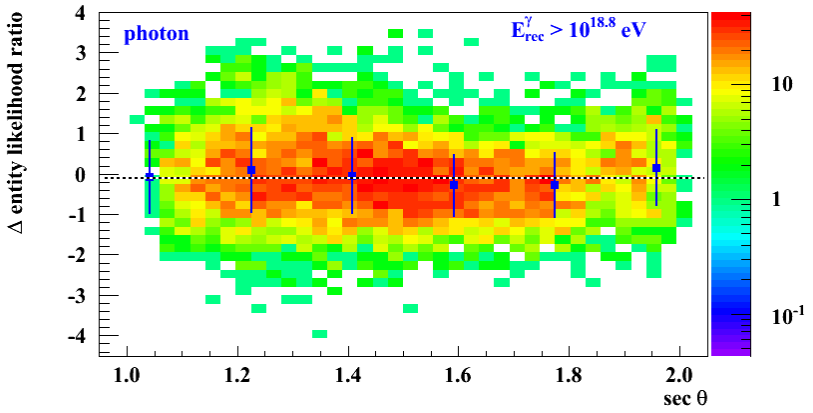


Figure 6.2-1 Δ-entity likelihood ratio and zenith angle from photon simulations

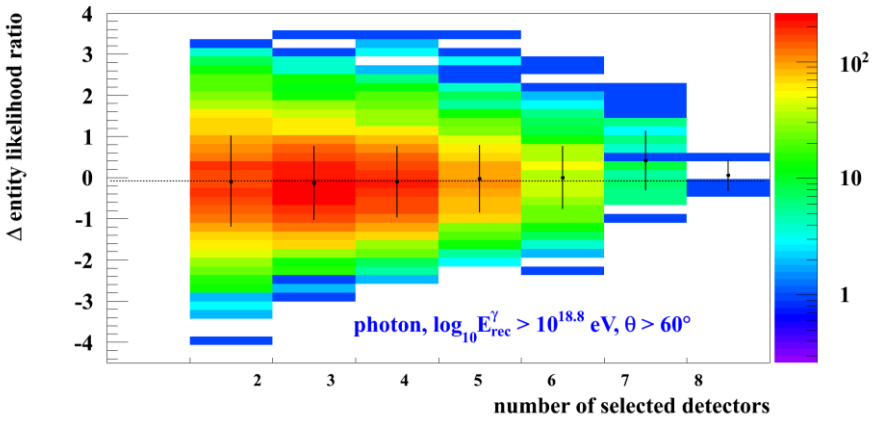


Figure 6.2-2 Δ-entity likelihood ratio and number of stations selected in the event from photon simulations

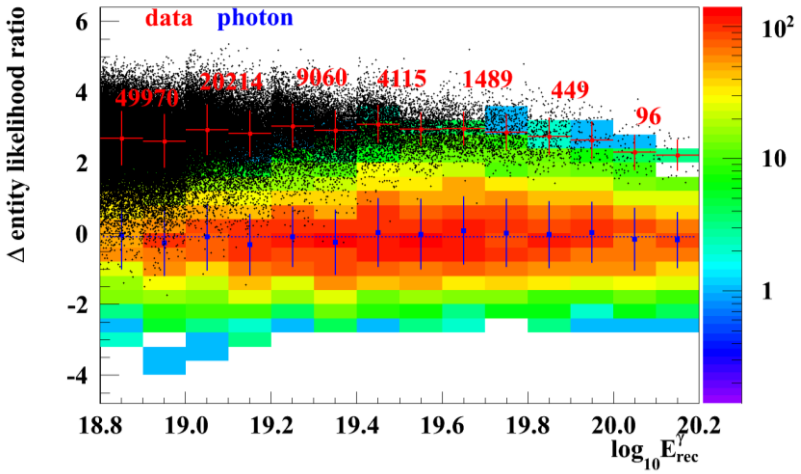


Figure 6.2-3 Δ-entity likelihood ratio and reconstructed energy in the photon scale E^{γ}_{rec}. Black points are from data with number of entries in each energy band shown. The colour shows the density of number of events from photon simulations. The mean of photons are in blue and fitted with a horizontal line.

The photon-hadron separations based on Δ-entity likelihood ratio in 6 differential energy bins are shown in Figure 6.2-4.

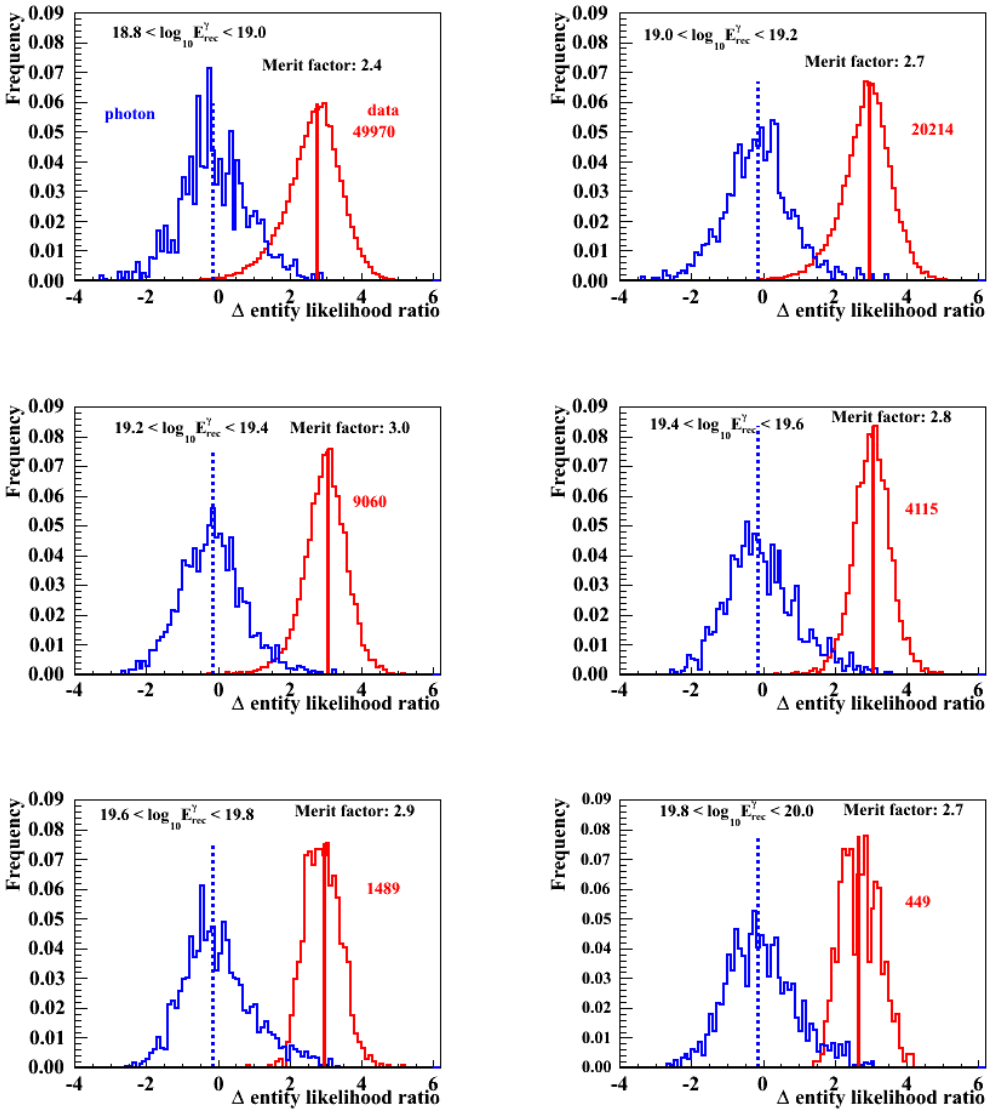


Figure 6.2-4 Photon-hadron separation based on the Δ-entity likelihood ratio parameter. Events with Δ-entity likelihood ratio < -0.15 are identified as photon candidates. There are103, 10, 3, 0, 0, and 0 candidates found respectively of energy intervals from $10^{18.8}$ eV to 10^{20} eV

The merit factors are shown for each energy band. The photon-hadron separation for integrated energy $E^{\gamma}_{rec} > 10^{19}$ eV is shown in Figure 6.2-5 and the merit factor is ~ 2.7. As a comparison, merit factors for other parameters based on surface detectors are shown in Table 6.2-1. These results use the same photon simulation library as for this work and are for $E^{\gamma}_{rec} > 10^{19}$ eV, $\theta > 30^{\circ}$ and the index of the photon spectrum -2. The cuts included are rather strict compared to what are required for the entity method. For example the radius of curvature cannot be fitted unless there are more than 5 stations in

the event. The risetime parameter requires a minimum of 4 non-saturated stations with S > 10 VEM. More importantly, pre-showers (section 3.1) are not included in the calculation of the merit factors for risetime, radius of curvature or S_b. Pre-showers have smaller X_{max} that are more data-like (Figure 3.1-5), which means that by excluding pre-showers the merit factor increases.

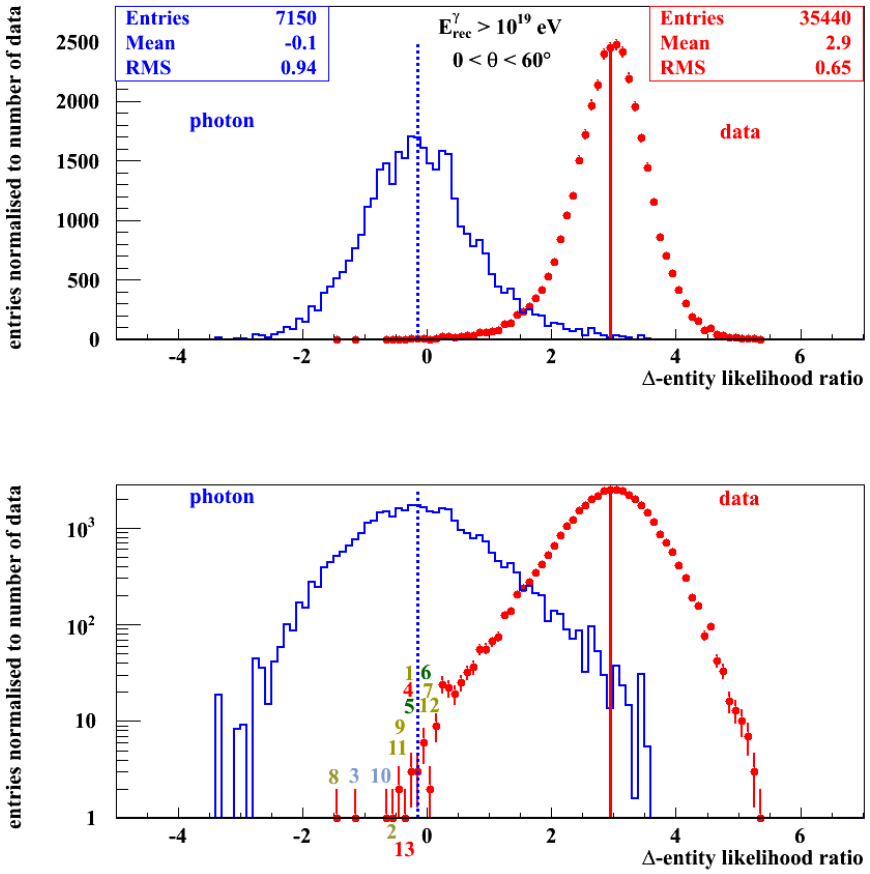


Figure 6.2-5 Photon-hadron separation from Δ-entity likelihood ratio for E^{γ}_{rec} > 10^{19} eV. The two plots are the same except that the one below is in logarithmic scale. The median of data is 2.95 and -1.5 for photon simulations. The merit factor is ~ 2.7. The energy spectrum of photon simulations is -2 and pre-showers are included. There are in total 13 events from data with Δ-entity likelihood ratio < -1.5. These events are labelled from 1 to 13 and positioned according to the value on the x-axis. The colours are defined in Table 6.4-1.

Table 6.2-1 Merit factor from other parameters that based on the surface detectors[104]

Parameter	Merit factor
$\Delta \log_{10} S_4$	1.3
$\Delta t_{1/2}(1000)$	2.6
ΔR_C	1.9

To summarise, during the calculation of the merit factor, there are four advantages of the entity method compared to the traditional SD methods, radius of curvature R_c, risetime $t_{1/2}$ and the parameters based on the lateral distribution function S_4.

i). More relaxed cuts and higher selection efficiency (than R_c and $t_{1/2}$)

ii). Valid for the full zenith angle range 0 – 60˚

iii). The entity method can also be applied to showers with lower energies ($10^{18.8}$- $10^{19.0}$ eV, which is the first time that a SD-only parameter has been used to reach such low energy.

iv). Pre-showers are included for the calculation of the merit factor of the entity method. The effect starts to be important from $10^{19.6}$ eV.

The merit factor of the entity method is higher than the highest value from the table. The distribution of photon-hadron separations according to the entity method in some of the other integrated energy bins are shown in the appendix B.

The next section is about cuts on the station level that have been applied in the photon search. The cuts are designed to remove stations from mis-reconstructed events and stations with wrong start times.

6.3 Station–quality cuts

It was known from inspecting detectors with extreme large $\log_{10}\chi^2$/ndf that it is possible to have FADC traces that are either data-like or photon-like but generating huge values of $\log_{10}\chi^2$/ndf. In this section, two types of such detectors are shown. It is necessary to remove detectors with those features from the data sample before the photon-search, to minimise the number of pseudo-candidates found in the end.

The first situation is when the traces are very fast. The time range used in the $\log_{10}\chi^2$/ndf calculation is from 87.5 to 237.5 ns. If the trace stops at around 250 ns or before, the entity method fails due to time bins have (nearly) no signals. Photons showers have a more delayed electromagnetic component which only makes the trace longer than data instead of being the opposite. An example of such station is shown in Figure 6.3-1. It is highly unlikely to have a station at 550 m from a shower of 40˚ to have a trace with risetime of 45 ns. The second station is located at 957 m but with the risetime 35 ns, which is below the risetime of a single muon going through the water-Cherenkov detector (40 ns). The suspicion is that the event was from an

inclined shower but with so few numbers of triggered stations, the zenith angle was miss-reconstructed.

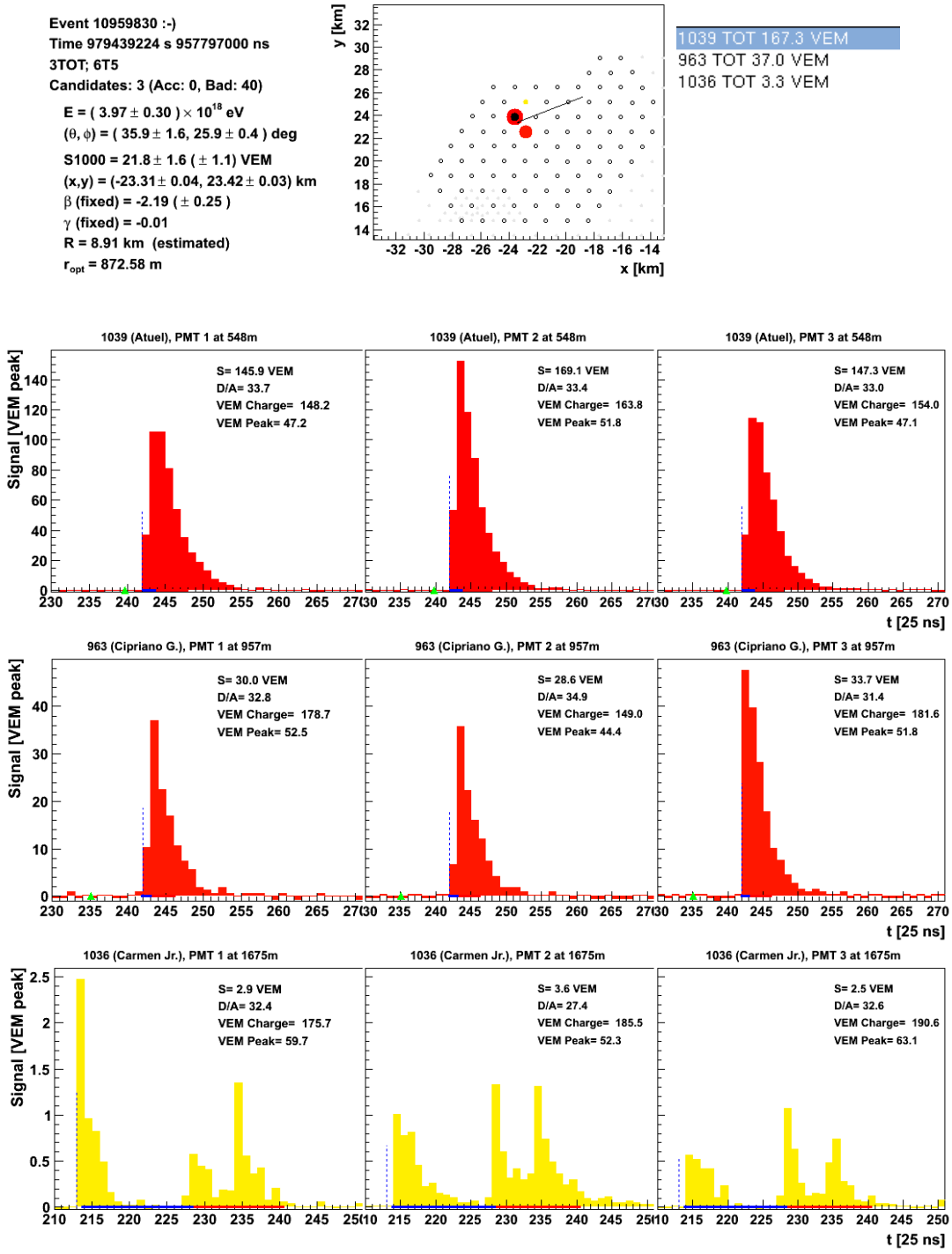


Figure 6.3-1 An example of event with more than one station have extremely fast traces. The risetime of three stations are 45, 35 and 370 ns respectively (after accounting risetime asymmetry due to polar angles, risetime are 46, 32 and 286 ns). The $\log_{10}\chi^2$/ndf of three stations are 2.1, 1.5 and -0.6. No events of such feature were found in photon, proton or iron simulations.

The other case considered is that the start time of a station is miscalculated due to photomultiplier problems such as an unstable baseline. One example

is shown in Figure 6.3-2. Three plots are time-traces obtained from three PMTs at the same station. The dashed line with the green triangle at the bottom indicates the start time of the trace. Even by eye it is obvious that there is no signal at the marked start time.

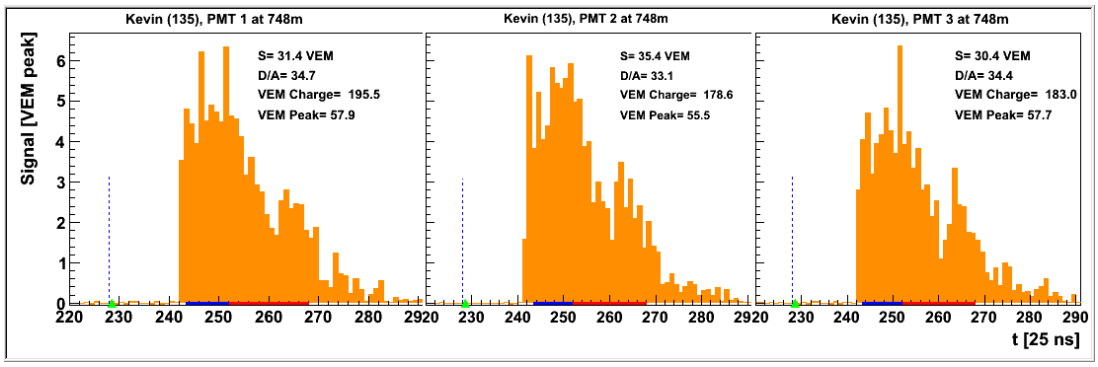


Figure 6.3-2 Example of PMTs with wrong start time due to fluctuations of the baseline. The dashed lines with the green triangle at the bottom show the position of the start time. Three traces were obtained in three PMTs from the same station.

Since identifying this problem a new algorithm has been tested[111]. The method is designed to reject stations of accidental signals from muons and all parameters are optimised for very inclined showers ($\theta > 60^\circ$). The dashed lines in Figure 6.3-3 show the new start time of the station.

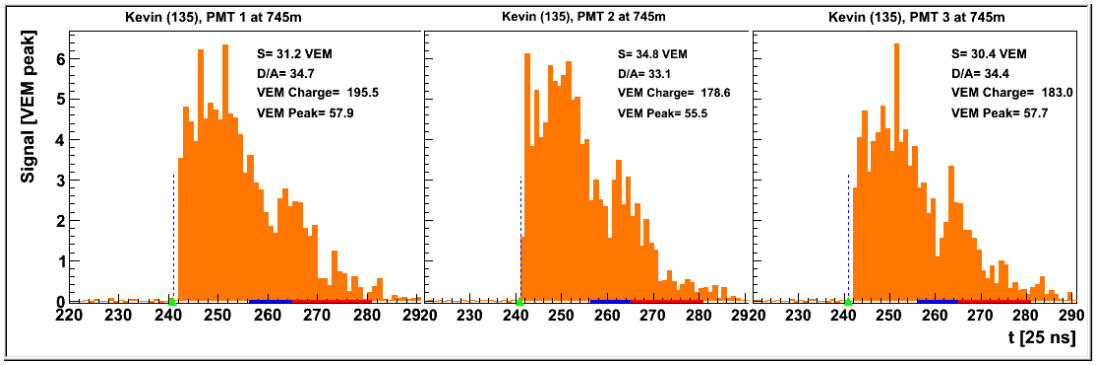


Figure 6.3-3 The corrected start time from using a new algorithm that optimised for rejecting accidental muons for the neutrino searches. Distance and the signal of the station does not change compared to before the correction, but the zenith angle has changed 1.5° for the event.

To use this accidental-signal-rejection algorithm for events with $\theta < 60^\circ$, systematic studies on how much does this change impacts on the other reconstructed quantities are needed. The parameters should also be optimised for vertical showers.

As a temporary solution, to remove stations with 'extremely fast' traces and ambiguous start time, cuts that are based on signal fractions of the FADC traces are introduced:

- The integrated signal up to the 4th time bin (time bin start from 1), S_4, over the signal included in the first 8 time bins, S_8, must be smaller than 70%
- S_8 over the total signal S must be smaller than 80%
- S_8 over the total signal S must be larger than 2%.

The distributions of S_4/S_8 and S_8/S of showers that have $0 < \theta < 30^\circ$ and $50 < \theta < 60^\circ$ are shown in Figure 6.3-4.

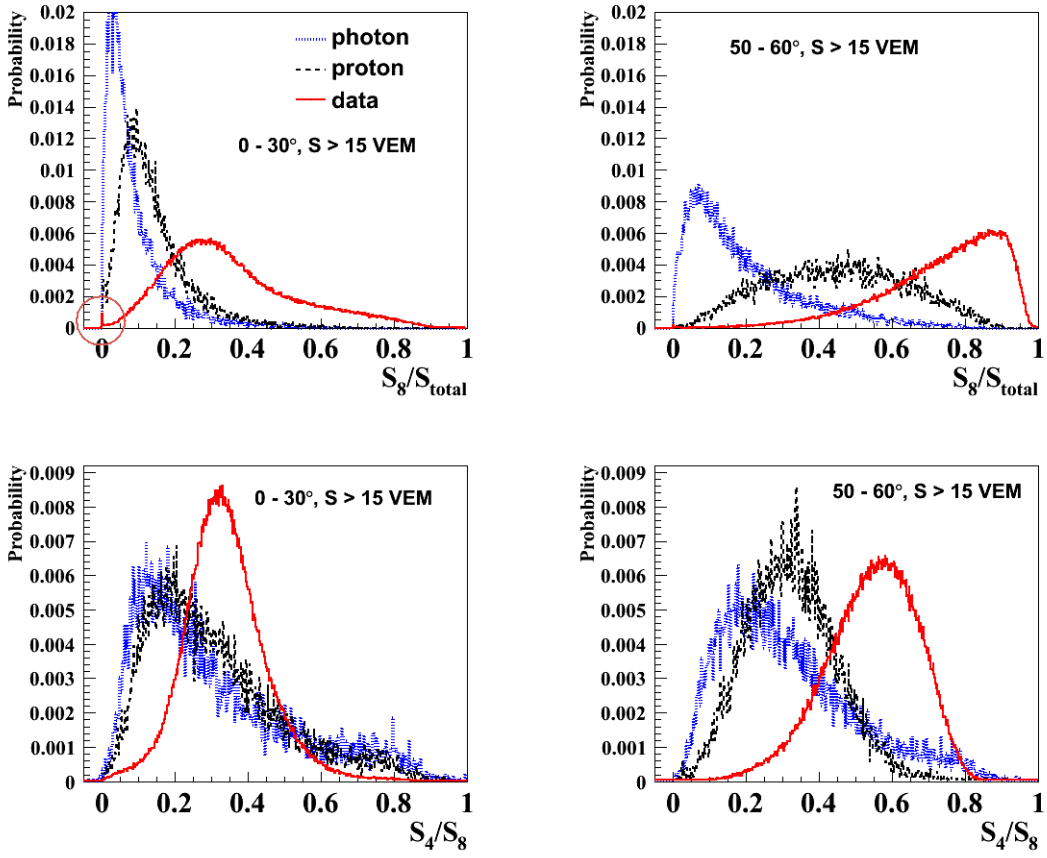


Figure 6.3-4 Distributions of S_4/S_8 and S_8/S for proton, photon and a subset of data at different zenith angles

Values of S_8/S for data is much larger than that from photon simulations, which means by cutting $S_8/S < 80\%$, only data are greatly affected and by doing this the 'extremely fast' traces are filtered. A similar reason is applied for the cut $S_4/S_8 < 70\%$. It should be noted that there is a small peak from the distribution of S_8/S of data at vertical zenith angle. This peak is due to the start time problems as shown in this section. By removing events with $S_8/S < 2\%$, the start time problem is suppressed but not solved. All the three cuts

are applied for data and photon simulations that are used for photon-hadron separations in section 6.2. The cuts should be removed once the problems of the reconstructions are solved.

The next section is about the photon-like events found from data that are with Δ-entity likelihood ratio less than -0.15.

6.4 Photon-like events and categories

There are 103 and 13 photon-like events found for $10^{18.8} < E^{\gamma}_{rec} < 10^{19.0}$ eV and $E^{\gamma}_{rec} > 10^{19.0}$ eV respectively. In this section we look into these events and classify them into four categories. In particular we focus on the 13 highest energy events and all FADC traces are shown for those events.

Table 6.4-1 Photon-like events that have $E^{\gamma}_{rec} > 10^{19}$ eV from data

no.	tank	Event ID	θ	$E_{rec}{}^{\gamma}$	$E_{rec}{}^{p}$	ΔLLR	r	S	S_{LDF}	$\log_{10}(\chi^2/ndf)$	LLR (detector)	cause
1	1	4586998	40.0	19.37	18.92	-0.25	590	251	243	1.8	-2.8	
	2						617	218	211	1.4	-1.0	
	3						1154	28	26	2.9	-4.3	Pre-pulse
2	1	4841436	34.0	19.06	18.54	-0.54	766	49	49	0.7	1.9	
	2						817	38	39	3.2	-7.1	Pre-pulse
3	1	6253275	42.7	19.04	18.64	-1.13	816	42	41	3.8	-7.9	Start time
	2						1082	20	16	1.1	0.4	
4	1	7250946	26.6	19.25	18.62	-0.22	494	278	277	0.9	1.4	
	2						1029	23	22	3.4	-6.4	PMT LG
	3						1041	22	21	2.2	-3.0	
5	1	7543164	53.8	19.32	18.93	-0.21	703	90	79	1.9	-2.8	Slow-rise
	2						1170	24	18	2.2	-2.5	
6	1	9110108	41.3	19.01	18.59	-0.17	752	51	50	1.5	-1.2	Slow-rise
	2						756	49	49	1.3	-0.5	
	3						800	42	41	2.2	-3.7	
7	1	10611670	31.9	19.03	18.49	-0.19	611	96	95	1.4	-0.9	
	2						872	29	29	3.0	-6.1	Pre-pulse
	3						971	21	20	0.8	1.3	
8	1	11865575	30.5	19.14	18.57	-1.45	501	224	222	1.7	-2.1	
	2						866	38	36	3.1	-6.7	Pre-pulse
9	1	13516784	30.6	19.09	18.52	-0.46	777	47	47	2.6	-5.3	Pre-pulse
	2						808	41	41	1.0	0.8	
	3						862	33	33	2.0	-2.7	Pre-pulse
10	1	15517241	27.0	19.09	18.48	-0.65	621	95	95	0.6	2.5	
	2						879	30	29	3.6	-8.1	Start time
11	1	15978235	22.7	19.05	18.40	-0.44	656	69	68	2.1	-3.8	Pre-pulse
	2						921	22	21	2.5	-4.2	Pre-pulse
	3						931	20	20	1.0	0.8	
12	1	22418492	40.0	19.09	18.65	-0.18	508	206	203	1.1	0.6	
	2						948	32	28	3.0	-5.8	Pre-pulse
	3						1001	21	23	1.3	-0.3	
13	1	26206054	60.0	19.12	18.51	-0.39	424	55	62	4.9	-5.9	PMT D/A
	2						925	19	10	0.8	1.4	

Some of the important quantities are listed including the event id, zenith angle, energy in the photon scale, energy in the hadronic scale, the final parameter Δ-entity likelihood ratio, core distance, total signal of the station, total signal expected from the LDF fitting, $\log_{10}\chi^2$/ndf from fitting the time-trace (Figure 5.3-1), the likelihood ratio of the station (Figure 5.3-9) and the cause for the event that has ΔLLR < -0.15. The colours in the table are assigned according to the reason that causes the event identified as a photon-like event. The four categories are: wrong start times of the station (blue), problems with PMT (red), pre-pulses at the beginning of the trace (yellow) and otherwise just traces with slow-rise (green).

6.4.1 Category I: Wrongly determined start times

The first category is about events that have start time wrongly determined. A flat piece is normally found at the beginning of the trace. These events are caused by fluctuations of baselines of PMTs and have been discussed in section 6.3. Two events are found with such signature and shown in Figure 6.4-1 and Figure 6.4-2.

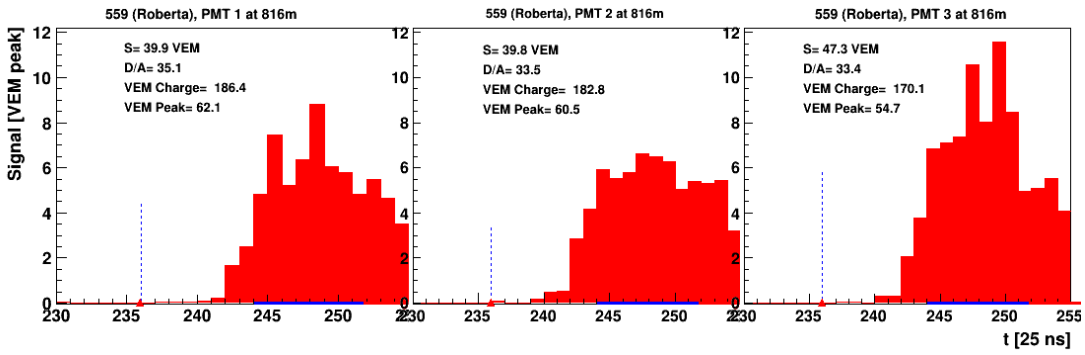


Figure 6.4-1 Event 6253275 – flat piece at the beginning of the trace

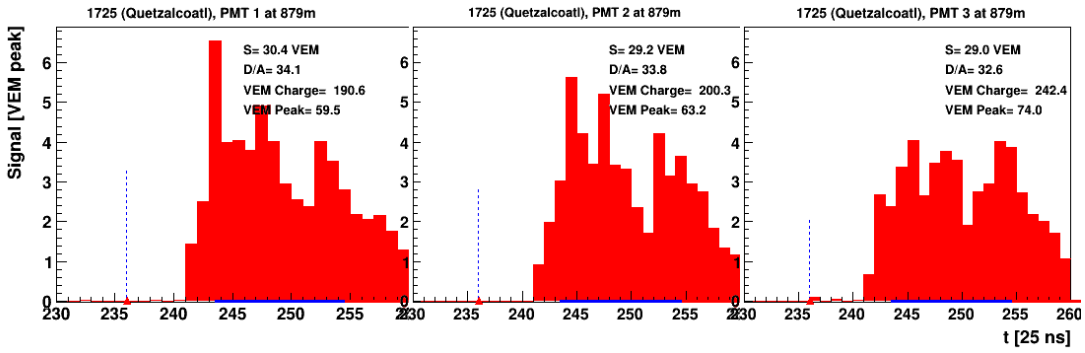


Figure 6.4-2 Event 15517241 – flat piece at the beginning of the trace

$\text{Log}_{10}\chi^2$/ndf of the two stations are both larger than 3.5, while other stations in the event are consistent with data-like traces. It is decided to remove these

two events from the candidate list since it is clear they are not photons but due to PMTs with highly fluctuated baselines.

6.4.2 Category II: PMT problems

The PMT qualities are monitored by observing the trigger rate, high-gain / low-gain ratio and many other variables, which excludes most of the PMT problems from the data. However very rarely there are bad PMTs that are not identified by the standard monitoring procedure. Since the entity method is very sensitive to exotic traces, some remaining PMT problems are the reasons for few outliers of ΔLLR.

PMT1 of station 201 from the event 26206054 (Figure 6.4-3) has dynode / anode ratio ~ 175, which is ~ 5 times higher than the averaged value. The problem is found as the anode was having large noises at the time of the event.

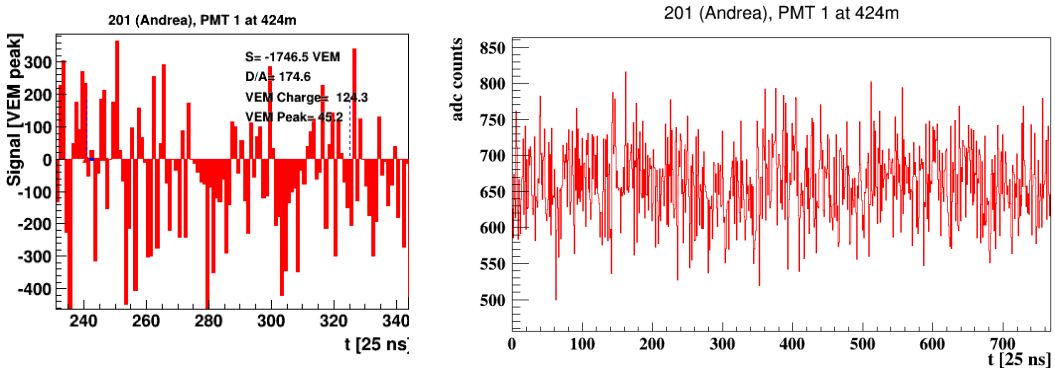


Figure 6.4-3 PMT 1 of station 201 – problems of low-gain (anode) channel of the PMT. On the left is the FADC trace and on the right is the baseline from the anode. Event 26206054 is removed from the photon-candidate list.

Event 7250946 was also selected due to problems with the anode. Figure 6.4-4 shows the FADC trace of three PMTs of station 157. The high-gain channel (dynode, Figure 6.4-5) was saturated and therefore the low-gain channel (anode, Figure 6.4-6) was used to determine the signal.

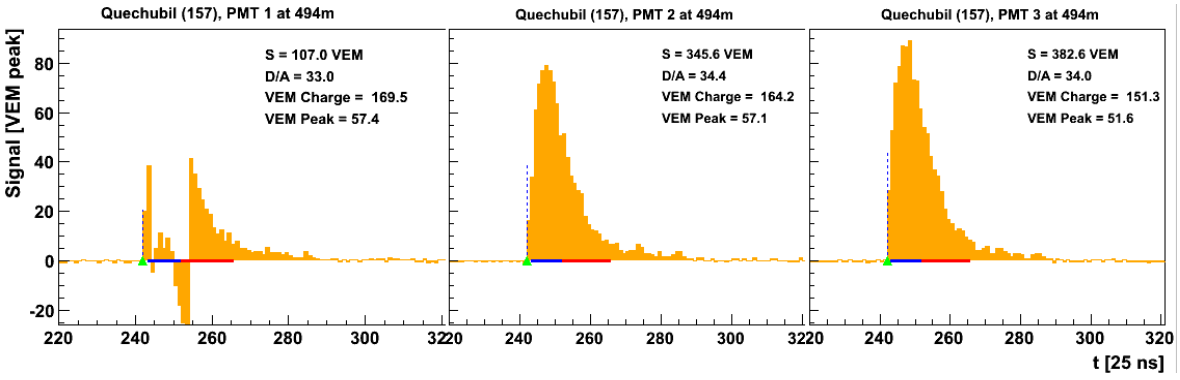


Figure 6.4-4 FADC trace of three PMTs of station 157 from event 7250946

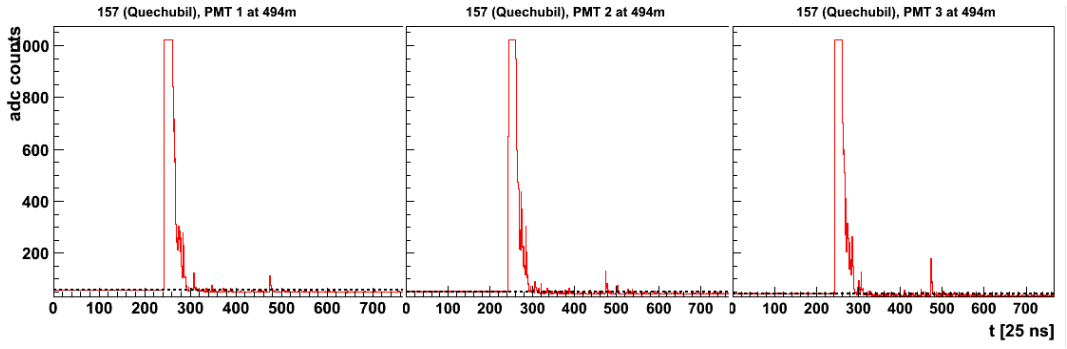


Figure 6.4-5 Dynode (high-gain) of station 157 from event 7250946

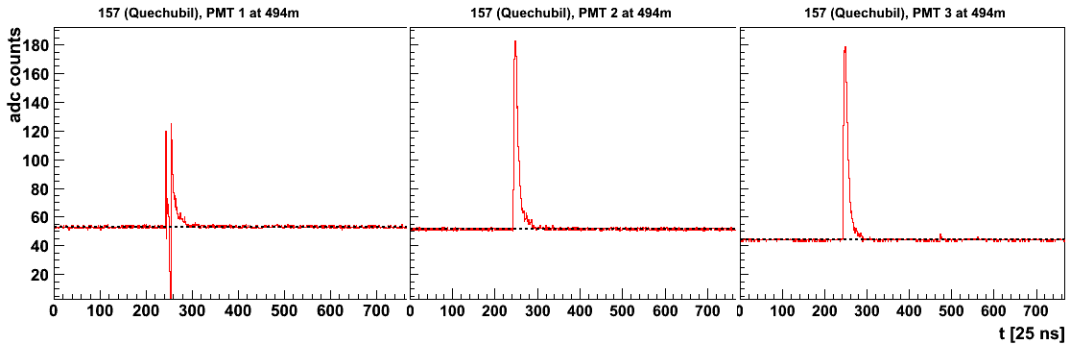


Figure 6.4-6 Anode (low-gain) of station 157 from event 7250946. It is clear that PMT 1 is not stable. This event is excluded from the photon candidate event list.

The first two categories are related to qualities of data and reconstructions. They have traces that neither data-like nor photon-like. However, it should be noted that the entity method has great potential of finding problems of PMTs. In the future one could try to optimise the entity method for the monitoring of qualities of PMTs.

6.4.3 Category III: Pre-pulses

Pre-pulses are used to describe a FADC trace with a small pulse at the beginning. To illustrate the signature, event 15978235 is used as an example and shown in Figure 6.4-7. It was vertical (θ = 23˚) and three stations were triggered.

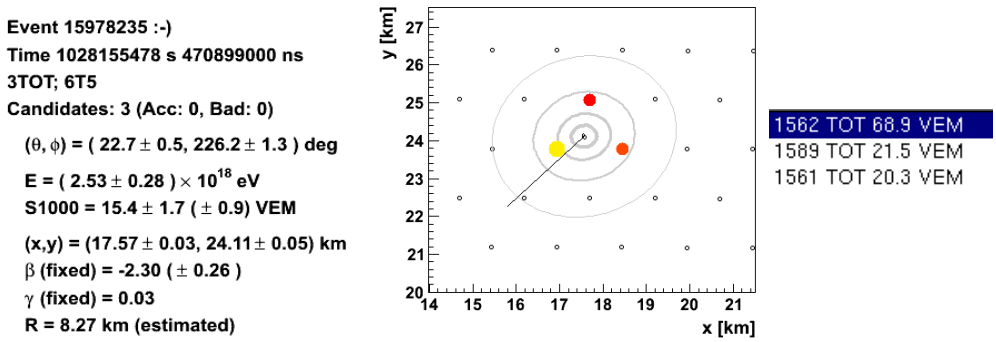


Figure 6.4-7 Example event (15978235, vertical) have all three stations with pre-pulses. The station marked in yellow is the first triggered and with the largest signal, the second triggered station is marked in orange with S = 21.5 VEM.

The FADC traces in each station are shown in Figure 6.4-8, Figure 6.4-9 and Figure 6.4-10. Signals on the y-axis are in units of VEM peak.

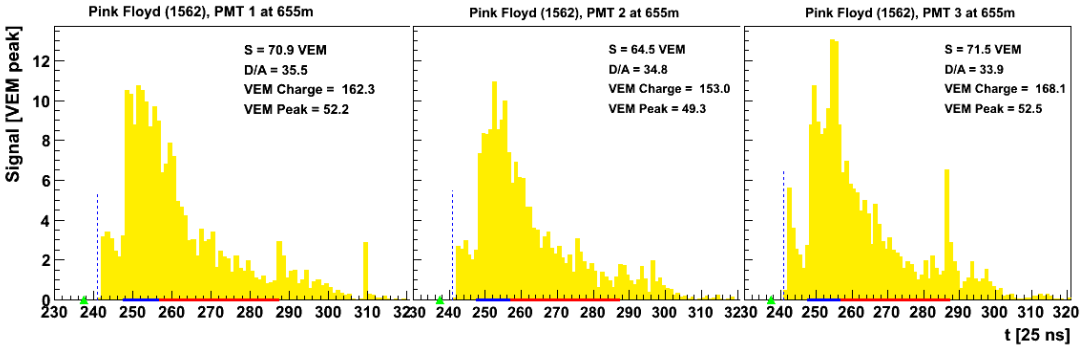


Figure 6.4-8 FADC trace of station 1562, $\log_{10}\chi^2/\text{ndf} = 2.11$

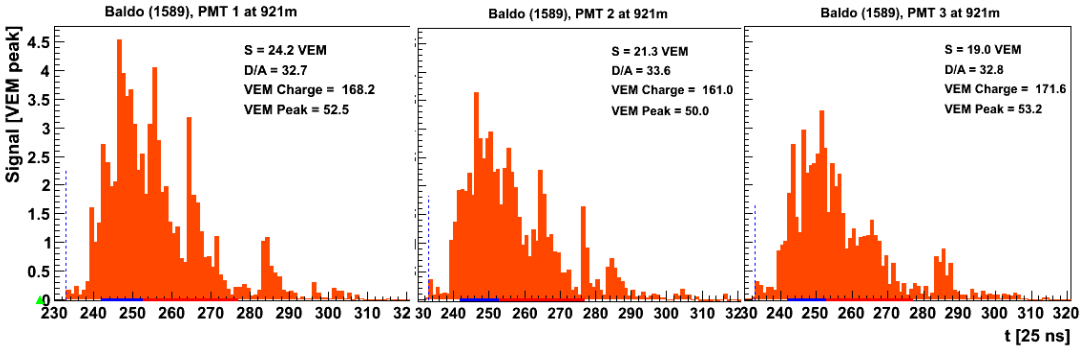


Figure 6.4-9 FADC trace of station 1589, $\log_{10}\chi^2/\text{ndf} = 2.48$

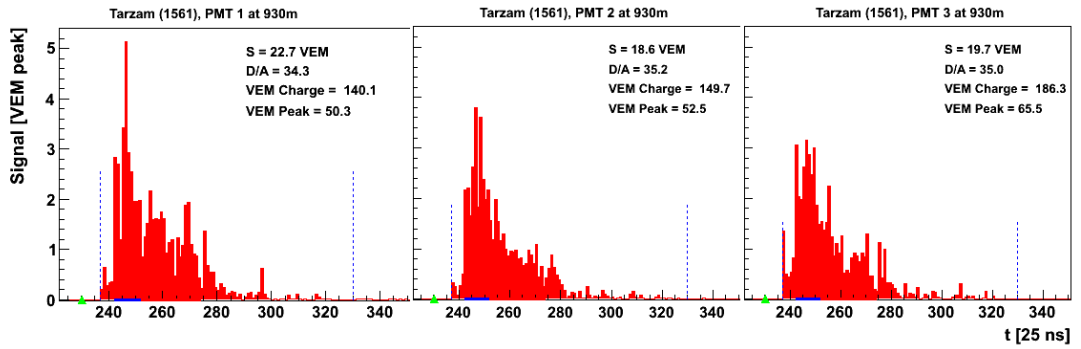


Figure 6.4-10 FADC trace of station 1562, $\log_{10}\chi^2/\text{ndf} = 0.96$

All three stations have small bumps at the beginning of the main part of the trace. This feature is named 'pre-pulse'. For station 1562 the pre-pulse lasts ~ 200 ns. Stations 1589 and 1561 both have pre-pulses of ~ 75 ns.

There are 7 events categorised as pre-pulses for $E^{Y}_{rec} > 10^{19}$ eV. Traces are compared to the mean of data (red) and photon (blue). Also compared is the risetime. A benchmark of risetime is drawn on the right top corner of each plot. It is a function of risetime with respect to distance and obtained from data. Photons are likely to have risetime above the benchmark.

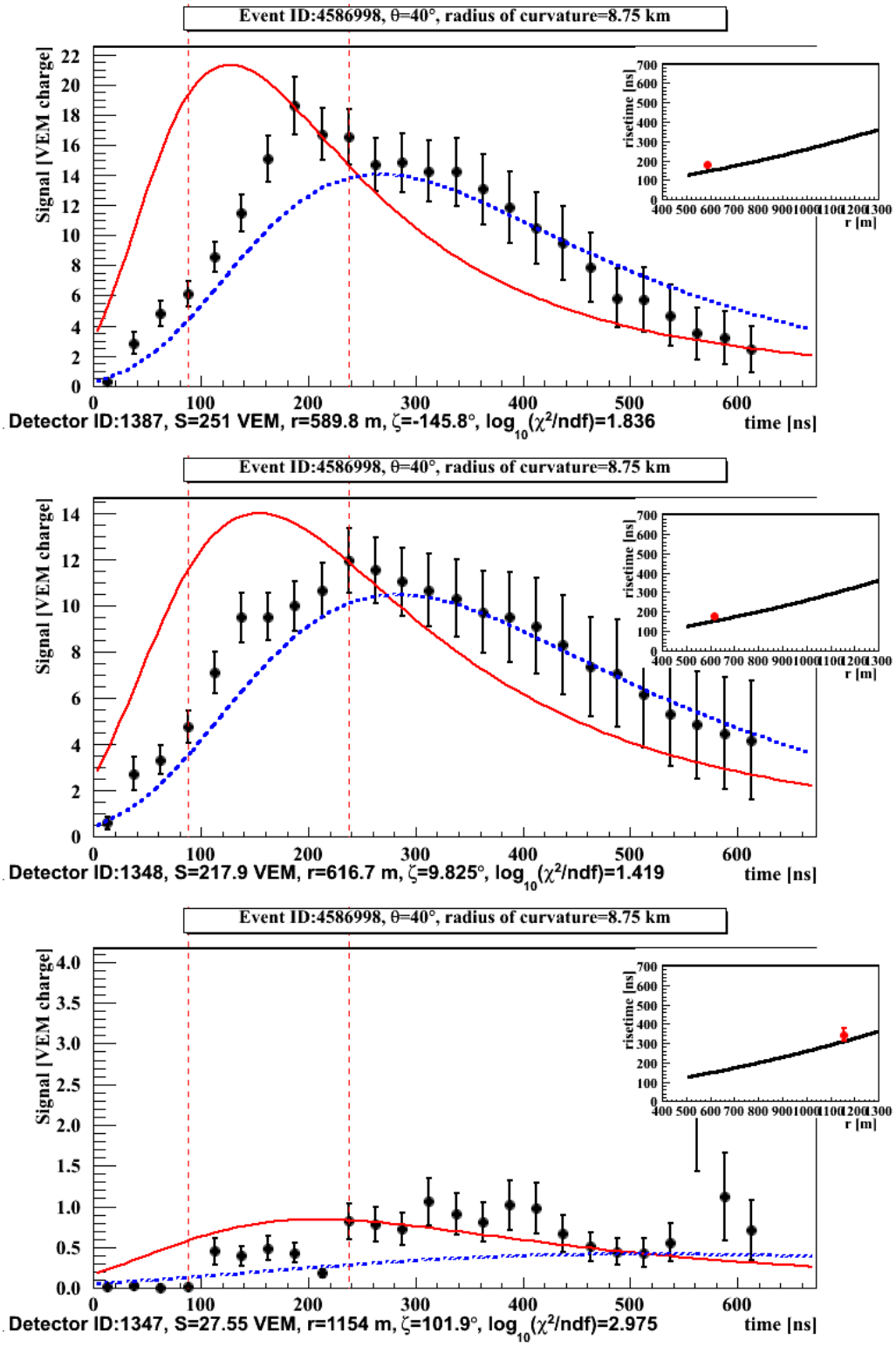


Figure 6.4-11 FADC traces of station 1378 (a), 1348 (b) and 1347 (c) of event 4586998. Only station 1347 has pre-pulse. The other two stations have slow trace and are photon-like. The red curve is the parameterised mean trace using data. The blue dashed curve is the mean trace from photon simulations and only used for illustration. The dashed interval in red is the time period used for the calculation of $\log_{10}\chi^2$/ndf. Detector signal, core distance, polar angle and $\log_{10}\chi^2$/ndf are shown below each plot.

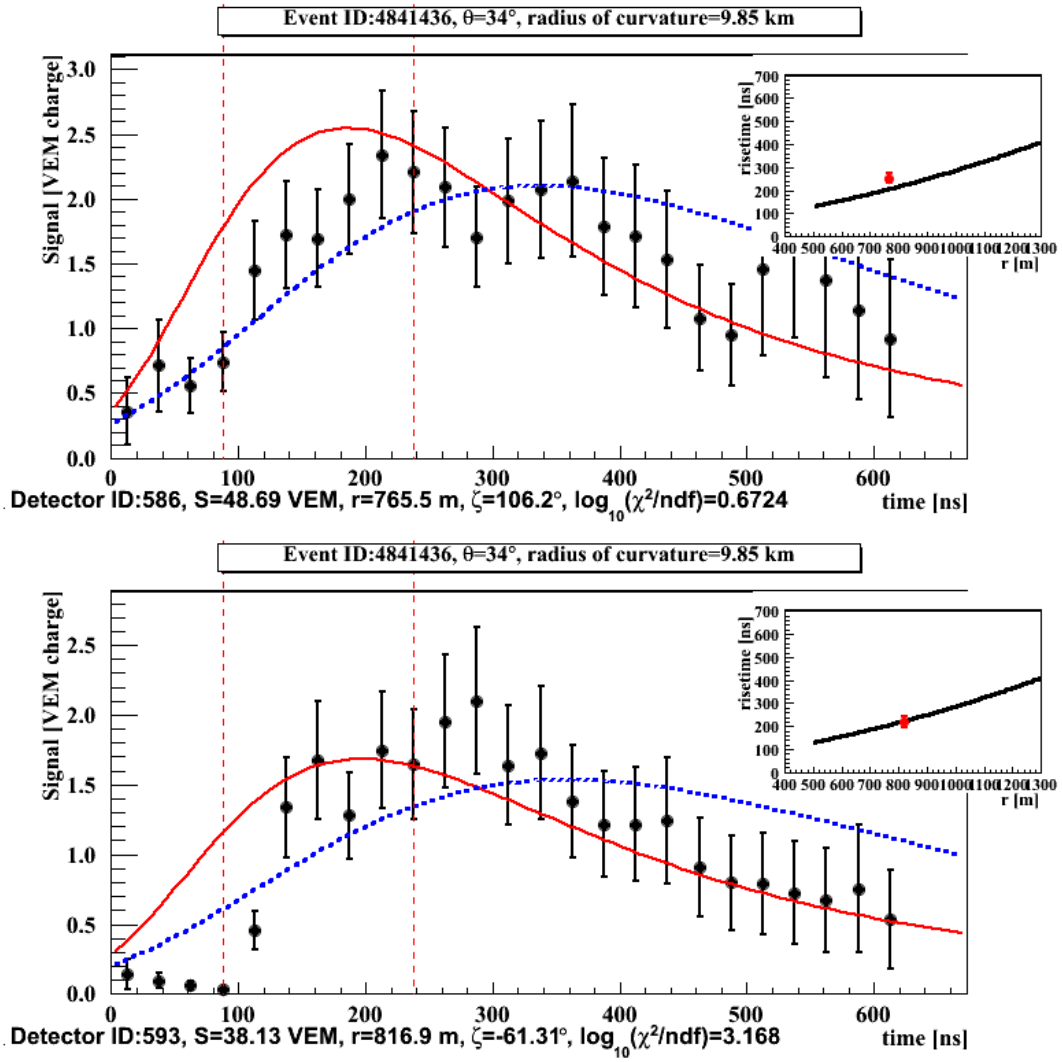


Figure 6.4-12 FADC time-traces of the station 586 (a) and 593 (b) of the event 4841436. On the right top corners are plots about risetime and the black line is the mean from data (also called the benchmark). The red point is the measured risetime of the station. Photons are expected to have larger risetime than data. Only station 593 has the pre-pulse feature and the risetime is consistent with the expectation from data. Station 586 has slow trace and is photon-like. The red curve is the parameterised mean trace using data. The blue dashed curve is the mean trace from photon simulations and only used for illustration. The dashed interval in red is the time period used for the calculation of $\log_{10}\chi^2$/ndf. Detector signal, core distance, polar angle and $\log_{10}\chi^2$/ndf are shown below each plot.

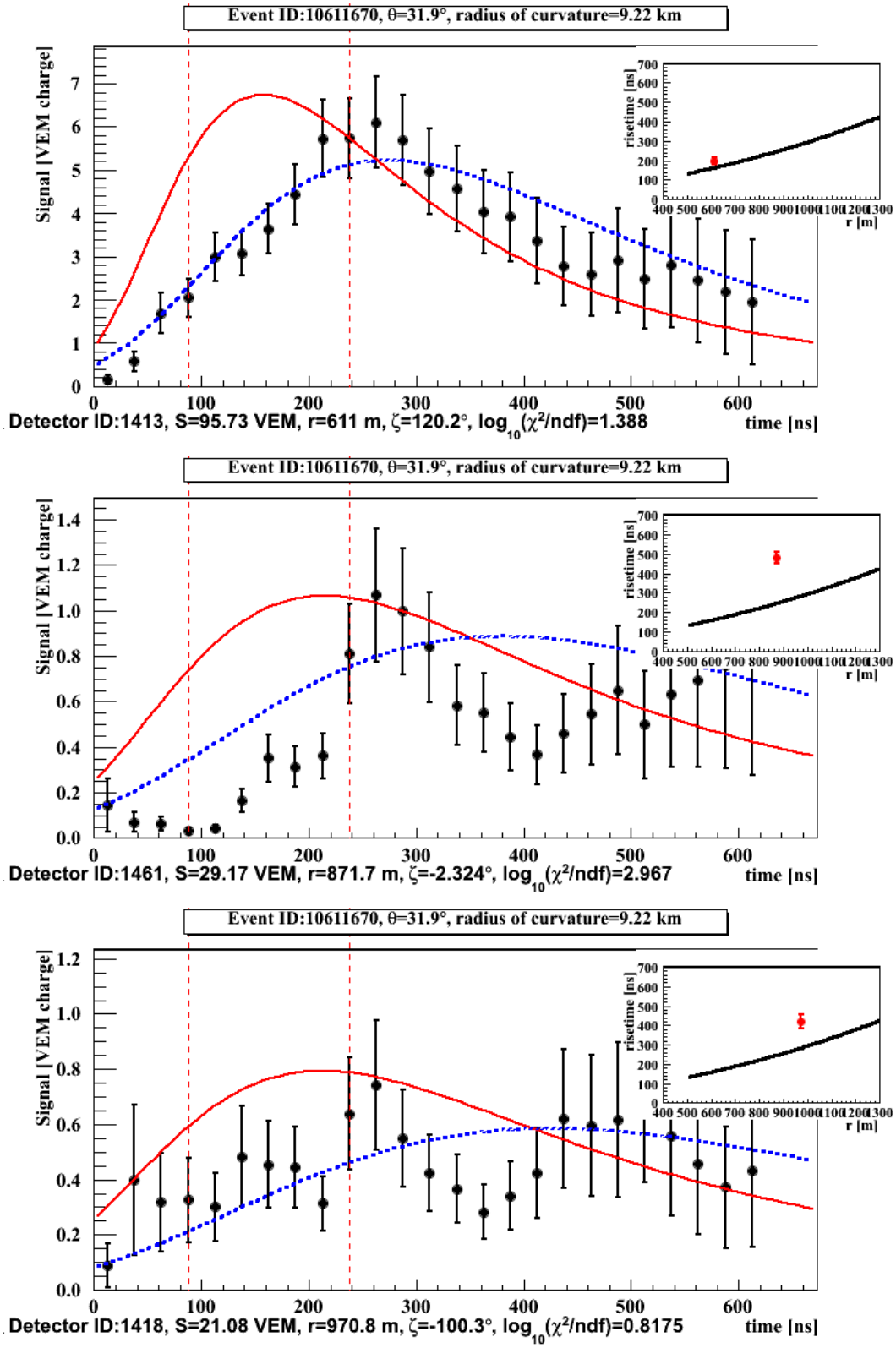


Figure 6.4-13 Time- traces of station 1413 (a), 1461 (b) and 1418 (c) of event 10611670. Only station 1461 is with pre-pulse. The other two stations are with trace slowly rising. All risetime values are larger than the benchmark from data.

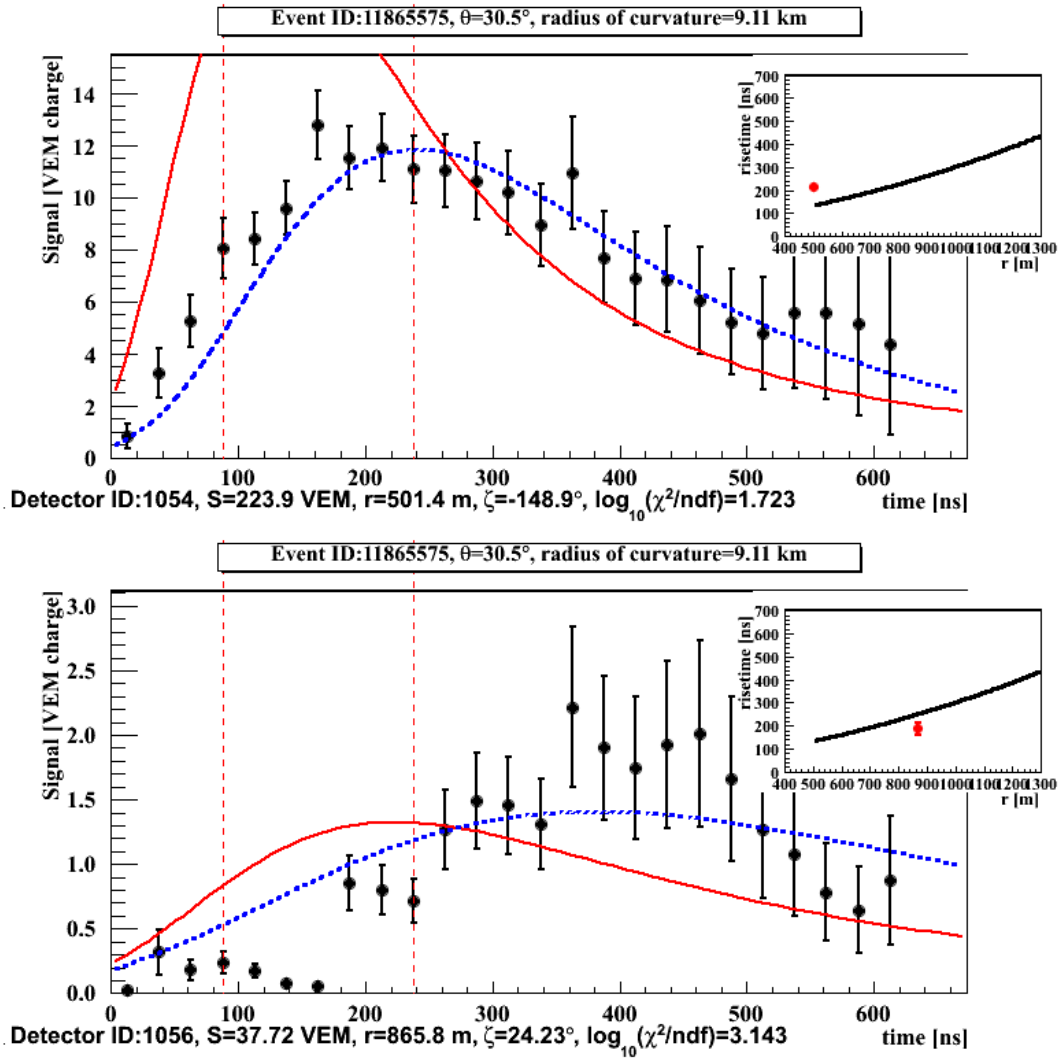


Figure 6.4-14 FADC time-traces of station 1054 (a) and 1056 (b) of the event 11865575. Only detector 1056 is with pre-pulse. Station 1054 has trace rises slowly (large risetime) and is photon-like.

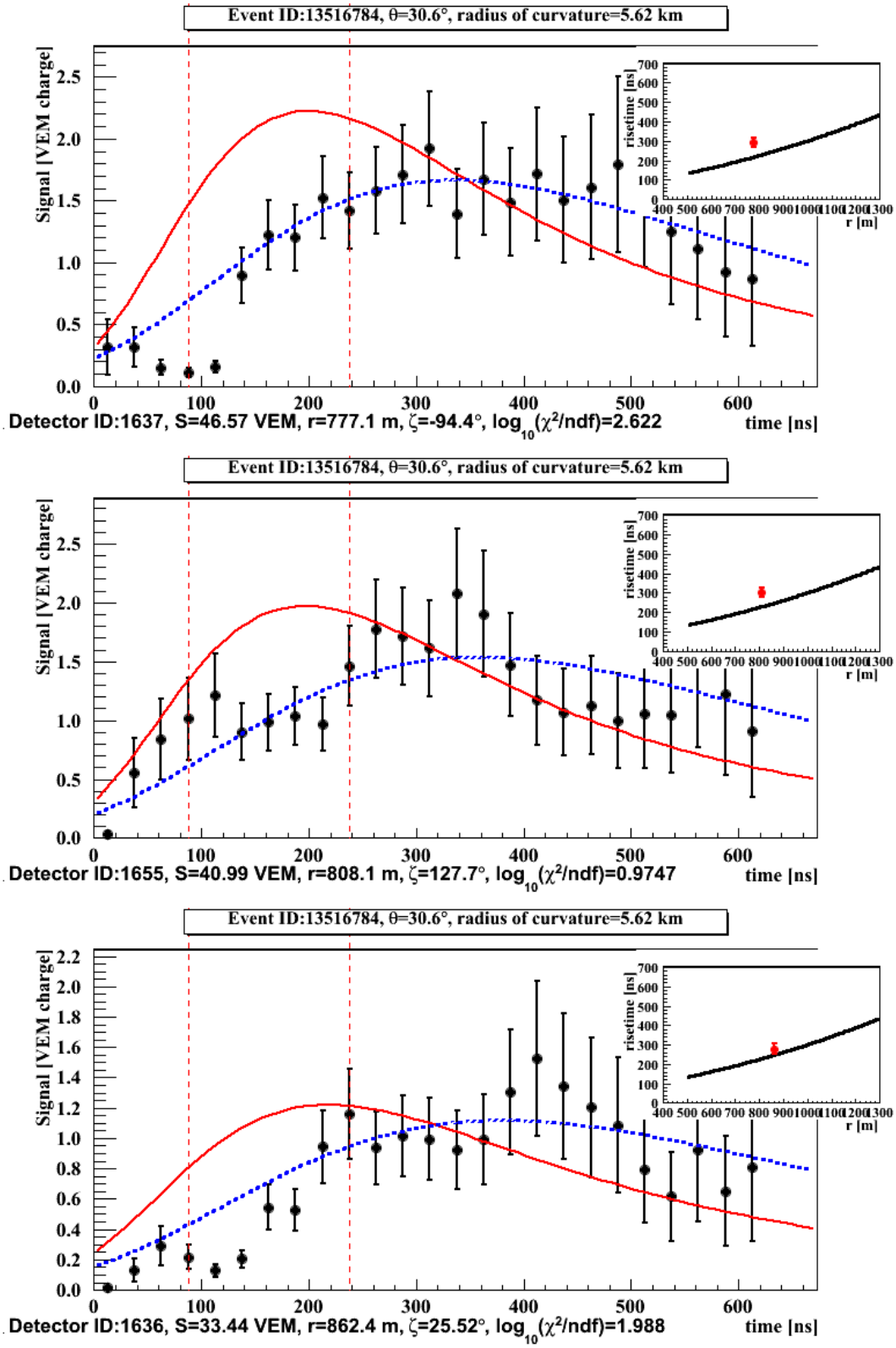


Figure 6.4-15 Time- traces of station 1637 (a), 1655 (b) and 1636 (c) of event 13516784. All three stations have pre-pulses and risetime above the benchmark.

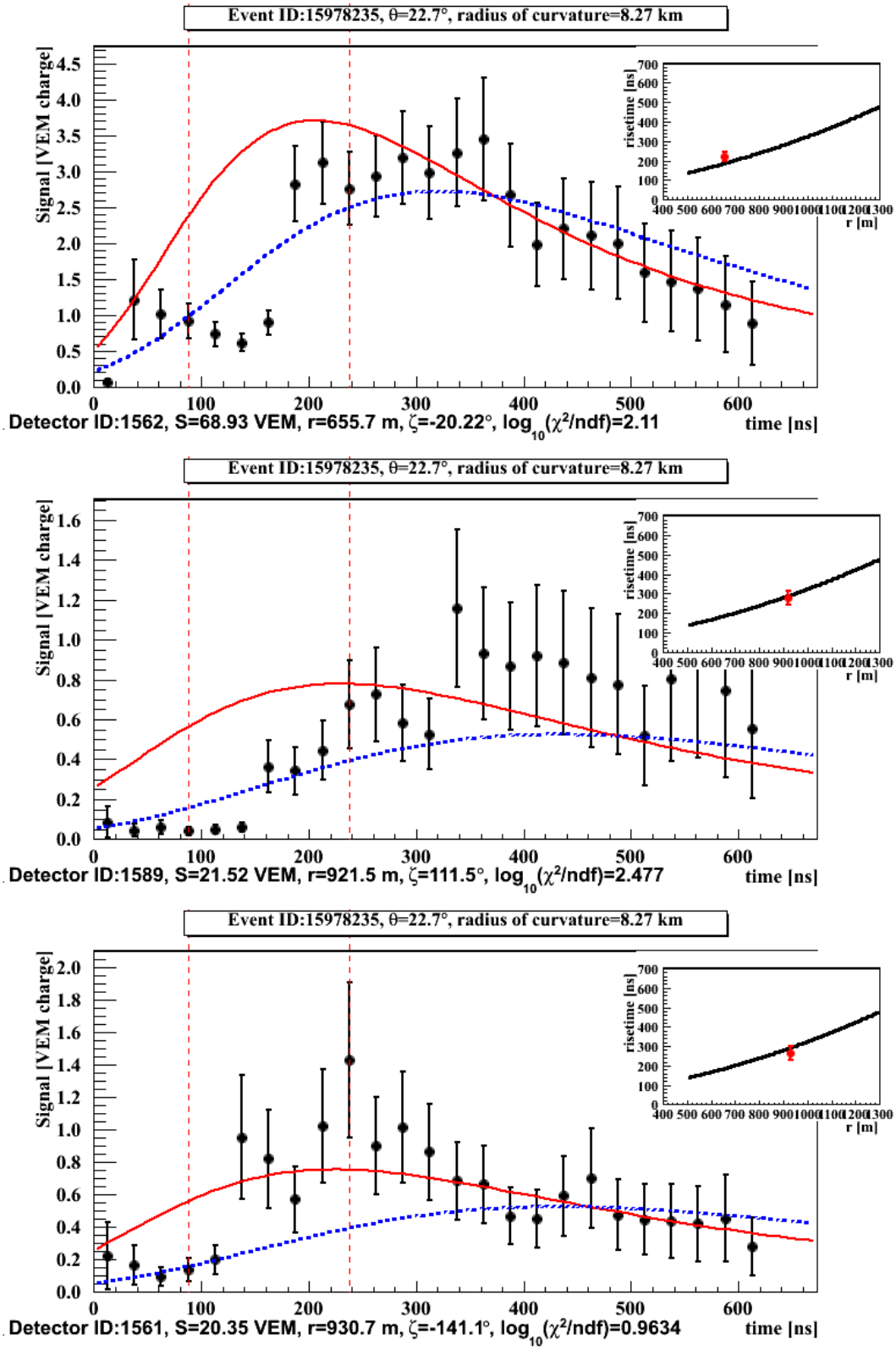


Figure 6.4-16 Time- traces of station 1562 (a), 1589 (b) and 1561 (c) of event 15978235. This is the event that was used as an example in Figure 6.4-7. It should be noted that traces with pre-pulses do not necessarily have larger risetime. It is important to understand the physics behind the pre-pulse to decide if that is it occurs for photon showers.

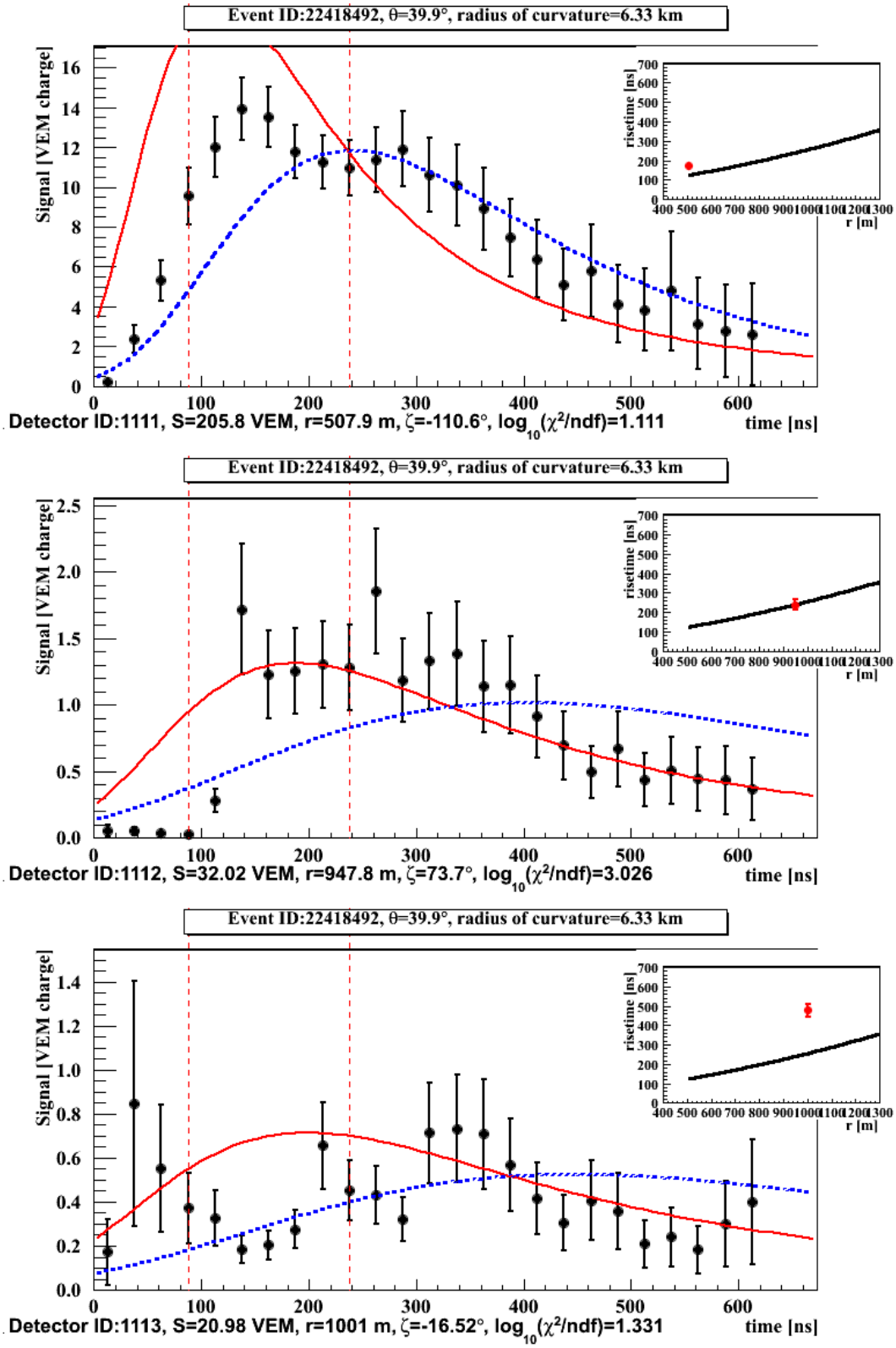


Figure 6.4-17 Time- traces of station 1111 (a), 1112 (b) and 1113 (c) of event 22418492. Station 1111 has the FADC trace between the mean of data and the mean of photon simulations. Station 1112 has the pre-pulse feature. Station 1113 has a trace slowly-developed and risetime much larger than the expectation value from data.

To summarise, the pre-pulse category has at least one station with a small pulse at the beginning of trace. They are kept as photon candidates as they are seen in photon simulations. Details are introduced in section 6.5.2.

6.4.4 Category IV: Slow traces

The last category is events that have no problems of PMTs or pre-pulses but only with slowly developed FADC traces. This is the most interesting category and includes two events in Table 6.4-1.

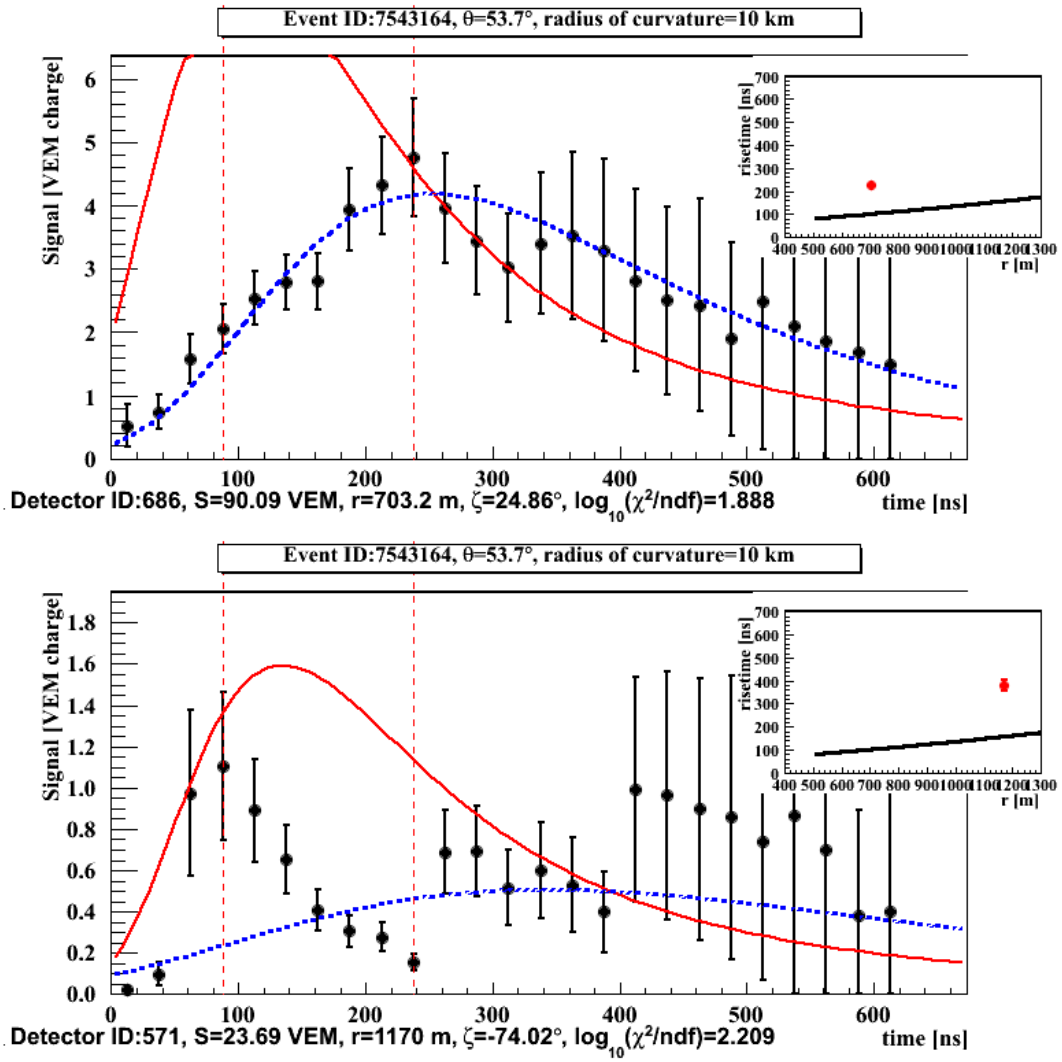


Figure 6.4-18 FADC time-traces of station 686 (a) and 571 (b) from the event 7543164. Both stations have risetime much higher than expectations from data. Trace from station 686 is consistent with the mean of photon. The trace from station 571 is peaky but develops slowly. Station 571 is not identified as the pre-pulse because the signal deposit at the first few time bins is too high.

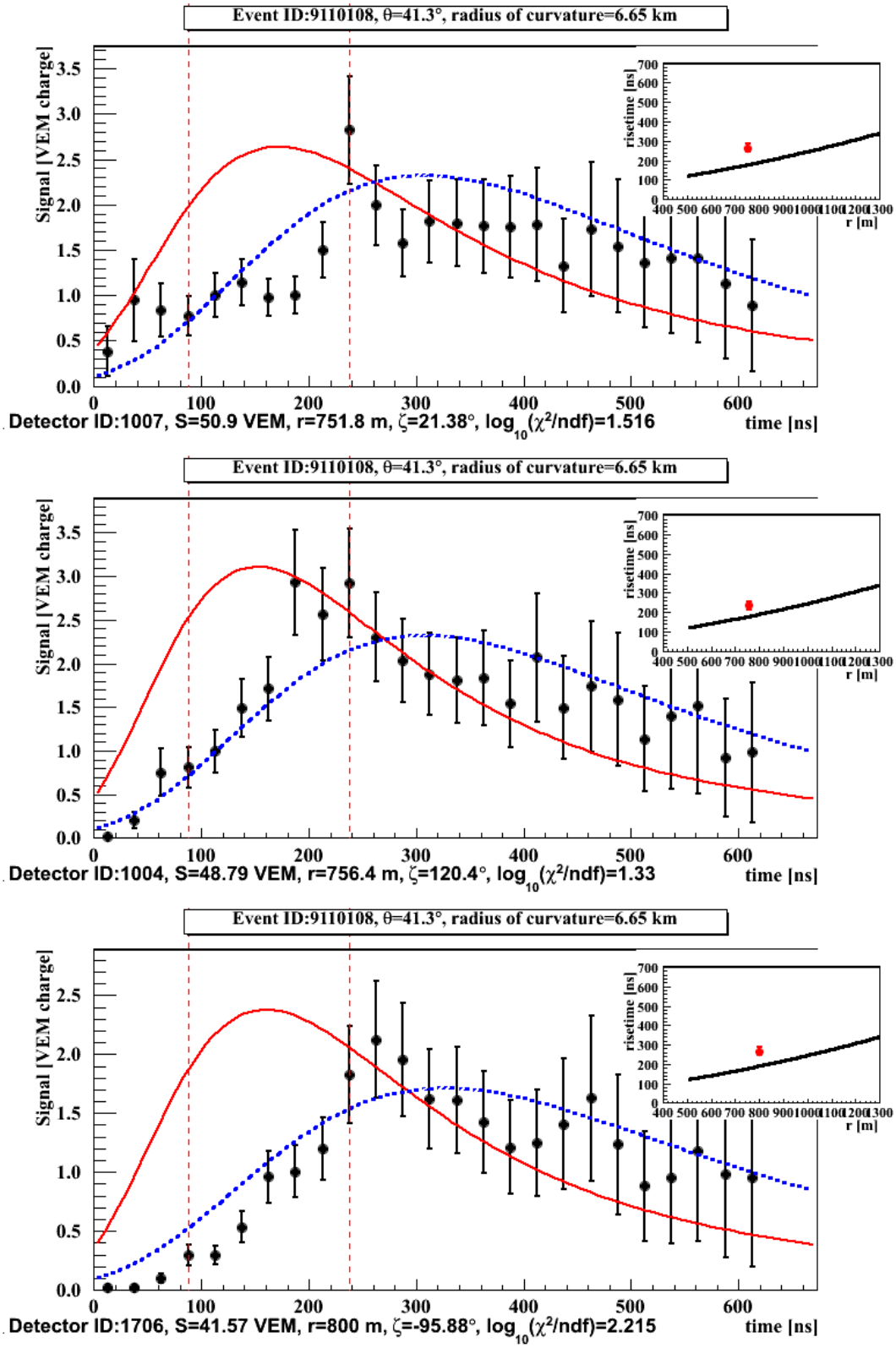


Figure 6.4-19 Time- traces of station 1007 (a), 1004 (b) and 1706 (c) of event 9110108. This is the most interesting event found in the thesis. All three traces show agreement of being photon-like. All three risetime are above the benchmark from data. The radius of curvature is significantly smaller than the expectations from data (11 km), which is another photon feature.

Events in this category cannot be excluded as photons candidates although they might be from hadronic background of light nuclei (Chapter 8).

For showers with $10^{18.8} < E^{\gamma}_{rec} < 10^{19.0}$ eV, all photon-like events can also be summarised into the four discussed categories. There are 31 photon-like events identified with problems of PMT / start time and 72 events with pre-pulse / slow trace. The events are saved at (Auger password 2014)

http://www.ast.leeds.ac.uk/~phyll/Thesis/Candidate_18.8-19.0.root

The entity method is seen to be very sensitive to traces with pre-pulses. It is important to understand if the pre-bumps are from the shower or from the background muons. The physics behind could lead to some interesting results.

6.5 Discussions of pre-pulses

As shown in the section 6.4.3, pre-pulses in the traces are frequently found in the FADC traces associated with photon-candidates. They could come from the shower itself, or are just signals from random muons in the atmosphere. In this section, we first explore the possibility of the effect being due to accidental muons. Then by making use of Monte Carlo simulations, we explore if the pre-pulses could be caused by fluctuations from sampling a small number of particles in proton, iron and photon showers. The idea that the pre-pulses are from EM halos is also introduced.

6.5.1 Accidental muons

Each surface detector on the ground is being hit by muons from the atmosphere with frequency of ~ 2.5 kHz. A simple signal threshold has been set for the first level trigger (Thr1, Figure 2.1-4) reduces the trigger threshold to ~ 100 Hz. The energy loss due to ionisation for a muon passing through the detector vertically is ~ 240 MeV and is used for the signal calibration of 1 VEM (vertical equivalent muon). The pre-pulses found in the candidate events in principle could be coincidences due to atmospheric muons that arrive just before the start of the trace. In such case the start time of the station would be estimated incorrectly. In the first part of this section, four types of accidental signals are shown and then the frequency and amplitude of the accidental signals are summarised. Data are used for the estimation of the rate for having such accidental signals at the start of the trace.

6.5.1.1 Examples of accidental signals found in data

The Central Data Acquisition System (CDAS) records signals from each PMT ~ 6 μs before the trigger of the event for the calculation of the baseline of traces. This time period is called pre-trace and is used in this section for estimating accidental signals.

The accidental signals are identified using a simple algorithm. A sliding-window with a fixed width of 100 ns (4 bins) goes through the trace in steps of 25 ns. An accidental signal is flagged if the total signal within the 100 ns is larger than 0.2 VEM (charge) and the signal in each bin is larger than 0. The overlaps of segments of accidental signals are merged to form the combined segments. Very rarely more than one merged segment exists (1% of the cases that there is only one merged segment) before the start of the trace and such events are excluded in this exercise.

A single muon going through the detector via different path lengths are shown in Figure 6.5-1 and Figure 6.5-3. They are both located ~ 150 bins (3600 ns) away from the majority of the trace. The accidental signals are then zoomed in and shown in Figure 6.5-2 and Figure 6.5-4 accordingly.

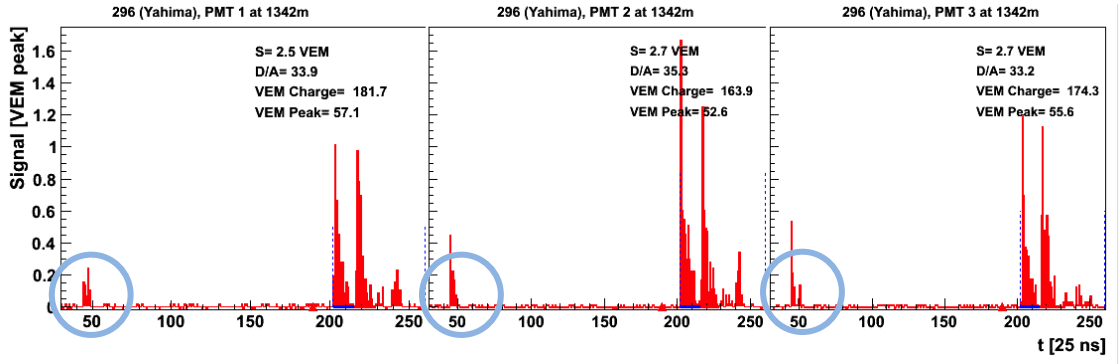


Figure 6.5-1 An event with a station of accidental signal (Single muon travels through the edge of the detector and the merged segment is indicated in the blue circle)

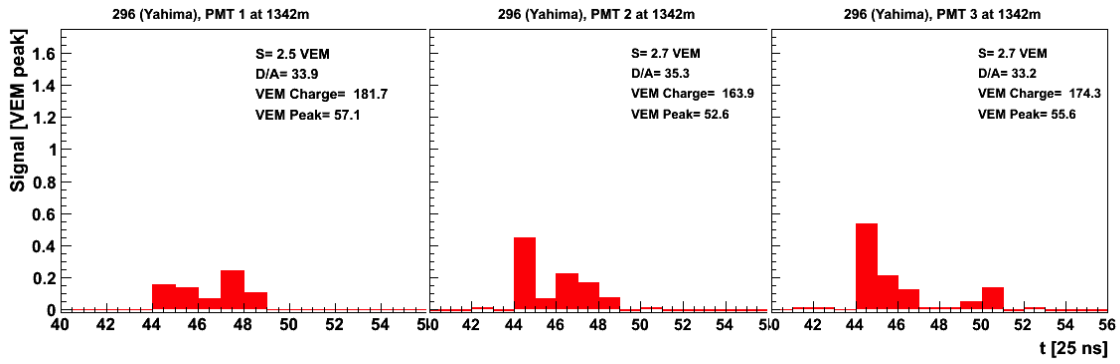


Figure 6.5-2 Zoom-in of the accidental signal

The signals drawn on the plots are with unit VEM peak, which is ~ 3 times the size of VEM charge. 1 VEM charge is the signal calibrated from one vertical muon. As stated in Chapter 2, 'VEM' is always referred to VEM charge in this thesis. The first event has the accidental signal ~ 0.5 VEM and the second event has the accidental signal of ~ 1 VEM. Both signals have the time interval of ~ 125 ns. If any of the two cases happen at the beginning of the trace (~ 75 ns ahead of the true start time), they will be identified as pre-pulses.

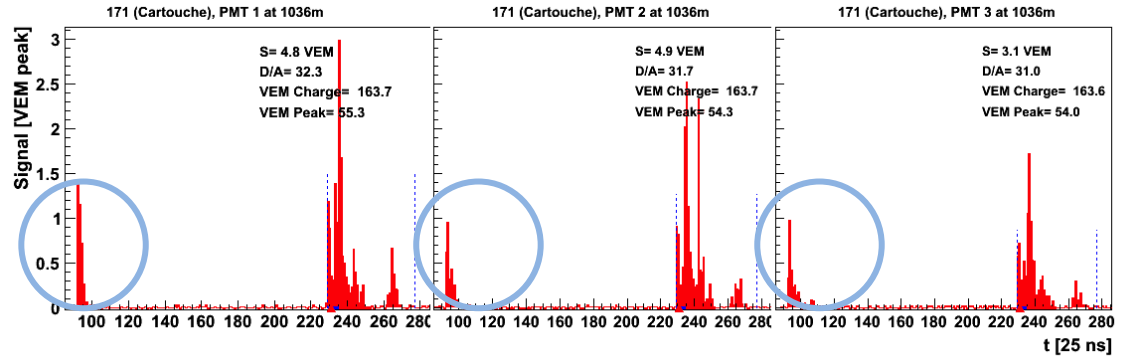


Figure 6.5-3 An event with a station of accidental signal (vertical muon) (The accidental signal is ~ 1 VEM and is indicated in the blue circle)

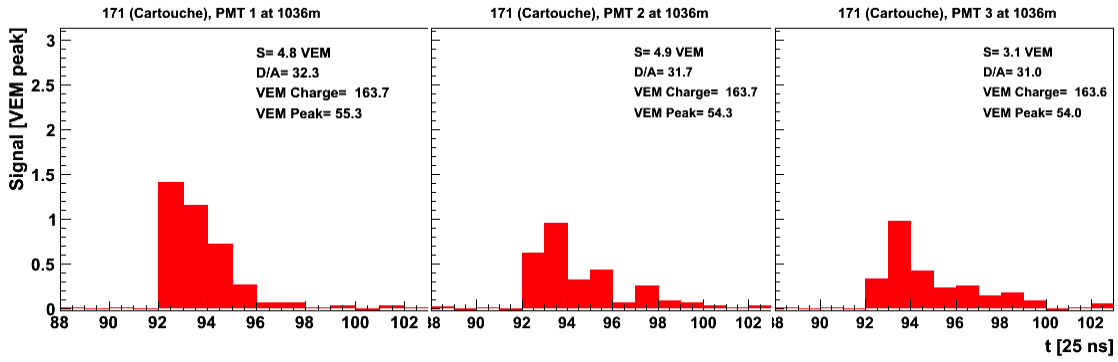


Figure 6.5-4 Zoom-in of the accidental signal (vertical muon)

The next two cases are less frequent. It was discovered that it is possible to have accidental signals with large signal sizes and large time spreads. Shown are two examples in Figure 6.5-5 and Figure 6.5-7. The first event has the accidental signal covering the time period of 350 ns and signal 10 VEM, the second event has the time period 420 ns and signal size 2 VEM (it has two peaks).

If there is a small shower coming within few μs of the current event, the signals might be recorded by CDAS. It is possible to select a wrong piece of the trace as the signal for the current event and have incorrect reconstructions of the event.

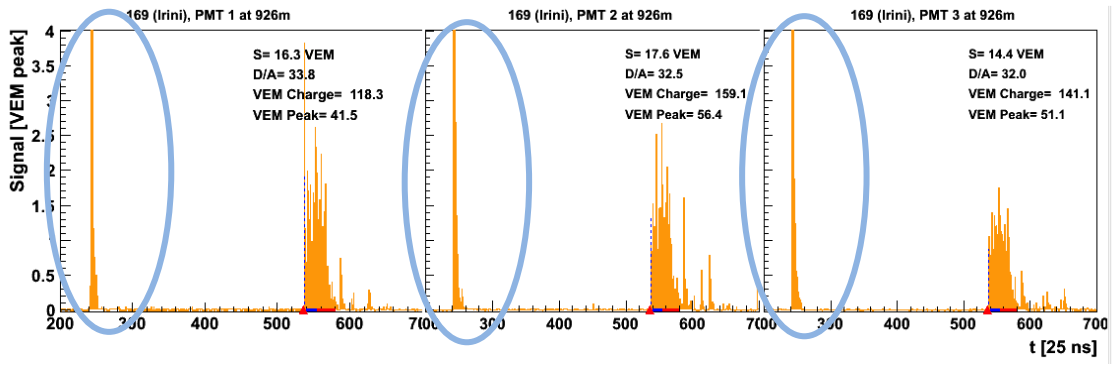


Figure 6.5-5 An event with a station of accidental signal (~ 10 VEM, 350 ns, circled in blue)

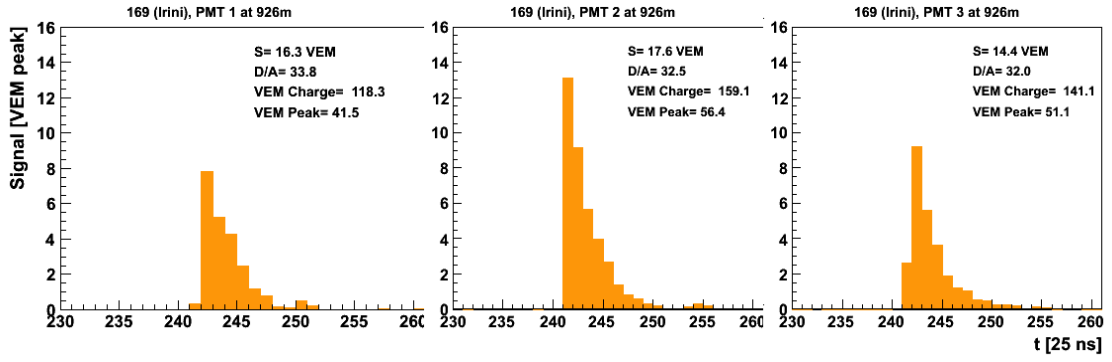


Figure 6.5-6 Zoom-in of the accidental signal (~ 10 VEM, 350 ns, risetime ~ 40 ns)

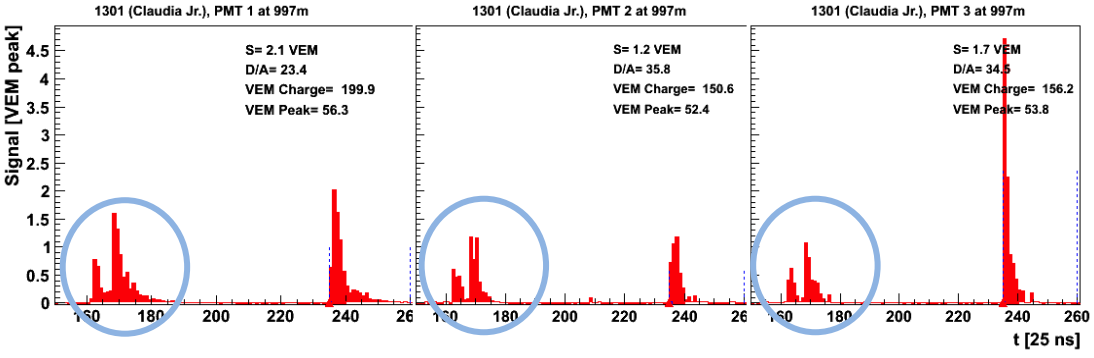


Figure 6.5-7 An event with a station of accidental signal (~ 2 VEM, 420 ns, circled in blue)

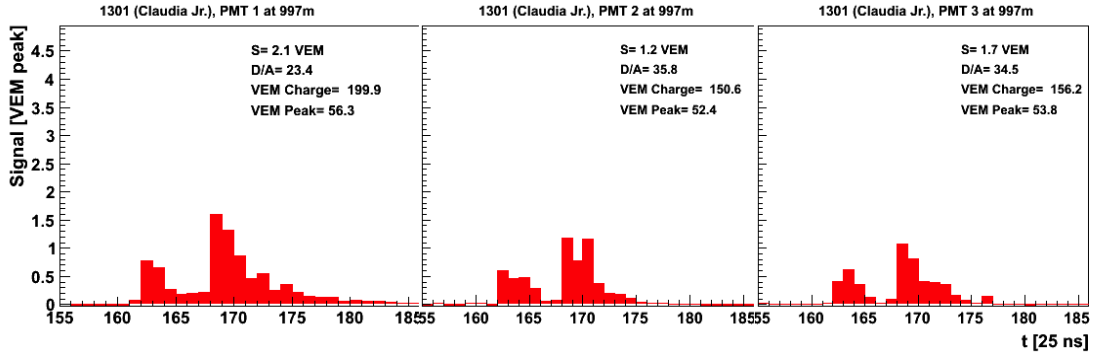


Figure 6.5-8 Zoom in of the accidental signal (~ 2 VEM, 420 ns)

These two examples are probably from accidental showers instead of single muons. All detectors surrounded have been checked and no accidental signals are found from their traces. If there is an accidental signal, the merged signal size only has ~ 0.2% of chance that it is larger than 6 VEM. However, if they occur at the beginning of the trace, one expects a very poor reconstruction of the event.

6.5.1.2 Distribution of the start time of the accidental signals

By applying the algorithms described in the previous section to data that have zenith angle $\theta < 30^{\circ}$, accidental signals are identified within the time period from 5000 ns to the trigger of the PMT. Figure 6.5-9 shows the distribution of the start time of the accidental signals.

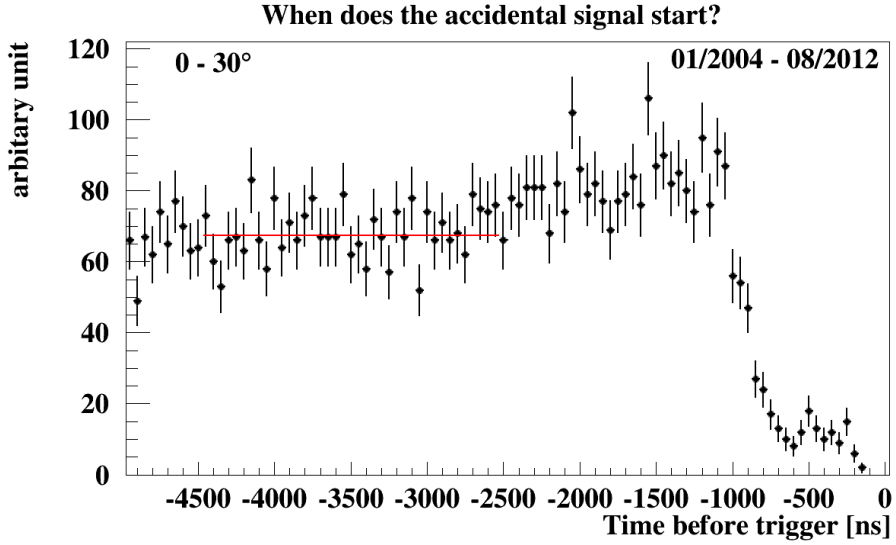


Figure 6.5-9 Distribution of the start time of the accidental signals. The trigger time of the PMT is treated as 0 and the time period used for scanning the accidental signal is from -5000 ns to the trigger time. The horizontal red line is a fit to the distribution between -4500 to 2500 ns. The reason that the rate of accidental signal drops from -1500 ns to 0 is perhaps due to triggers of the event. The time period -4500 to -2500 ns is used for further analyses.

The rate of having accidental signals is relatively flat from -5000 to -2000 ns. Then the rate is affected due to the trigger effect of the event. If the accidental signal is within ~ 1000 ns of the start time of the PMT, the signal is likely to be included as the signal of the shower event. The time period -4500 to -2500 ns is selected for the calculation in the next section.

6.5.1.3 Signal size and time interval of the accidental signals

The average signal and time length of accidental signals are shown in Figure 6.5-10 and Figure 6.5-11. As expected the majority of the signals are from single muons and the signal is therefore ~ 1 VEM. This plot is consistent

with the calibration curve shown in Figure 2.1-3.The mean of the time interval corresponds to the signal response of a muon, which is ~ 200 ns (the entity method makes use of trace between 87.5 ns to 235.5 ns). All calculations are done for two period of time for data-taking to ensure there is no dependence on the aging of detectors.

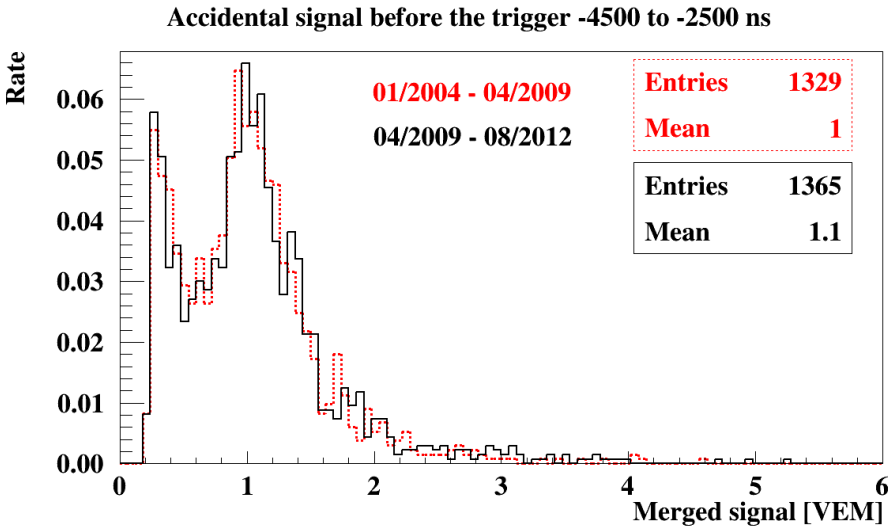


Figure 6.5-10 The averaged signal carried by accidental signals, which is consistent with Figure 2.1-3

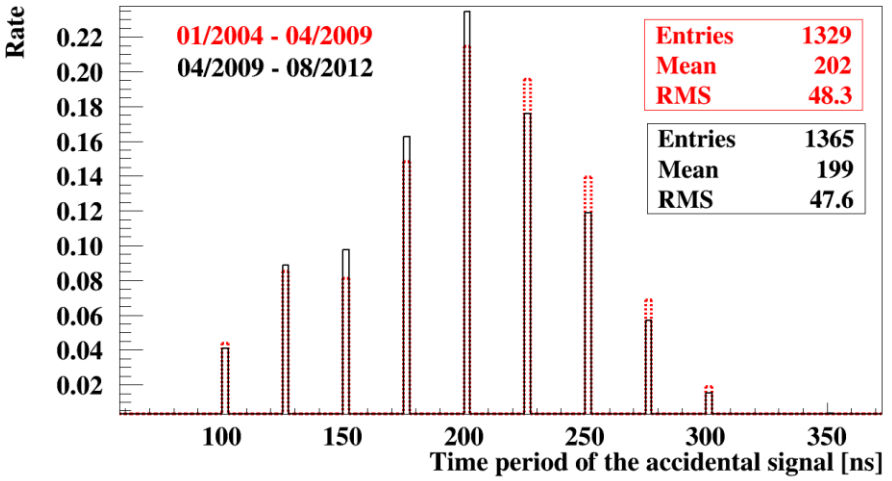


Figure 6.5-11 The averaged time spread of accidental signals. The vertical muon has response ~ 200 ns in the detector.

No disagreements have been found between different time periods of the data-taking and thus it is believed that the result is robust. The relationship between the time interval and the size of the accidental signal is shown in Figure 6.5-12, the colour shows the number of stations that fall in each bin. This plot is based on data from January 2004 to August 2012.

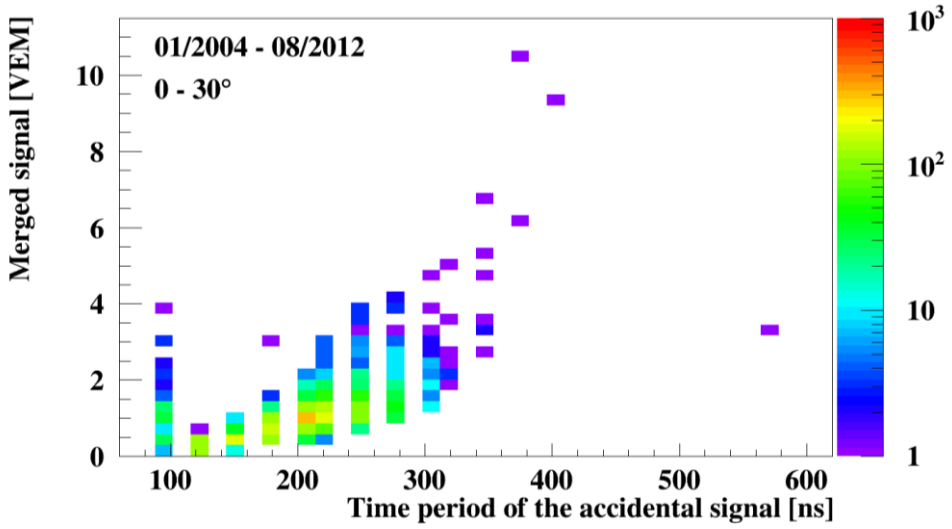


Figure 6.5-12 Relationship between signal size and time interval of accidental signals . Colours indicate the number of entries in each bin. For example there is only one event with time period larger than 500 ns and only 4 events with signal size larger than 6 VEM. Only accidental signals occur 4500 to 2500 ns before the trigger of the PMT are accounted. Data are from 01/04 to 08/12.

There are 273084 and 269261 FADC traces selected from the time period 01/2004 – 04/2009 and 04/2009 to 08/2012 respectively. 1329 and 1365 traces are found with accidental signals occurred by scanning a time interval of 2000 ns. The average length of each segment is ~ 200 ns. The probability of having an accidental signal that creates a signal of 25 ns is (1329 × 200 × 25) / (273084 × 2000) ~ 1.2%, which agrees with the result using the data from 2009 to 2012. Therefore it is unlikely to have more than one station in the event to have pre-pulses due to the accidental muons.

6.5.2 Pre-pulses from the shower

The pre-pulses could also be from productions of the shower. In this section, the physics behind various processes that could lead to pre-pulses are introduced. Two possibilities are explored: the EM halo which is caused by decayed muons and sampling fluctuations due to small number of particles received within the area of the detector.

6.5.2.1 EM halo

Although the EM component from π^0 decay is, on average, delayed with respect to the muon component, there is a possibility, that electromagnetic signal decayed from muons will arrive at about the same time as the muons. This EM component is referred as the EM halo. Shown in Figure 6.5-13is a schematic illustration of the 'halo' that muons turn to. Electrons from the decay develop into small showers and if this happens at the beginning of the trace, a pre-pulse is possible to form.

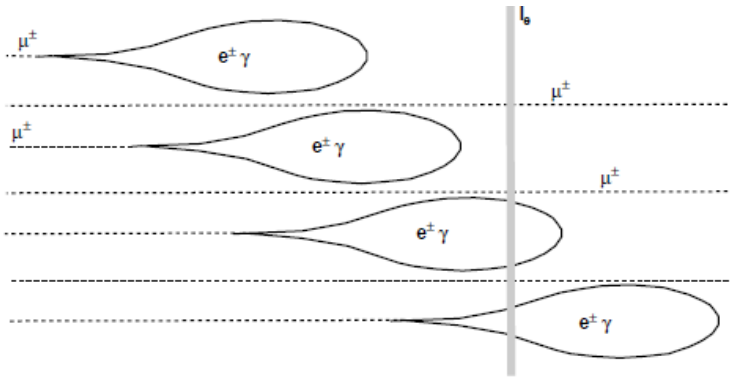


Figure 6.5-13 The EM halo: electromagnetic shower formed by a muon that decayed [112]. l_0 is the position of the detectors on the ground and the traces in the detector are different depending on the position of the EM shower. For the first two cases, signals in the water-Cherenkov detector are not affected. The two scenarios down, where decays occur close to the detector, EM signals could create FADC traces with a small signal pulse (pre-pulse) at the beginning of the trace.

Interactions between muon and matter depend on the primary energy of the muon. Below 100 GeV, the energy loss is dominated by ionisation and the energy loss rate does not depend on the primary energy of the muon. 'Knock-on' electrons, also called δ-rays, typically carry energies too low (~ 10 times less than the energy of photons from ionisation) to propagate in the atmosphere. Bremsstrahlung, pair-production and the photonuclear interactions are less frequent for low energy muon. Details are shown in Figure 6.5-14.

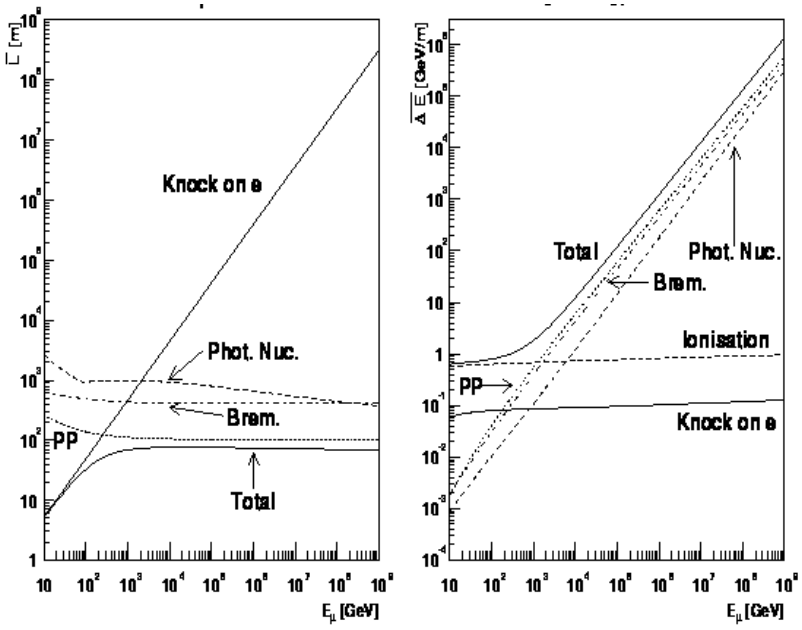


Figure 6.5-14 Mean free path and the energy loss rate for muons at different energies[113]. The curves are describing interactions between muon and rock, for interactions with water / air, values are needed to be scaled according to the matter density.

The average energy of muons arriving at the detector is ~ 1 GeV. Muons that arrive first are likely to be produced higher in the atmosphere and carry larger energies. To obtain reliable results of energy spectrum of particles arriving on the ground, artificial fluctuations from the 'de-thinning' procedures should be avoided. The concept of 'thinning' was introduced to reduce the amount of CPU time and memory needed for air-shower simulations by assigning weights to particles above an energy threshold. De-thinning is the procedure to un-weight particles on the ground level and allows them to be sampled by the detectors. Details of impact of thinning are introduced in section 2.4 and 9.4. Using an unthinned proton shower at 0°, muon energies are plotted against arrival times. No detector simulations are included and the time drawn is without detector responses. The receiving area of the muons is a ring with the radius 745 m to 755 m, the colour code of Figure 6.5-15 shows the number of muons fall in each energy-time bin.

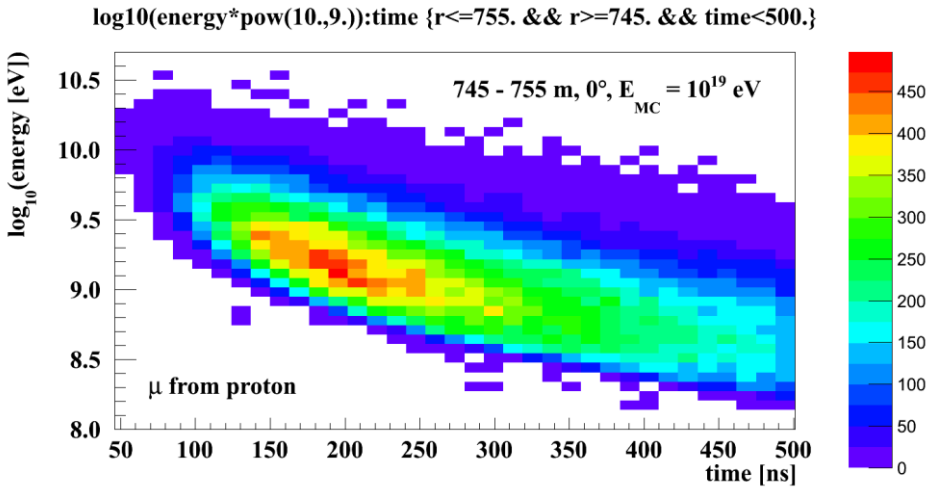


Figure 6.5-15 Arrival time and energy of muon from a vertical proton shower (unthinned). It is from direct CORSIKA outputs without simulating detector responses

The energies of muons are well below the critical energy, where bremsstrahlung starts to dominate. However, muons could also decay during the propagation. The muon decay follows

$$\mu^{\pm} \rightarrow e^{\pm} + \nu_e/\overline{\nu_e} + \overline{\nu_\mu}/\nu_\mu \qquad 54.$$

The lifetime of muon is $\tau \sim 2 \times 10^{-6}$ s. To find out the probability that muons decay to electrons, one needs to know the production height of the muon. Using proton showers simulated using the EPOS-LHC， Figure 6.5-15 shows the change of number of muons with respect to the depth of the shower. The simulations were done for 60° and Monte Carlo energy 10^{19} eV.

If the muon comes from ~ 30,000 m, the probability it decays at the ground is

$$P = e^{-t/\tau\Gamma} \quad 55,$$

where Γ is the Lorenz factor and t is the time that the muon has travelled. The probability for a muon with energy 10 GeV to decay is ~ 60%.

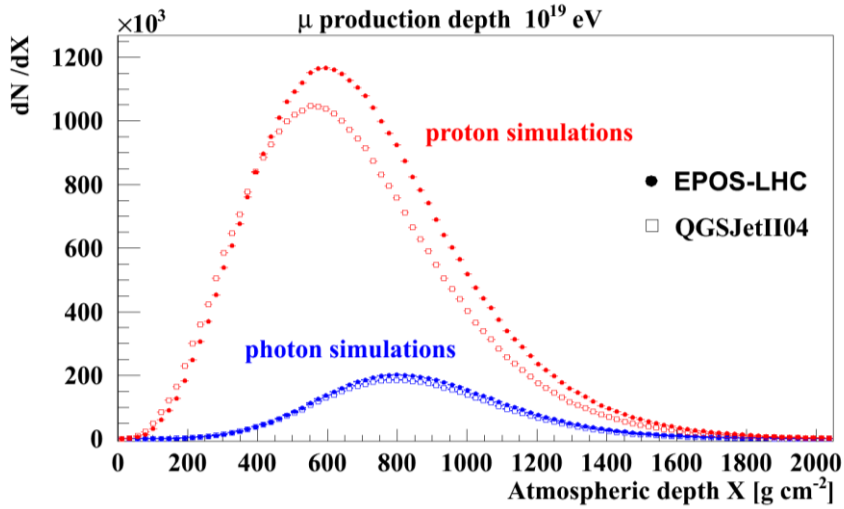


Figure 6.5-16 Change of muon number with respect to shower depth for photon and proton simulations using different hadronic interaction models

This feature has also been observed in a detailed study of the components of the traces using Monte Carlo simulations. Shown in Figure 6.5-17 is the averaged four-component trace using simulations of protons with 0° and 10^{19} eV at 1000 m.

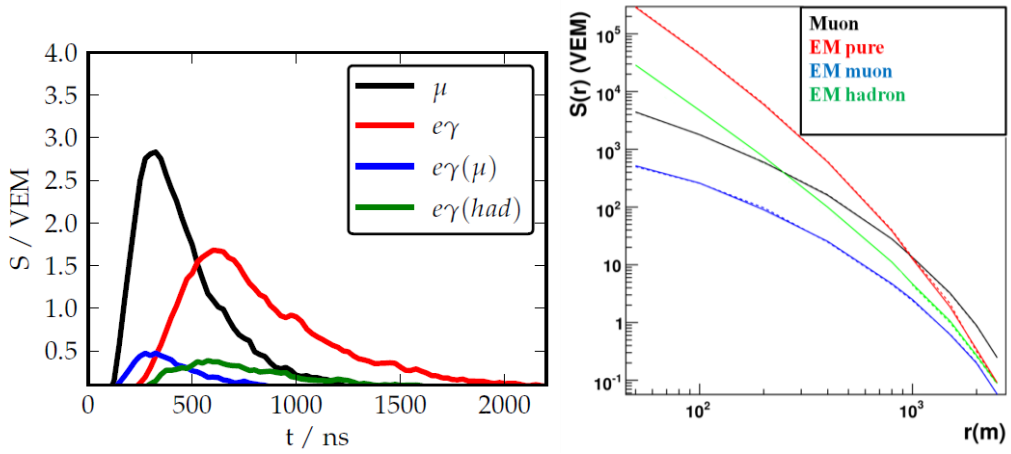


Figure 6.5-17 Component traces of vertical proton simulations of 10^{19} eV, r = 1000 m[114] (left). There are three sources of electromagnetic component: from muon decays (EM halo), decay of low energy hadron jets, and the pure electromagnetic component from decay of π^0. Lateral distribution function of the four-components[114] is shown on the right. EM muon is the electromagnetic component from muons (EM halo)

Compared to the component traces shown in this work, the four-component trace separates electromagnetic components into three categories. The red is from the electromagnetic shower during the development, the green is

electromagnetic from hadrons and the blue is, as discussed, that are dominated from the decay of muons.

6.5.2.2 Can the pre-pulse occur due to the difficulty of sampling small number of particles?

The area of each surface detector is ~ 10 m^2 and suffers Poisson fluctuations when there are a small number of particles arriving at the detector. As in the previous section, unthinned showers are used to avoid artificial fluctuations (details in section 9.4). The idea of this study is to investigate if the Poisson fluctuations from sampling particles could cause the pre-pulse feature. Proton (Figure 6.5-18), iron (Figure 6.5-19) and photon (Figure 6.5-20) simulations of $\theta = 0^\circ$ and $E = 10^{19}$ eV were used. An area of ring with radius 745 to 755 m was selected around the core position and then normalised to an area of 10 m^2. Detector responses were not simulated.

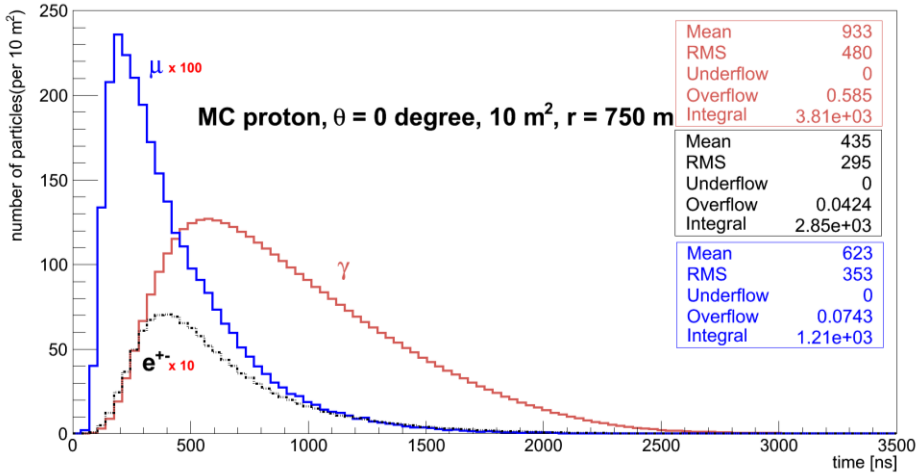


Figure 6.5-18 Number of arrival particles with respect to time for three components of an unthinned proton shower. Number of muons and electrons are scaled just for the illustration purpose.

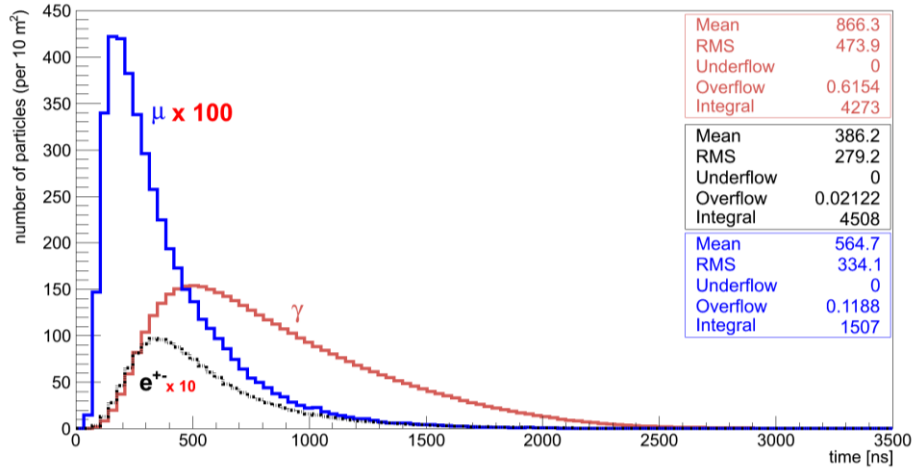


Figure 6.5-19 Number of arrival particles with respect to time for three components of an unthinned iron shower

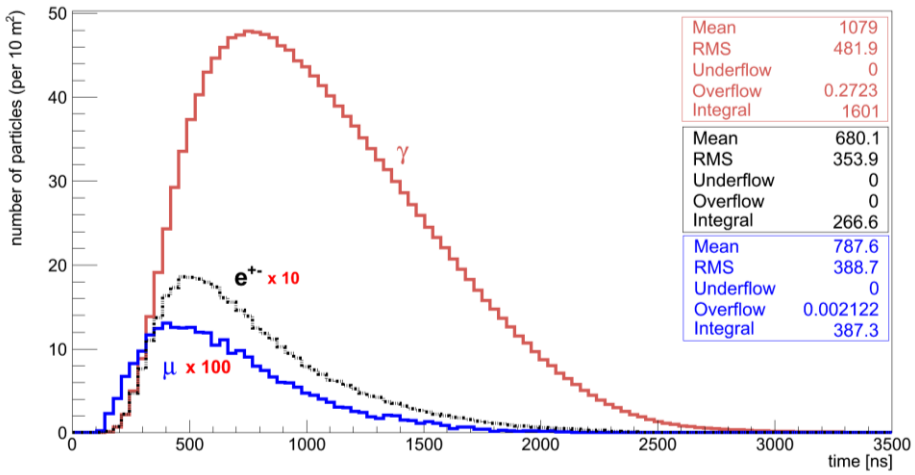


Figure 6.5-20 Number of arrival particles with respect to time for three components of an unthinned photon shower.

Numbers of muons and electrons have been scaled in the above plots for the illustration purpose. It is interesting to see that muon always arrive early no matter what the primary of the cosmic ray is. The arrival of muon is delayed in the case of photon and the number of muons is ~ 4 times smaller than what is seen from the iron shower. Of course this is just an example for one specific distance, energy and zenith angle. More studies are needed to give quantitative conclusions.

It could also be concluded that the pre-pulse might be from the muon component of a photon shower from Figure 6.5-20. Because the signal carried by the muon is insignificant compared to the signal from electromagnetic component, the rise of the muon signal could be interpreted as a pre-pulse.

6.6 Summary

In this section, it has been demonstrated that the entity likelihood ratio is a very competitive parameter for the photon-hadron separation. As illustrated in section 6.5, the pre-pulse feature is not excluded from either photon or hadronic behaviour. More simulations and detail studies are need to for further estimations. However, the first two categories of candidates (PMT and start time problems) are removed from the candidate list.

To summarise, 72, 7 and 2 photon candidates are found that between the energy range $10^{18.8} < E^{\gamma}_{rec} < 10^{19.0}$ eV, $10^{19.0} < E^{\gamma}_{rec} < 10^{19.2}$ eV and $10^{19.2} < E^{\gamma}_{rec} < 10^{19.4}$ eV respectively. No candidates are found that with energy beyond $10^{19.4}$ eV.

In the next chapter, results of the upper limit of the photon flux are shown with the assumption that all candidates found are hadronic background.

Chapter 7. Results of the entity method

By applying the entity method to data that were taken between January 2004 to March 2014 with $0 < \theta < 60^\circ$, 72, 7 and 2 photon candidates are found that lie in the energy range $10^{18.8} < E^{\gamma}_{rec} < 10^{19.0}$ eV, $10^{19.0} < E^{\gamma}_{rec} < 10^{19.2}$ eV and $10^{19.2} < E^{\gamma}_{rec} < 10^{19.4}$ eV respectively. No candidates are found that with energy beyond $10^{19.4}$ eV. Through inspecting all events, it is clear that the candidate events are possibly of light composition.

However, to claim the discovery of UHE photons, more work is needed for estimating the backgrounds arising from proton primaries. For instance the background could be from accidental muons (section 6.5.1), from the EM halo (section 6.5.2.1) and from π^0 produced as the leading particle from the first interaction in the hadron showers. The latter case is discussed in details in Chapter 8. The uncertainties of hadronic interaction models and statistical thinning procedures (section 2.4, 9.4) make results such as muon fractions and FADC traces based on simulations less reliable. In this thesis, we consider all photon candidates found are from the hadronic background.

This chapter is on results of the upper limit of photon flux and the ingredients used for the calculation. First the energy reconstruction on the photon scale (E^{γ}_{rec}) is introduced, then the cuts that applied in the entity method are summarised. The upper limit from the entity method is compared to the standard SD photon search ($t_{1/2}(1000)$ and R_c, section 3.3) using the same period of events and zenith angle range. Finally the upper limit to the photon flux using the entity method covering data from January 2004 to March 2014 and $0 < \theta < 60^\circ$ is shown.

7.1 Reconstructing energy in the photon scale

The calibration of energy for SD events has been discussed in section 2.3. It is based on the relationship between the calorimetric energy measured by FD and the signal at 1000 m from the core on the ground, S(1000). Data were used for the energy conversion which makes the result consistent with hadron showers due to the photon fraction[107] of data is < 0.5% for showers with energy larger than 2×10^{18} eV.

Photons have different shower developments and different lateral distributions on the ground. As shown in Figure 3.2-10, from a vertical shower with Monte Carlo energy $E_{MC} = 10^{18.5}$ eV, S(1000) for a proton

shower is ~ 20 VEM and is ~ 10 VEM for a photon shower. The reconstructed energy of this photon shower is ~ 10^{18} eV, which underestimated the energy of ~ 70%. It is therefore mandatory to modify the energy reconstructions to have the energy correctly converted to the energy in photon scale E_{rec}^{γ}.

Techniques have been developed[115] for estimating the photon energy: a summary and demonstration follows using the photon simulation library that has been used in this thesis. The photon simulation library has photons with 0 - 60˚ and with index of the energy spectrum -1. It contains pre-showers at the highest energies (Figure 7.1-1). Only showers that triggered the array are included.

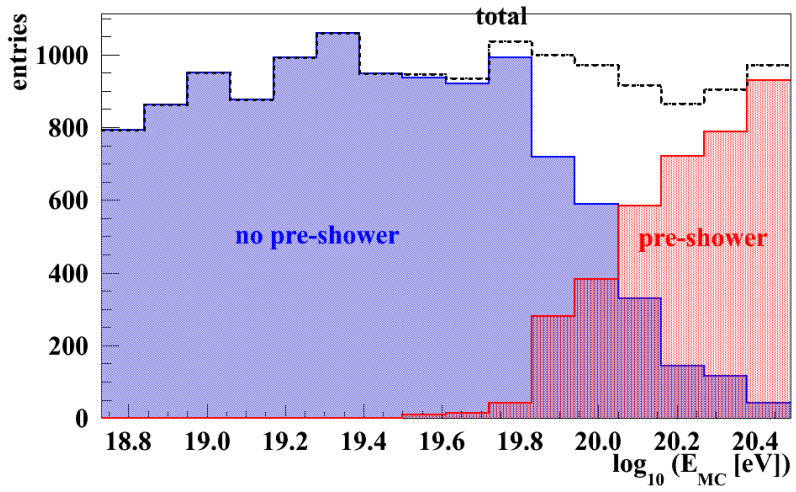


Figure 7.1-1 Distribution of MC energy E_{MC} of showers in the photon simulation library. The blue area is showers that without pre-showers, the red region are pre-showers and the black line is the total. The energy follows E^{-1} spectrum, from QGSJetII-03, simulated at Wuppertal. The energy spectrum is always scaled to E^{-2} in the analyses.

For electromagnetic showers, the signal at large distance (i.e. S(1000)) is indicative of particle densities and the evolution of the shower development. The shower evolution is often associated with the distance (DX) between the shower depth at the level of the detector and X_{max}. For showers with fixed energy E, a large DX means the size of the shower starts to decrease high in the atmosphere and therefore the signal on the ground S(1000) is smaller. However, this is only valid for showers with DX > 200 g cm^{-2}. For showers have X_{max} close to the ground, the correlation between S(1000) and E_{MC} is different from the cases that showers have been developing sometime after X_{max}. The relationship between DX, Monte Carlo energy E_{MC} and S(1000) can be described using

$$\frac{S(1000)}{E_{MC}} = \frac{r_0\left[1+\frac{DX-r_3}{r_1}\right]}{1+\left(\frac{DX-r_3}{r_2}\right)^2} \qquad 56.$$

Values of r_0, r_1, r_2 and r_3 are obtained by fitting photon simulations. The curve, which is also known as the 'universality' curve, is shown in Figure 7.1-2.

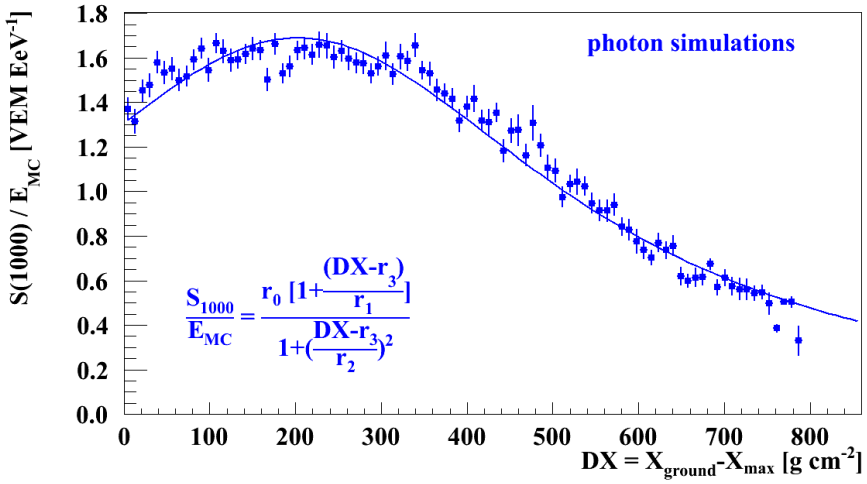


Figure 7.1-2 The relationship between S(1000), E_{MC} and DX, from photon simulations. DX is the depth of the atmosphere from X_{max} to the ground. Showers have large zenith angles are with large DX. The curve is also known as the 'universality' curve and fitted with a Gaisser-Hillas function.

Because X_{max} is also sensitive to the energy of the shower (Figure 3.1-5), it can be expressed as a function of E_{MC}

$$X_{max} = r_4 + r_5 \log_{10} E_{MC} \qquad 57.$$

Therefore since DX = X_{ground} – X_{max} and X_{ground} only depends on the zenith angle of the shower, DX can be described as a function of E_{MC}.

For an event that has S(1000) and θ, several iterations are made for fitting the universality (equation 56) and the elongation rate (equation 57) curves to get the best combination of X_{max} and the energy. The iteration is terminated if DX is smaller than - 50 g cm^{-2} because otherwise X_{max} is too deep under the ground. The bias of the energy reconstruction is shown in Figure 7.1-3 and the comparison of the reconstructed energy on the hadron and photon scale is shown in Figure 7.1-4.

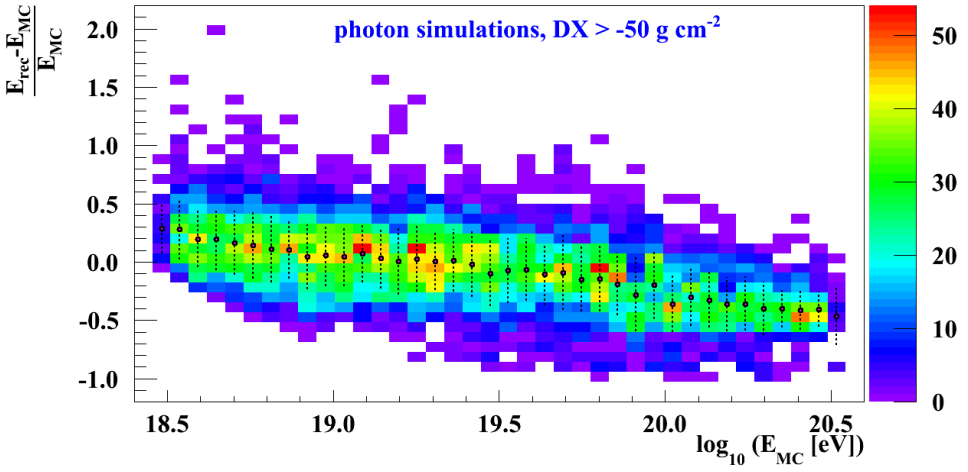


Figure 7.1-3 The bias of the calibrated energy for photon simulations used in the thesis. The plot includes pre-showers. The points show the mean in each energy range and the error bars show 1 sigma. The colour shows the number of entries in the bin.

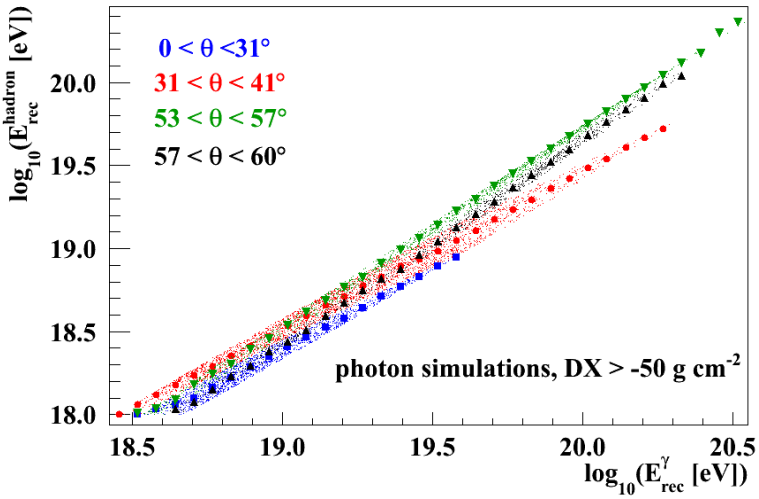


Figure 7.1-4 Relationship between reconstructed energy on photon and hadron scale for showers have different zenith angles

As expected there is a shift of the E_{rec}^{γ} from E_{rec}^{hadron} , the latter is lower than the former. In this thesis, to find photon candidates, both data and photon simulations firstly have energies reconstructed in the photon scale. Then if the event from data is identified as non-photon, the energy is calculated again but in the hadron scale. In this way we can calculate the photon fraction with the minimum interference from the different energy scales.

7.2 Cuts and efficiencies

Events are required to pass the T4 and 6T5 triggers which were introduced in section 2.1. Bad periods when data are of poor quality are not used (including the comms-crisis in 2011). The optimisation cuts on detector

distance and signal range have been determined in section 5.3. The station-quality cuts have been discussed in the section 6.3. The last cut is the requirement that X_{max} has to be not below the ground by more than 50 g cm^{-2}. The list of all cuts is summarised as:

a). T4, 6T5, 0 – 60˚, no bad periods, no lightning.

b). $X_{max} - X_{ground} > -50$ g cm^2 and possible to have a reconstructed energy on the photon scale (i.e., fit converges)

c). Optimisation cuts: At least one station with 400 < r < 1300 m and 30 < S < 800 VEM; other stations selected 400 < r < 1500 m, 15 < S < 800 VEM. At least two stations selected per event and neither of the stations is low-gain saturated.

d). Station-quality cuts: S4 / S8 < 70% and 2% < S8 / S < 80%.

For a typical photon shower that has $E_{MC} = 10^{19}$ eV and θ = 0˚, the shower of maximum is X_{max} ~ 950 g cm^{-2} which is below the ground by ~ 75 g cm^{-2}. Therefore vertical showers with such energies are likely to be cut away from the sample. Figure 7.2-1 shows the events that survived the DX cut. Events with θ > 48˚ survive this cut regardless the energy.

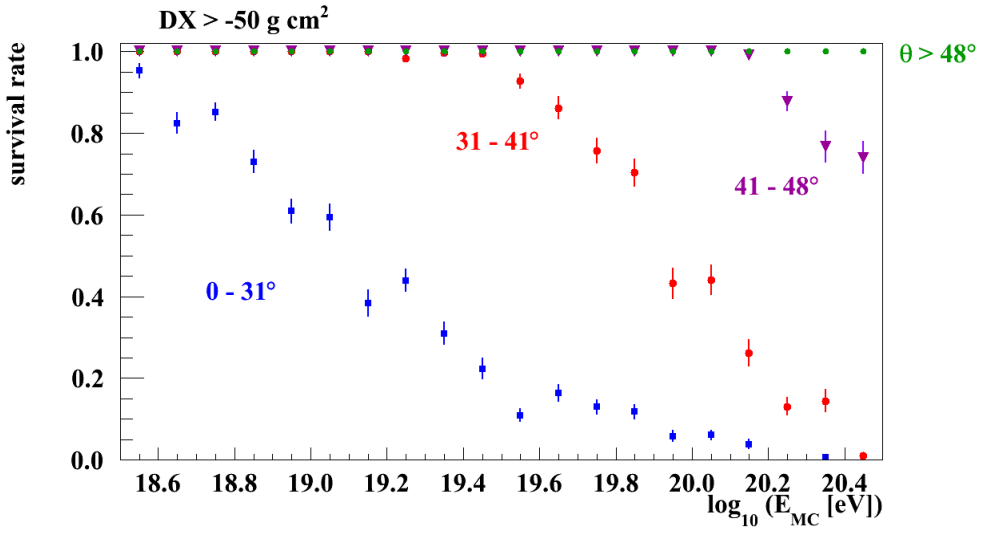


Figure 7.2-1 Survival rate from the cut DX < -50 g cm^{-2} of photons that have different θ. The results are based on photon simulations that triggered the events. Vertical showers are less likely to survive the DX-cut.

The selection efficiency after all cuts is shown for 30 < θ < 60˚ and 0 < θ < 60˚ in Figure 7.2-2.

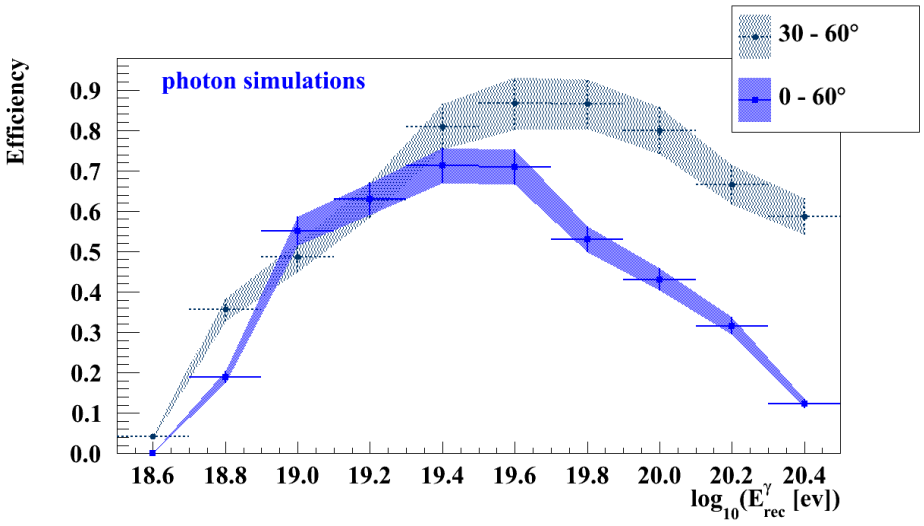


Figure 7.2-2 Selection efficiency for photon simulations of $30 < \theta < 60^\circ$ and $0 < \theta < 60^\circ$. The gain of the efficiency by excluding vertical showers is because vertical showers at the highest energies have DX < -50 g cm^{-2}. The calculation includes pre-showers. The error bars show one sigma of the distribution in each energy bin.

At the lower energies, the trigger efficiency restricts the selection of events. At the higher energies, X_{max} is below the ground. By including 0 – 30°, the selection efficiency decreases especially at large energies. However, the aperture is 1.5 times larger than if the zenith angle range used is only 30 – 60°.

7.3 Upper limit to the photon flux and fraction

To calculate the upper limit of the photon flux, the formula used is

$$\Phi_{95}\,(E > E_0) = N_\gamma^{95}\,(E_{rec}^\gamma > E_{min}) \,/\, (A \times \varepsilon \times 0.5) \qquad 58.$$

N_γ^{95} is the number of photon candidates at 95% confidence level. A is the aperture used for the flux. ε is the selection efficiency of the photon simulations and 0.5 is due to the cut is on the median of the photon. The reason of cutting on the median is to be consistent with the previous analyses of the photon search in Auger.

To compare the result to the standard SD photon search, the integral flux limits using data that have $30 < \theta < 60^\circ$ and taken between January 2004 to March 2011 are used. 5, 2, 0, 0 candidates are found for integrated energy bins $E_{rec}^\gamma > 10^{19}$ eV, $E_{rec}^\gamma > 2 \times 10^{19}$ eV, $E_{rec}^\gamma > 3 \times 10^{19}$ eV and $E_{rec}^\gamma > 4 \times 10^{19}$ eV respectively. Shown in Figure 7.3-1 are results from the entity method (black) and the standard SD search (combined analysis of radius of curvature and risetime).

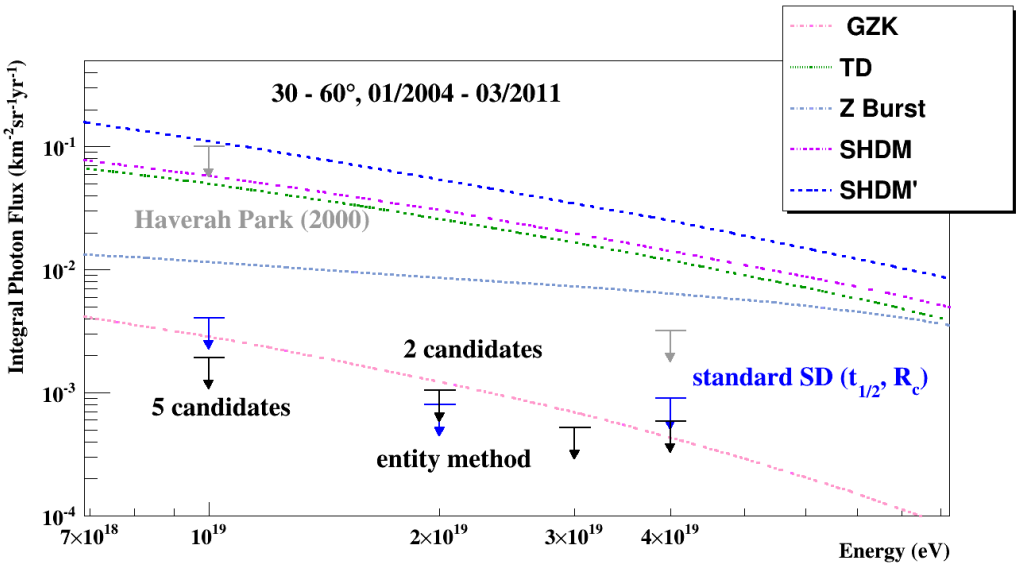


Figure 7.3-1 The comparison of upper limits of the photon flux. Same data period and zenith angle range are used for both the standard SD and the entity method (01/2004 – 03/2011). The prediction lines from different models are consistent with what are used in Figure 1.4-4.

The entity method has the best merit factors without combining with any other parameters. It gives a better limit than the standard multivariate analysis (section 3.3) in all but one bin. The reason for having a worse result at $E > 2 \times 10^{19}$ eV is because there were no candidates found in the standard analysis. The fact that entity method does not need a large number of stations for each event allows one to apply the method to lower energies and smaller zenith angles. An update of the photon flux limit using the entity method applied to data until March 2014 is shown in Figure 7.3-2.

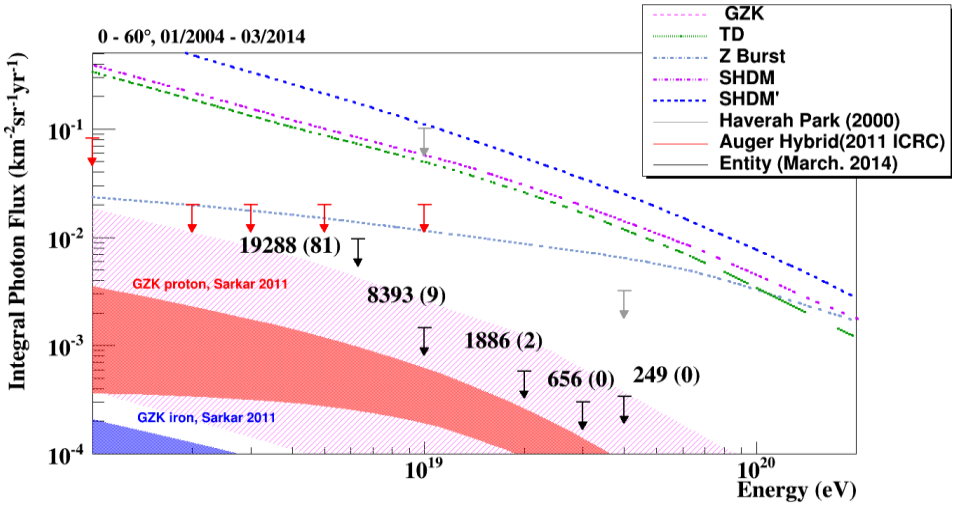


Figure 7.3-2 The result of upper limits of integrated photon flux from the entity method. Data period used is 01/2004 – 03/2014. The red band is the prediction of GZK-photons from proton-primaries and the blue is the GZK-photons from iron-primaries[116]. They were calculated based on Auger energy spectrum from 2011. The pink band is the GZK-photon predicated from proton-primary and is used as the reference[117] in the previous publications of the Auger photon searches.

This calculation includes zenith angle 0 – 60°. The number of candidates increased from 5 to 9 for the energy bin $E^{\gamma}_{rec} > 10^{19}$ eV and the efficiency decreased from 80% to 61% due including vertical showers. Data that were not identified as photons then had energy reconstructed in the hadron scale, which are labelled in Figure 7.3-2. A list of number of candidates, selection efficiencies, photon flux limit and fraction limit is shown in Table 7.3-1.

Table 7.3-1 Summary of the limits of integral flux and fraction based on events from 01/2004 – 03/2014, $0 < \theta < 60°$

(Integral) energy (eV)	$N_\gamma+N_{non\text{-}\gamma}$	N_γ photon candidate	N_γ^{95} photon candidate	ε	Flux limit ($km^{-2}sr^{-1}yr^{-1}$)	fraction
$>10^{18.8}$	19288	81	97.4	57%	0.0097	1.8%
$>10^{19.0}$	8393	9	15.3	61%	0.00145	0.6%
$>2 \times 10^{19}$	1886	2	6.3	61%	0.00058	1.1%
$>3 \times 10^{19}$	656	0	3	56%	0.0003	1.6%
$>4 \times 10^{19}$	249	0	3	50%	0.00034	4.8%

The upper limit of the photon flux in differential energy bins based on data taken from January 2004 to March 2014 and $0 < \theta < 60°$ is shown in Figure 7.3-3.

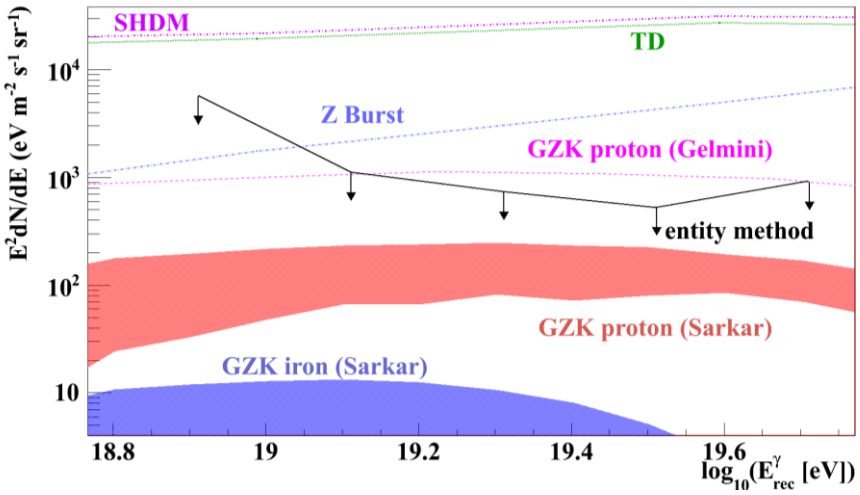


Figure 7.3-3 Limit of the photon flux in differential energy bins using the entity method. Astrophysical scenarios were taken from same models as drawn in Figure 7.3-2

The result is encouraging that upper limit obtained in this thesis is below all top-down models and already entering the GZK region (model from Gelmini) which has protons as UHECR.

Chapter 8. A possible source of background from protons with π^0 as the leading particle

A key factor, in addition to inelasticity, multiplicity and cross-section associated with the first interaction, is the nature of the particle, called the leading particle, which takes most of the energy from the collision. For hadronic interactions, the leading particles are often protons or neutrons

$$p + p \rightarrow p + p + N(\pi^0 + \pi^+ + \pi^-) \quad 59,$$

$$p + p \rightarrow p + n + N(\pi^0 + \pi^-) + (N+1)\pi^+ \quad 60,$$

and less frequently pions or other particles. When the leading particle is a π^0, considerably more energy of the shower is channelled electromagnetic channel, which thus makes the shower photon-like. Searches for background from protons with a leading π^0 that are photon-like have been studied for applications of ground-based gamma-ray observatories[118]. In this chapter, we explore the possibility of back-grounds from proton showers with the leading π^0 at ultra-high energy, including the dependence on hadronic interaction models.

8.1 Leading particles from the first interactions

From the toy model (section 1.3) of shower development, a heavy nucleus A, can be considered as the superposition of nucleons and creates an air shower that is more dominated by the hadronic component. Since the aim of this chapter is to explore the possibility of having photon-like showers from hadronic primaries, it is natural to concentrate on proton showers first.

The leading particle from the first interaction is most likely to be a proton (25% according to EPOS-LHC and 35% with QGSJetII-04) in a typical interaction. Figure 8.1-1 shows probabilities of various particles being the leading particle from the first interaction. The results were calculated using simulations generated on the Pleiades cluster in Wuppertal during 2014. Protons are simulated using the CONEX program (section 2.4), with $\theta = 60^\circ$, E^{-1} spectrum and Monte Carlo energy from $10^{17.9}$ to $10^{20.5}$ eV. There are ~ 440,000 showers simulated for each hadronic interaction model. The particle IDs marked on the x-axis correspond to the CORSIKA encoding[119].

ID 7 represents π^0s which will decay to photons that fuel the electromagnetic component of the shower following

$$\pi^0 \rightarrow \gamma + \gamma \qquad 61.$$

It is seen that EPOS-LHC produces more pions but fewer protons in the first interaction compared to QGSJetII-04.

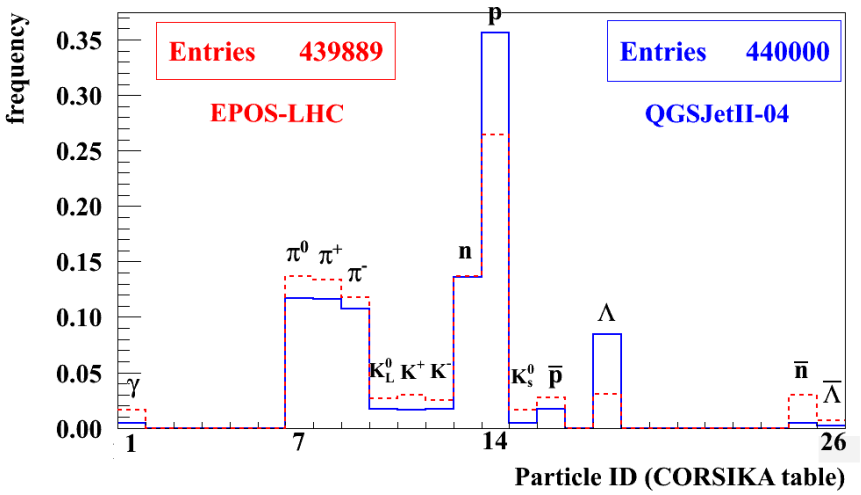


Figure 8.1-1 Types of leading particles from the first interactions for proton simulations. Protons have energy from $10^{17.9}$ to $10^{20.5}$ eV and follow an E^{-1} spectrum.

On average the possibility of having π^0s as the leading particle is ~ 12% from QGSJetII-04 and ~ 14% from EPOS-LHC (Figure 8.1-2).

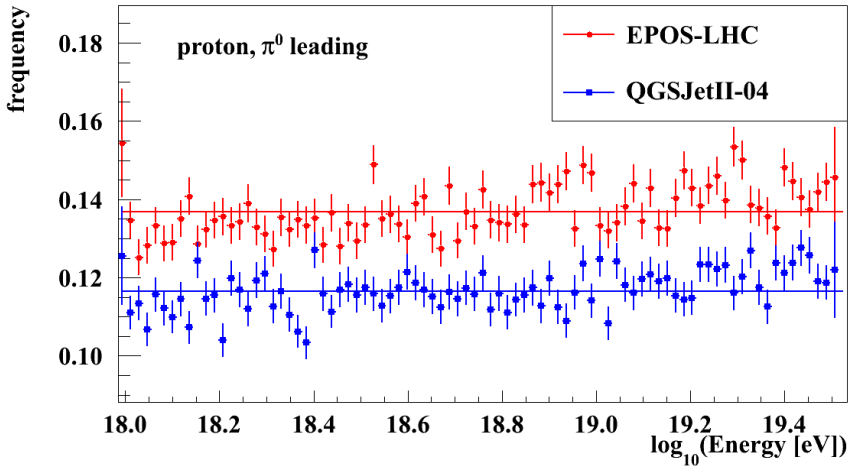


Figure 8.1-2 Energy dependence on the probability that π^0 leading the first interaction

The quantity to describe the amount of energy carried by the leading particle over the total energy of the shower is elasticity k_{ela}. The dependence of the elasticity on the energy of the shower is shown in Figure 8.1-2 and is nearly independent of energy for both models. The amount of energy that goes into the electromagnetic channel is determined by the elasticity when the leading

particle is π^0. The averaged elasticity with respect to the energy of the primary proton is shown in Figure 8.1-3.

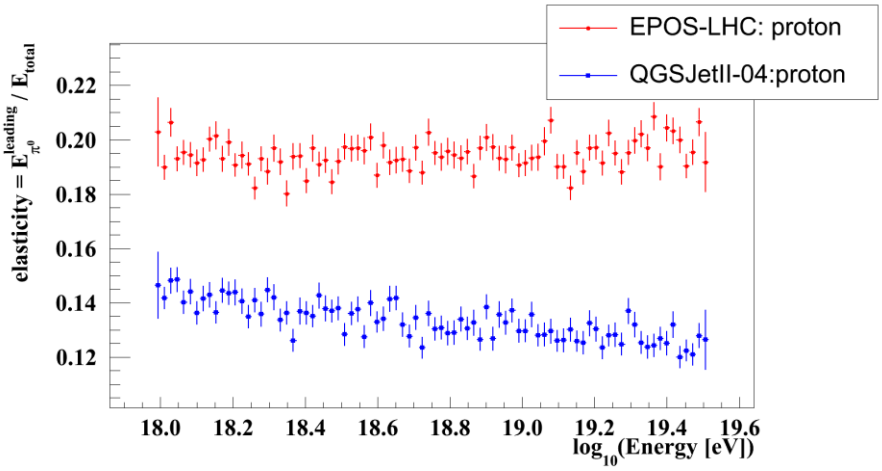


Figure 8.1-3 Elasticity of neutral pion as the leading particle in the first interaction with respect to the primary energy of the shower

The elasticity does not change with increasing energy for EPOS-LHC while it decreases slowly for QGSJetII-04. The energy carried by π^0 from the EPOS model is ~ 20% while from the QGSJet model is < 14%. The distribution of elasticity is shown in Figure 8.1-4.

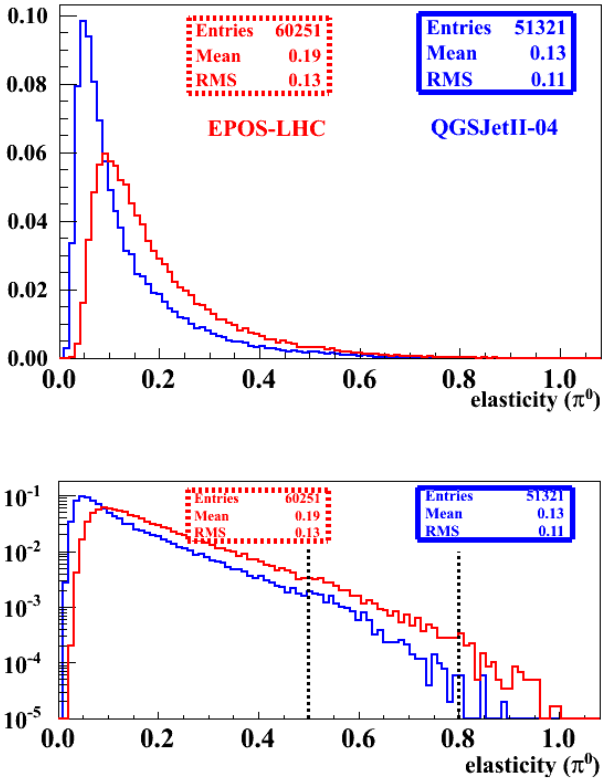


Figure 8.1-4 Distribution of elasticity of π^0 as the leading particle in the first interaction. (The lower plot is in logarithmic scale) Dashed lines show the elasticity = 50% and 80%. EPOS-LHC has more π^0 with higher elasticity than QGSJetII-04.

In principle the larger the elasticity, the more energy of the shower is deposited in the electromagnetic component and therefore the shower is more likely to appear as a photon candidate. In the next section the first study of the background based on proton showers with elasticity > 80% is introduced.

8.2 The averaged shower profile of proton showers with a leading π^0 in the first interaction

Fluorescence telescopes are used to determine the shower maximum which corresponds to the atmospheric depth where the number of particles reaches the maximum. The shower profile is observed as a function of number of particles with respect to the depth in the atmosphere. The electromagnetic component dominates the shower development. In Figure 8.2-1, profiles of the photon component and muon component are shown.

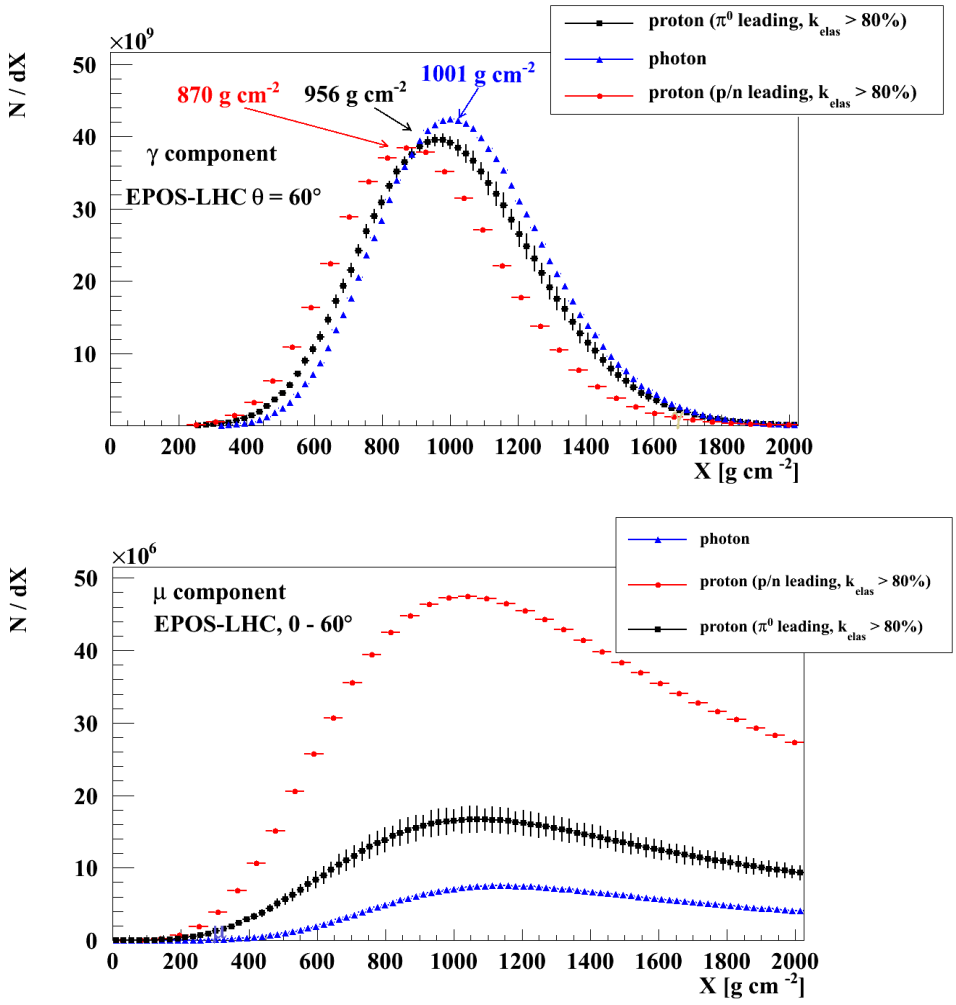


Figure 8.2-1 Component profile of photons and protons with different leading particles. Protons with proton / neutron leading are in red, with π^0 leading are shown in black. The elasticity is larger than 80% in both cases. Photon showers are in blue.

Red points are from proton simulations when the leading particle is a proton with elasticity larger than 80%, black points are also from proton simulations but with a π^0 as the leading particle of elasticity larger than 80% and blue points are photon simulations. EPOS-LHC is used for the simulations of showers with Monte Carlo energy 10^{19} to $10^{19.2}$ eV. The muon component and photon component are drawn separately. The numbers of secondary photons has its maximum at 870, 1001 and 956 g cm^{-2} respectively. Similarly it can be seen that protons with π^0 leading have the muon profile between the averaged muon profile for protons and photons.

8.3 X_{max} and N^{μ}_{max} of photon-like showers from proton simulations

As introduced in section 3.6, X_{max} and N^{μ}_{max} can be used as the direct measurement for photon-hadron separations. In this section, we quantify whether protons with leading π^0 are photon-like by using these two variables. The results are from the an idealised scenario since values were taken directly from simulations and no detector resolution or reconstruction procedures were included.

Proton simulations that have π^0 as leading particles and with elasticity larger than 80% are selected as 'photon-like' protons. The rest of proton simulations are treated as protons and the mean with a contour including 90% of the entire distribution are drawn in Figure 8.3-1 to Figure 8.3-4. Both EPOS-LHC and QGSJetII-04 have been used for simulations of proton, photon and iron.

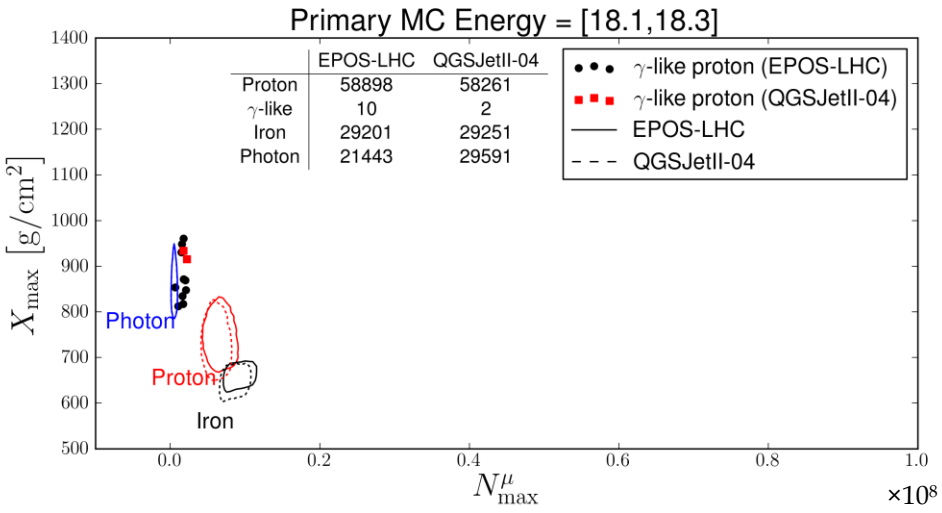


Figure 8.3-1 X_{max} and N^{μ}_{max} of γ, Fe, p and γ-like simulations of $10^{18.1} < E_{MC} < 10^{18.3}$ eV. The contour shows 90% of the population and the black points are photon-like protons which have π^0 as the leading particle in the first interaction and elasticity larger than 80%.

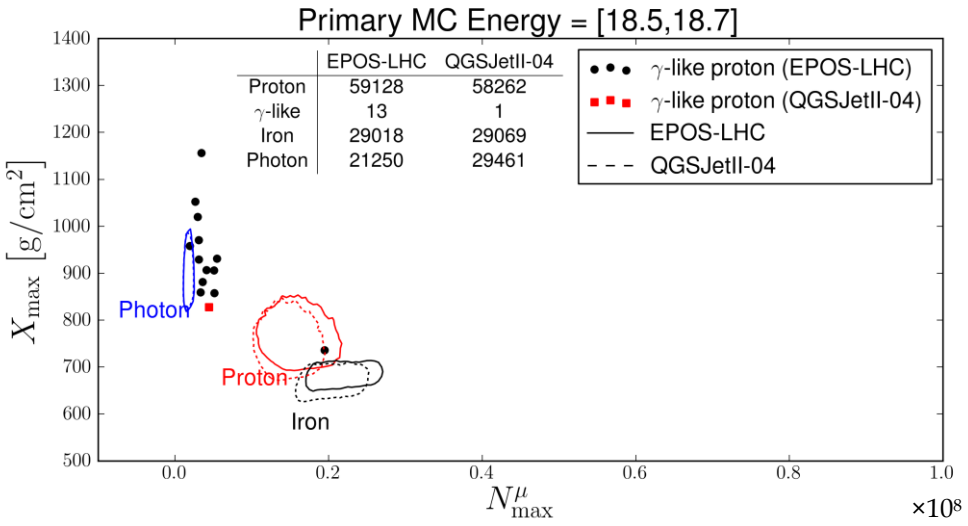


Figure 8.3-2 X_{max} and N^{μ}_{max} of γ, Fe, p and γ-like simulations of $10^{18.5} < E_{MC} < 10^{18.7}$ eV

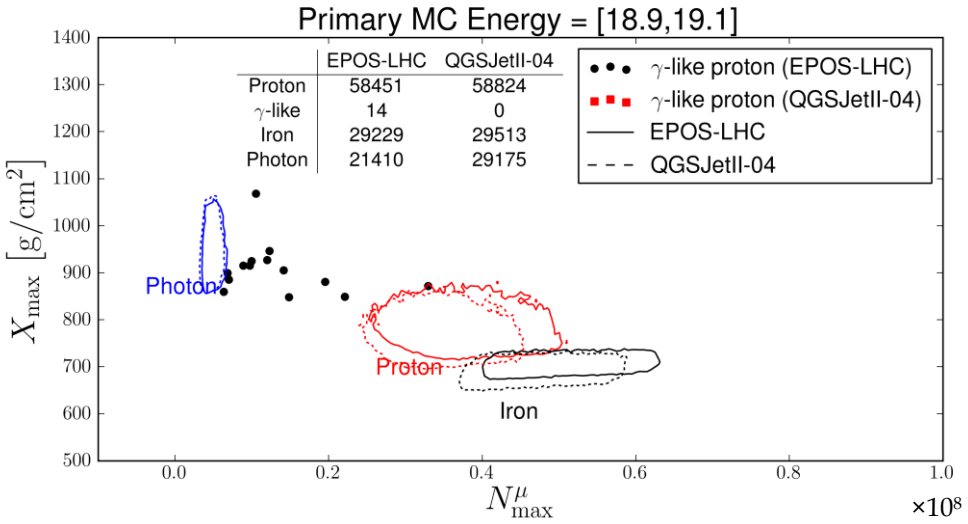


Figure 8.3-3 X_{max} and N^{μ}_{max} of γ, Fe, p and γ-like simulations of $10^{18.9} < E_{MC} < 10^{19.1}$ eV

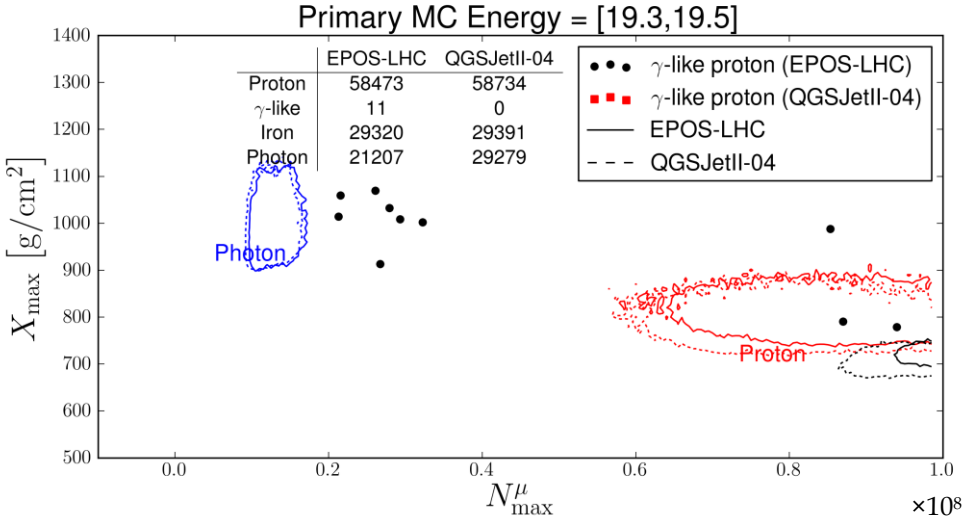


Figure 8.3-4 X_{max} and N^{μ}_{max} of γ, Fe, p and γ-like simulations of $10^{19.1} < E_{MC} < 10^{19.3}$ eV. Numbers of each type of shower are listed in the table on the figure.

The photon-like protons have a mean X_{max} similar to the distribution from photon simulations especially at low energies. N^{μ}_{max} of the photon-like protons are mostly between the photon and the proton distribution. Based on this idealised scenario, the conclusion is that N^{μ}_{max} is a better parameter than X_{max} in terms of identifying the photon-like events as protons.

Consistent with elasticity prediction, it becomes evident that there are less photon-like protons from QGSJetII-04. One could notice that the differences between contours of the same primary are not negligible in case of proton and iron but are nearly identical for photon simulations. EPOS-LHC generates more muons for non-photon events from iron and proton simulations. The reason can be attributed to fundamental differences in how baryons and also rho-mesons are hadronised in EPOS compared to QGSJetII. Still both models have been shown to produce too few muons in all air showers[109]. Shown in Figure 8.3-5 and Figure 8.3-6 is the elasticity when the leading particles is a proton / neutron or $\pi^{\pm}$.

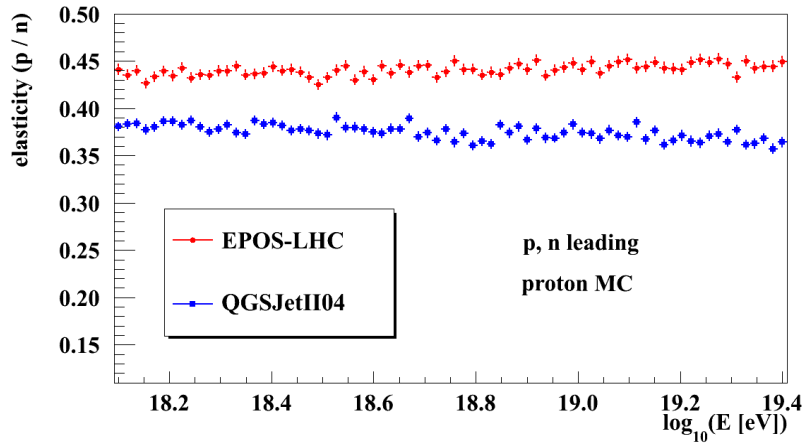


Figure 8.3-5 Elasticity of the leading particle (proton or neutron) of proton simulations

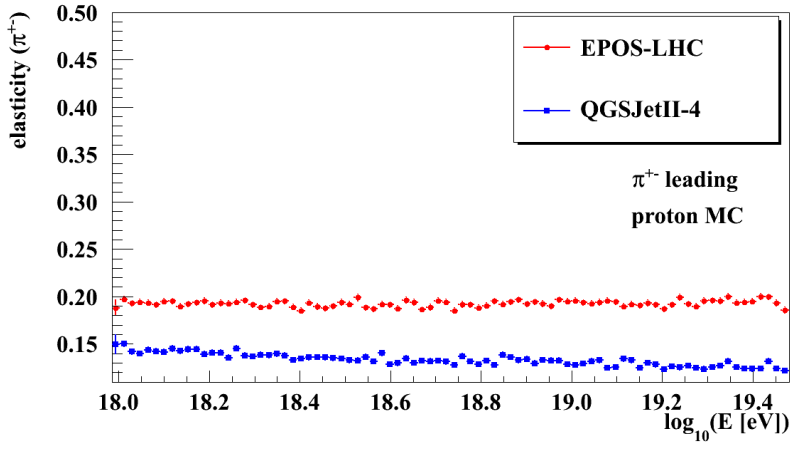


Figure 8.3-6 Elasticity of the leading particle ($\pi^{\pm}$) of proton simulations

However for searches based on SD parameter, the reconstructed energy has to be based on S(1000), which has different biases for proton showers and photon showers (section 7.1). In such cases a proton shower with a generated energy $10^{18.3}$ eV could be reconstructed with energy on the photon scale E^{γ}_{rec} ~ $10^{18.7}$ eV, which makes the differences between proton and photon even smaller.

8.4 Ground based variables of π^0-leading showers from proton simulations

To obtain observables that are measured by the SD, a full 3D simulation is needed. Thus the information on particles from the interactions is passed from the CONEX programme to CORSIKA. Full simulations were made for photon-like proton showers on the thinning level 10^{-6}. Then every shower was dropped 5 times on the ground at different core position of the array. In the end there are 49 showers that survive all cuts with energies 10^{19} – $10^{19.5}$ eV. The distribution of the entity likelihood ratio is shown in Figure 8.4-1.

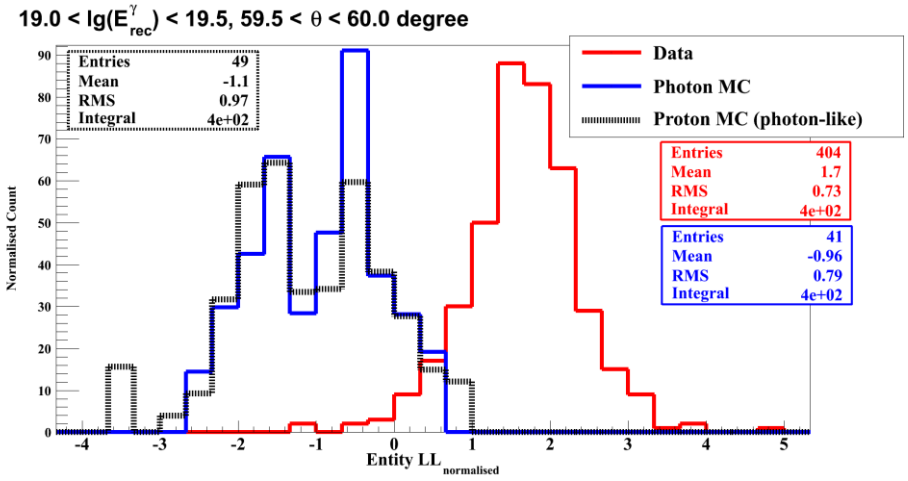


Figure 8.4-1 Entity likelihood ratio for data, photon simulations and proton simulations that have neutral pions as the leading particles

It is clear that it is not possible to distinguish photon-like protons from photons using the entity method. Another attempt is to use risetime which is shown in Figure 8.4-2. Photon showers are expected to have larger risetime than proton showers for a detector at the core position r. The red points are from data and the blue triangles are from photon simulations. The black squares are protons simulations with π^0 leading and elasticity larger than 80%.

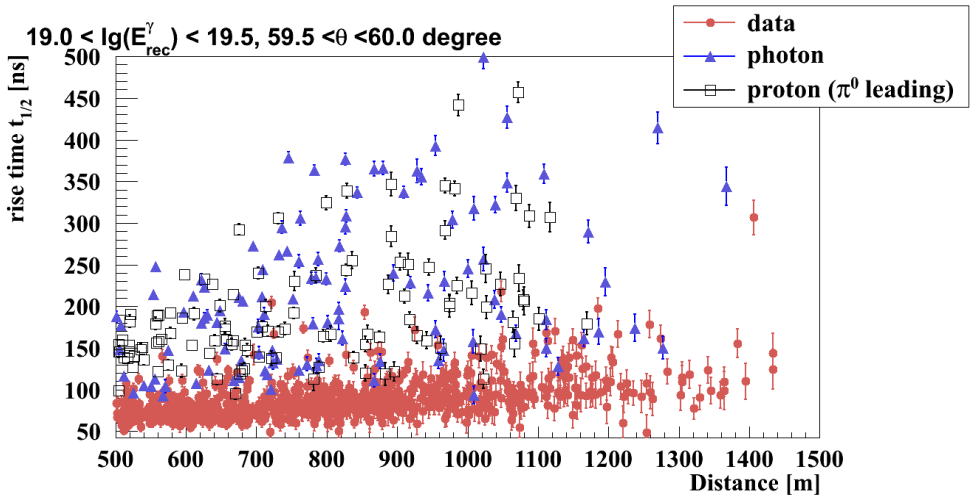


Figure 8.4-2 Risetime of data, photon simulations and photon-like events from proton

Risetime of photon-like protons are high above the mean of the data, which is mixed with the distribution of photon simulations. Therefore the conclusion could be made that at the moment there is no ground-based variable that is able to identify protons with π^0s as leading particles that carry > 80% elasticity.

8.5 Probability and future work on dependences of elasticity

So far it has been demonstrated that when protons have a π^0 as the leading particle that carries more than 80% of the primary energy, they can appear like photon candidates. The probability for such protons occurring is ~ 2 × 10^{-4} from the EPOS-LHC model and ~ 0.4 × 10^{-4} for the QGSJetII-04 model (Figure 8.5-1).

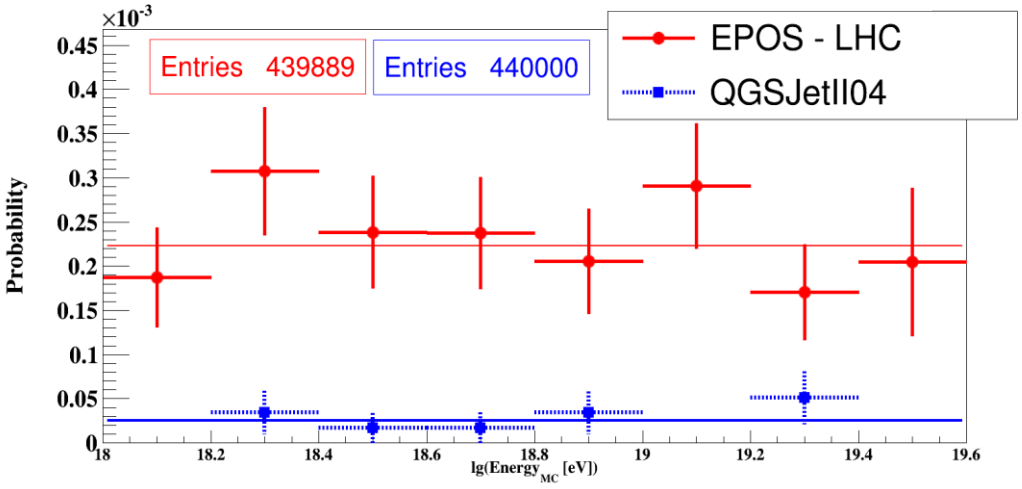


Figure 8.5-1 Probability of have proton showers with leading particle as π^0 and elasticity > 80%

The smaller elasticity the π^0 has, the less fraction of energy of the shower goes to the electromagnetic component and therefore creating a larger number of muons which makes the shower less photon-like. The relationship between N^{μ}_{max} and elasticity is shown in Figure 8.5-2.

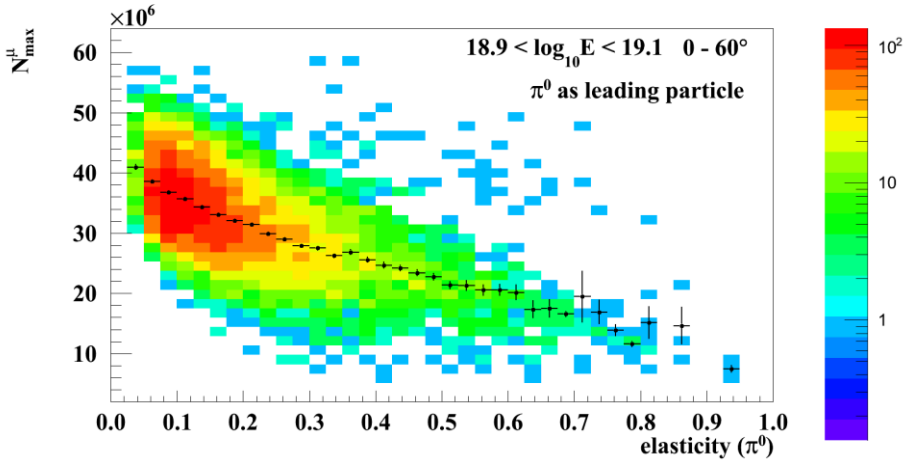


Figure 8.5-2 N^{μ}_{max} and elasticity of π^0 (as the leading particle in the first interaction)

The mean of N^{μ}_{max} at an elasticity of 50% is ~ 2.2 × 10^7, which is lower than the mean from the proton simulations ~ 3 × 10^7 (Figure 8.3-3). The probability of protons having leading π^0s with elasticity larger than 50% is a factor of 10 higher than that of π^0 with elasticity larger than 80%. CORSIKA simulations are needed for a detailed study of the impact of elasticity on the photon-hadron separation. In addition to give a conclusive estimate of the background estimation, selection efficiency and trigger efficiency should be taken into account in the full detector simulations.

In this the chapter, it has been demonstrated with proton simulations with a π^0 as the leading particle in the first interaction that photon-like air showers are created. The probability for such phenomena depends on the elasticity of the π^0, which is sensitive to hadronic interaction models. EPOS-LHC generates more candidates from proton than QGSJetII-04. If the elasticity of π^0 is larger than 80%, nearly all protons are identified as photon candidates. This happens at the rate of ~ 2 × 10^{-4} for the EPOS-LHC model and ~ 0.4 × 10^{-4} for the QGSJetII-04 model, which is at a level well below what we can presently detect.

Chapter 9. Outlook for the entity method

The entity method has been proven to be effective for the photon-hadron separation. However, it might well be helpful to combine the entity method with other parameters for constructing a multivariate analysis for the photon-search. The selection cuts of the entity method are relatively relaxed compared to cuts of risetime or radius of curvature methods so that it is more useful to find a parameter that also requires loose cuts that allows the selection efficiency to stay at the current value. A simple idea is to use the goodness of the fit from the lateral distribution function χ^2/ndf and is shown in the section 9.1. Alternatively for the photon-search based on hybrid data, the entity method could be combined with X_{max} measured from FD (section 9.2), which may contribute to the point-source searches. The entity method could also be used for searching for exotica such as events with two shower profiles (shell events). This is going to be introduced in the section 9.3. The end of the chapter is the discussion of impacts of statistical thinning effects (the concept is introduced in section 2.4).

9.1 Combining the entity parameter with LDF

The lateral distribution function and its role in photon-hadron separation has been introduced with details in section 3.2.3. In this exercise we use the value of $\log_{10}\chi^2$/ndf from LDF fittings and combine it with the entity parameter.

Events with more than 3 stations triggered, reconstructed $\theta > 50^\circ$ and $10^{19} < E^{\gamma}_{rec} < 10^{19.2}$ eV are selected from proton and photon simulations. The photon-hadron separation only based on $\log_{10}\chi^2$/ndf is shown in Figure 9.1-1 and the separation based on the entity method is shown in Figure 9.1-2. The Merit factor using the $\log_{10}\chi^2$/ndf from fitting LDF is 0.5 and the Merit factor for the entity method is 2.3. The Merit factor of the entity method is lower than what is shown in Chapter 6 because this exercise is based on proton simulations instead of data.

The correlation of $\log_{10}\chi^2$/ndf from the LDF fit and the entity likelihood ratio is shown in Figure 9.1-3. A typical way of defining the parameter that is based on two variables is to determine a Fisher axis, which has the largest separation between protons and photons after all points are projected on the

axis. The dashed line in Figure 9.1-3 is the Fisher axis. The Fisher response, which is the projected value on the Fisher axis could be expressed as

$$\text{Fisher response} = p_0 + p_1 x + p_2 y \qquad 62.$$

x is the entity likelihood ratio and y is the $\log_{10}\chi^2$/ndf from LDF fitting.

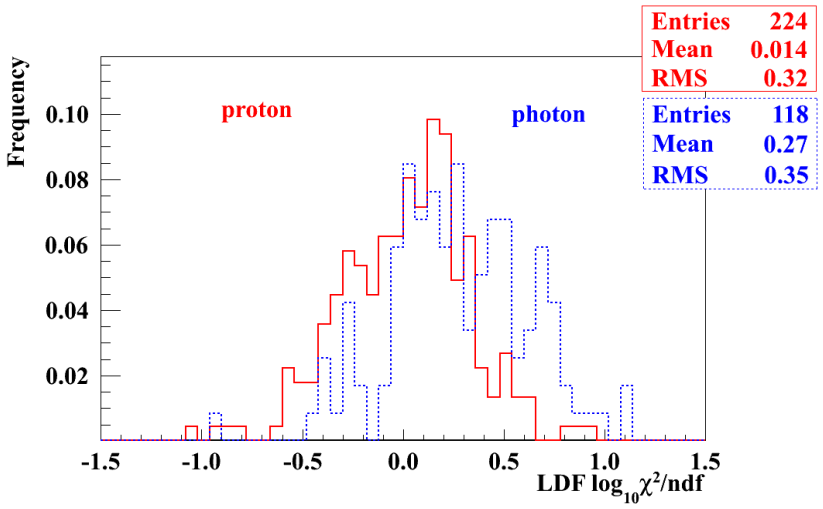


Figure 9.1-1 Distribution of $\log_{10}\chi^2$/ndf from the LDF fit of proton and photon

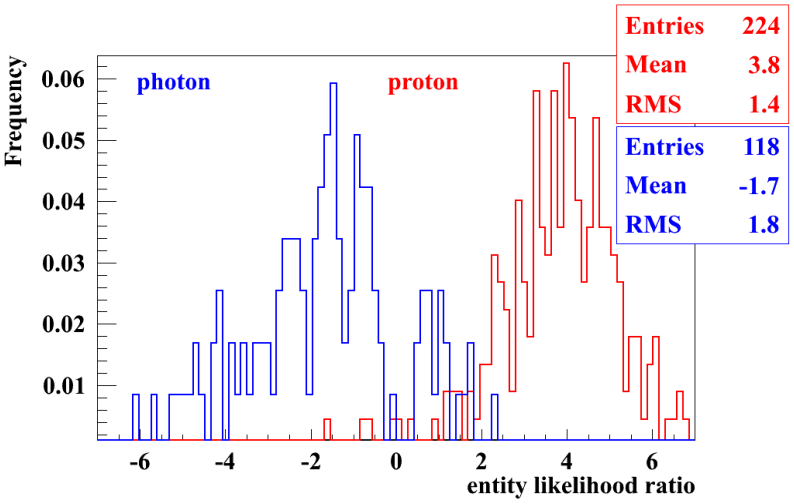


Figure 9.1-2 Distribution of the entity likelihood ratio of proton and photon

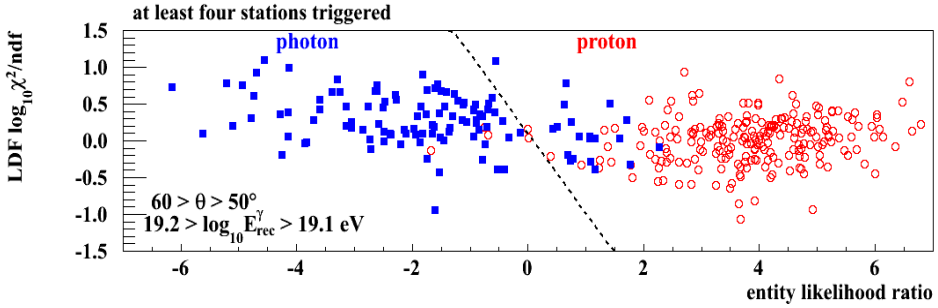


Figure 9.1-3 Relationship of $\log_{10}\chi^2$/ndf from LDF fits and the entity likelihood ratio. To achieve the maximum separation, all points are projected onto the Fisher axis, which is the dashed line.

The Fisher response of proton and photon simulations is shown in Figure 9.1-4. The merit factor is ~2.5, which is higher than the merit factor based on one parameter.

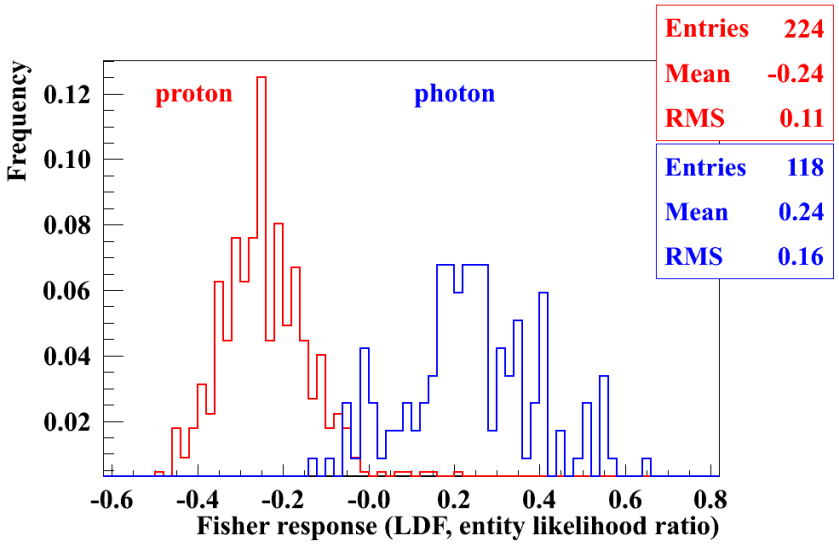


Figure 9.1-4 Fisher response of the combined analysis of LDF and the entity method. Only showers that have $\theta > 50^\circ$ and $E^{\gamma}_{rec} \sim 10^{19.1}$ eV are selected.

In principle an independent set of simulations should be used for calculating the Fisher axis. Alternatively one could use an event likelihood ratio that based on the LDF and the entity method. The signal information carried by LDF complements the trace information carried by the entity method. More should be explored in the future.

9.2 Combining the entity parameter with X_{max}

The entity method could also be extended to search for lower energy photons since it is possible to calculate the entity parameter even when there is only one station selected. The hybrid analysis requires the shower detected by at least one FD telescope and one surface detector. To apply the entity method for hybrid search, a simple test has been carried using data taken from Jan. to Aug. 2012. Only showers that survived the standard FD cuts (section 3.1) and the entity cut (except no minimum request on the number of stations) are kept. Events have zenith angle between 0 - 60° and energy from $10^{18.8}$ to 10^{19} eV. The distribution of X_{max} is drawn in Figure 9.2-1 and the distribution of the entity is in Figure 9.2-2. The relationship between X_{max} and the entity likelihood ratio is shown in Figure 9.2-3. The Fisher axis is the black dotted line, which allows the maximum photon-data separation once all points are projected on it. Figure 9.2-4 is the Fisher response as the combined parameter.

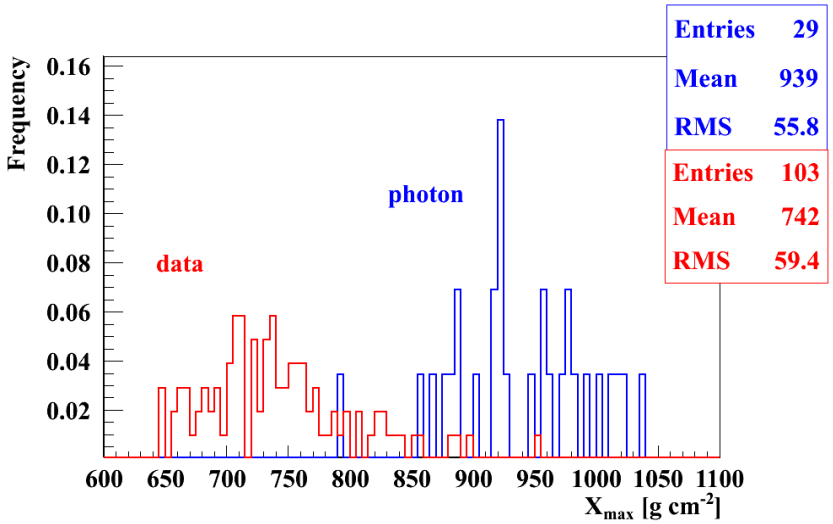


Figure 9.2-1 Distribution of X_{max} of the selected photon simulations and events

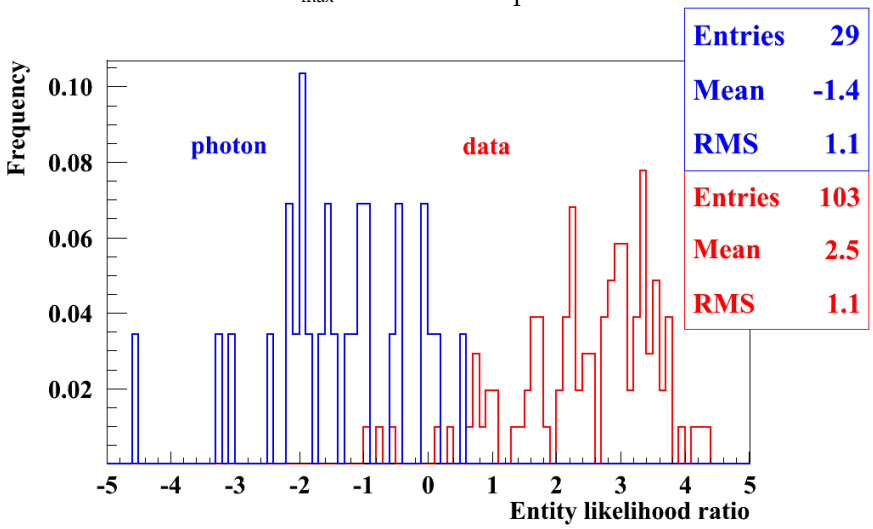


Figure 9.2-2 Distribution of the entity likelihood ratio of selected photon simulations and data

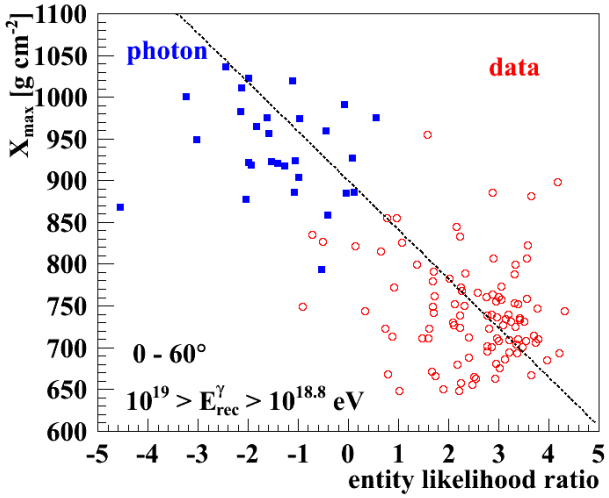


Figure 9.2-3 X_{max} and the entity likelihood ratio of photon simulations and data. The dashed line is the Fisher axis as defined in Figure 9.1-3.

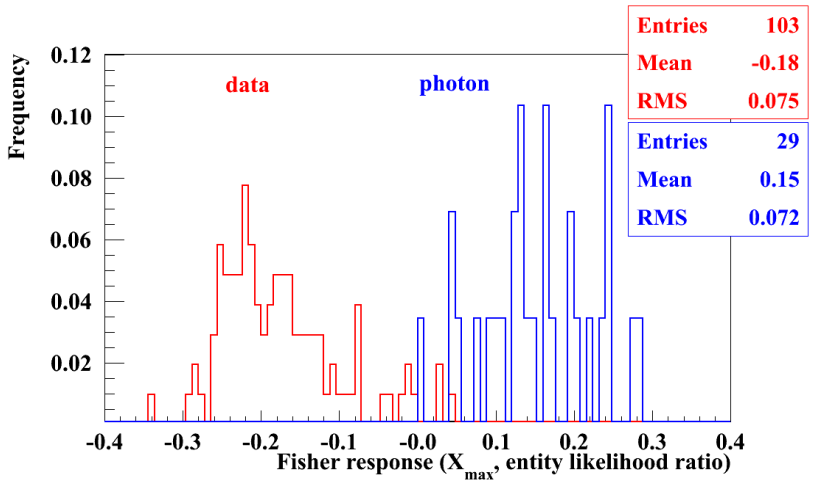


Figure 9.2-4 Fisher response as the combined parameter of X_{max} and the entity method

The merit factor of photon-hadron separation based on X_{max} and the entity likelihood ratio is 2.4 and 2.5 respectively. The X_{max} distribution seems to have a long tail which introduces photon candidates from data. The entity method on the other hand does not give any photon candidates if one sets the cut on the median of the distribution from photon simulations. The Fisher response has the merit factor ~ 3.2, which has been improved significantly. To conclude this simple test, the entity method could be applied to the photon search based on hybrid data and increase the separation power dramatically.

9.3 Applications for searching for exotic events

Previously in the section 6.4.2, it was shown that the entity method could be used to detect problems of PMTs of stations. In addition, the entity method could be used to search for exotic particles other than photons, such as to search for showers with the shower profile that have double-bumps[120]. Showers with double-bumps are also known as 'shell' events. If identified, they have the potential to help study hadronic interaction parameters such as cross-section.

An example of a helium shower with such profile is shown in Figure 9.3-1 (left). The helium has energy 10^{17} eV and the first interaction is at ~ 19 g cm^{-2}. The shower develops and reaches the maximum at ~ 700 g cm^{-2}. However, a 'spectator' nucleon that carries 25% of the primary energy of the helium travelled deep into the atmosphere and interacted at ~ 583 g cm^{-2}, which leads to the second peak of the profile at ~ 1200 g cm^{2}. The probability of

showers with double-bumps for proton, helium and iron that have different energies is shown in Figure 9.3-1 (right).

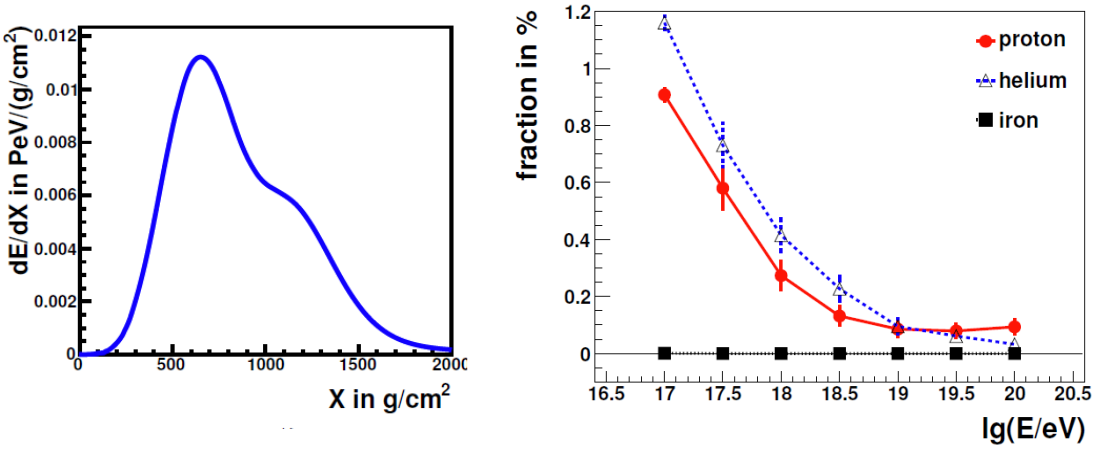


Figure 9.3-1 An example event with the shower profile that has two peaks (left) and the probability of such event to exist for p, He and Fe showers with various energies (right)[120]. There is no double-peak event from iron regardless of the energy.

The cross-section increases with energy so that the mean free path decreases. Therefore the separation between two peaks is less noticeable when the energy of the primary particle is higher. For protons with energy 10^{19} eV, the chance of having a shower with double bump is 0.1%. 'Spectator' nucleons only carry ~ 2% of the primary energy (1/56), which makes it difficult to have the double-bump feature.

Particles that were created at higher in the atmosphere arrive the ground early, which makes it possible to have FADC signals spread in two separated time periods.

Events with such features might be found using the entity method: an example event (13560169) shown in Figure 9.3-2 has multiple stations with double peaks. The event has triggered five stations and has $\theta = 54.5^\circ$. FADC traces of station 180, 668 and 124 all have more than one peak. The timing position and the production depth[121] are listed in Table 9.3-1.

Table 9.3-1 The peak position and production depth of the peaks of the FADC traces from the 'shell-like' event that was found using the entity method

ID	Peak time	Peak 1 muon	Peak 2 time	Peak 2 muon
180	37 ns	378 g cm^{-2}	487 ns	1476 g cm^{-2}
668	144 ns	364 g cm^{-2}	1244 ns	1204 g cm^{-2}
124	133 ns	368 g cm^{-2}	1183 ns	1378 g cm^{-2}

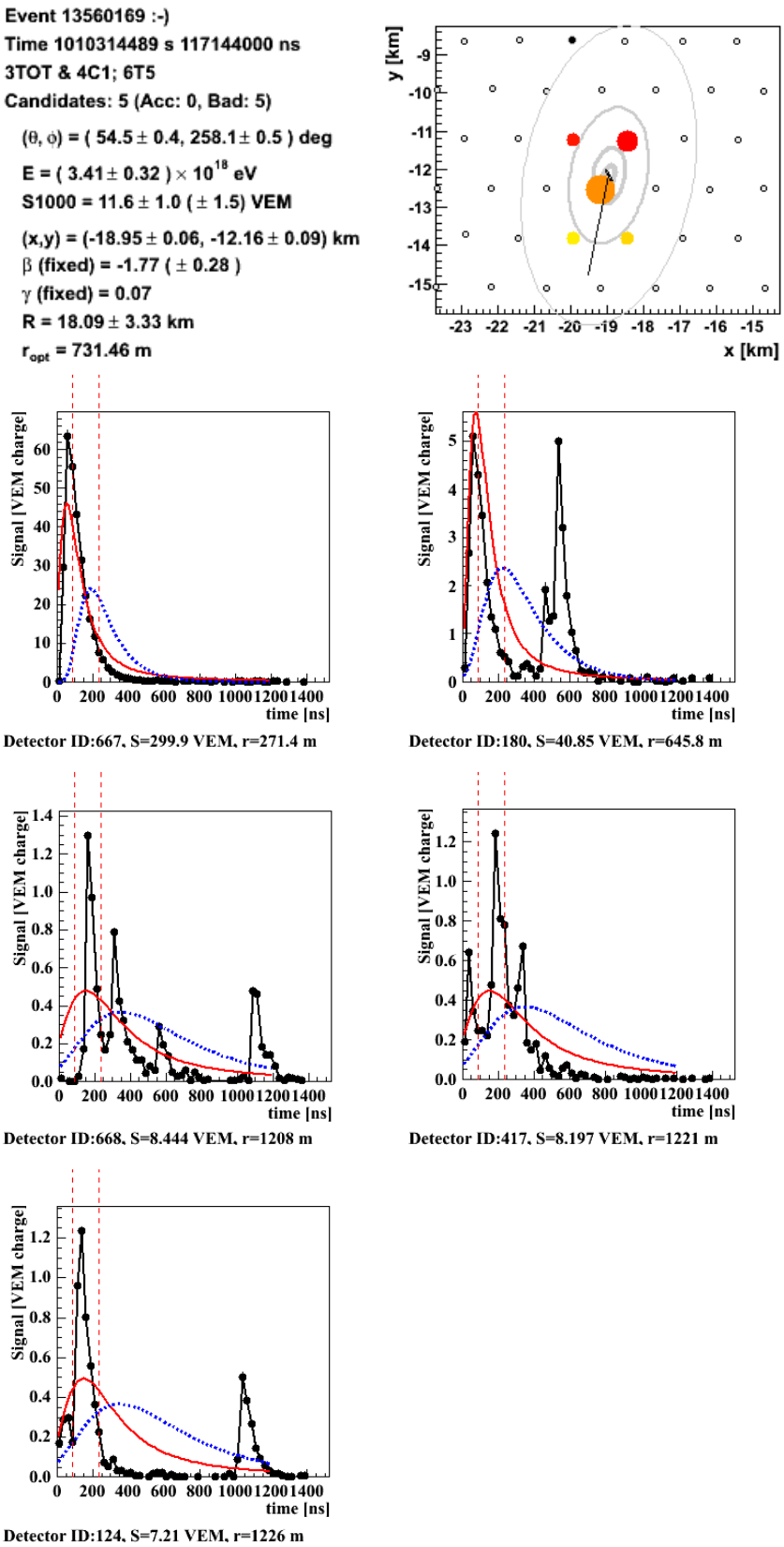


Figure 9.3-2 Event that has FADC traces of three tanks with the 'double-bump' feature. The red line is the mean from data and the blue line is the mean from photon simulations. The signals between dashed lines are used for the entity method.

The time difference between two peaks of station 180 is ~ 450 ns, for station 668 is ~ 1100 ns and for station 124 is ~ 1050 ns. The production depth of the two peaks has a distance of ~ 1000 g cm^2. It is evident that there are two streams of showers arriving on the ground for this event.

In the future, studies that are designed to search for shell events can benefit from the entity method.

9.4 Thinning effect and FADC traces from unthinned simulations

It was stated in section 2.4 that the thinning and de-thinning procedures could create artificial fluctuations. In this section, we compare FADC traces obtained in the same station of the same photon shower but simulated with different thinning levels. Next by calculating $\log_{10}\chi^2$/ndf of the trace-fitting from three unthinned showers of proton, photon and iron primary, we select photon-like events from hadron showers. It is clear that further studies are needed.

9.4.1 Can we trust thinned showers?

As explained in section 2.4, the sampling area on the ground has a fan shape and covers an area around the detector. In Monte Carlo simulations, it is common to simulate a ring of detectors at a fixed distance from the core position. However, to avoid the overlapped sampling area of stations, it is required to place a maximum of 20 detectors in one ring.

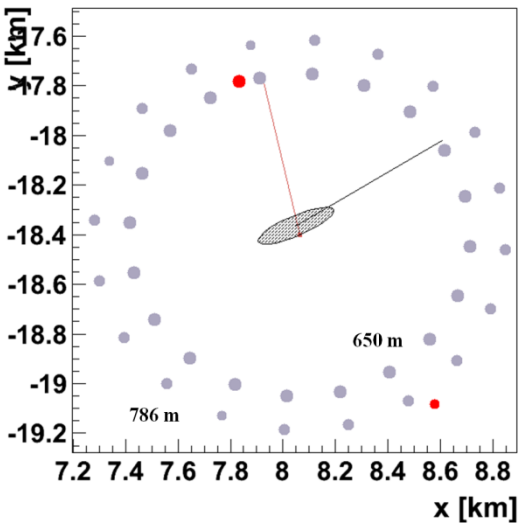


Figure 9.4-1 Distribution of simulated stations for a photon shower simulated at different thinning levels. There are 20 stations at 650 m and 20 stations at 786 m.

One photon shower with the initial energy 10^{19}eV and the zenith angle 0˚ has been simulated but saved with thinning level 10^{-6} (particles with energy larger than 10^{13} eV are tracked), 10^{-8} (particles with energy larger than 10^{11} eV are saved) and without thinning procedures applied. Stations are simulated at two rings surrounding the core with distance r = 650 m and r = 786 m and with 20 tanks per ring. In Figure 9.4-1, only tanks marked in grey are used in this study, stations in red are located in the regular array and not accounted for in the following calculations.

The first comparison is the mean of the time-trace. The shower is vertical which minimizes the effects of polar angles. Shown in Figure 9.4-2 and Figure 9.4-3, the averaged mean time-trace of stations at 650 m and 786 m are in blue, black and red to represent the shower that is unthinned, with thinning level 10^{-8} and thinning level 10^{-6}.

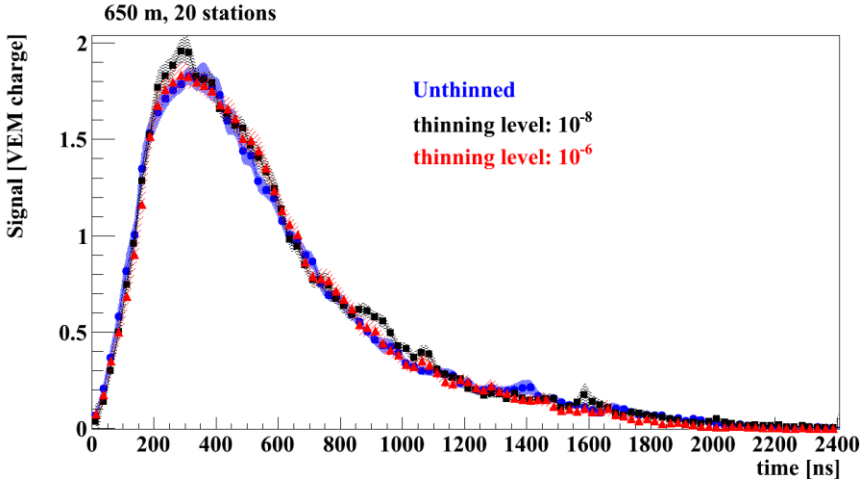


Figure 9.4-2 Mean time-trace from 20 ring stations at 650 m for the same photon shower with different thinning levels

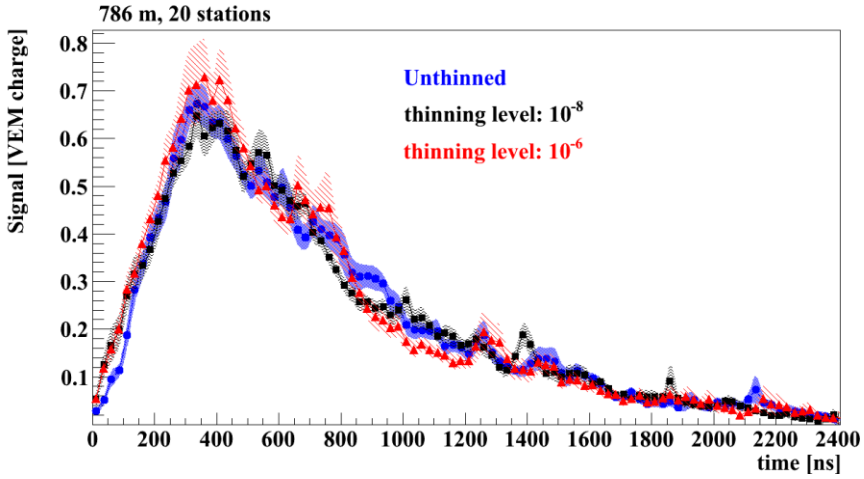


Figure 9.4-3 Mean time-trace from 20 ring stations at 786 m for the same photon shower with different thinning levels

It could be seen that the averaged time-traces of thinned showers agree with what was calculated from unthinned showers within uncertainties. The fluctuations of traces at 786 m are larger than at 650 m which is simply due to the smaller number of particles at larger distances from the shower core. The risetime $t_{1/2}$ of the traces are compared (Figure 9.4-4).

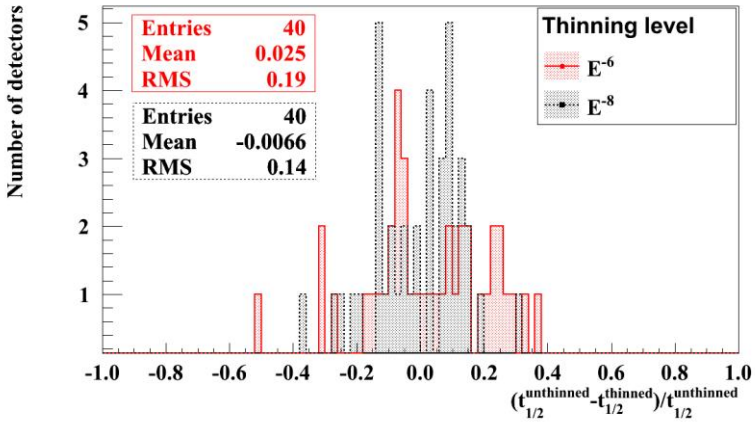


Figure 9.4-4 Differences of risetime for thinned and unthinned shower

As expected, the 10^{-6} thinning level has larger biases in the risetime than the thinning level at 10^{-8}. More statistics are required to give a more conclusive result. So far it turns out the averaged behaviour of FADC traces can be represented well by showers with thinning level 10^{-6}. The $\log_{10}\chi^2$/ndf (defined in section 5.3) of the entity method describes how well the shape of the time-trace agrees with data and are calculated for 40 detectors at different thinning levels (Figure 9.4-5).

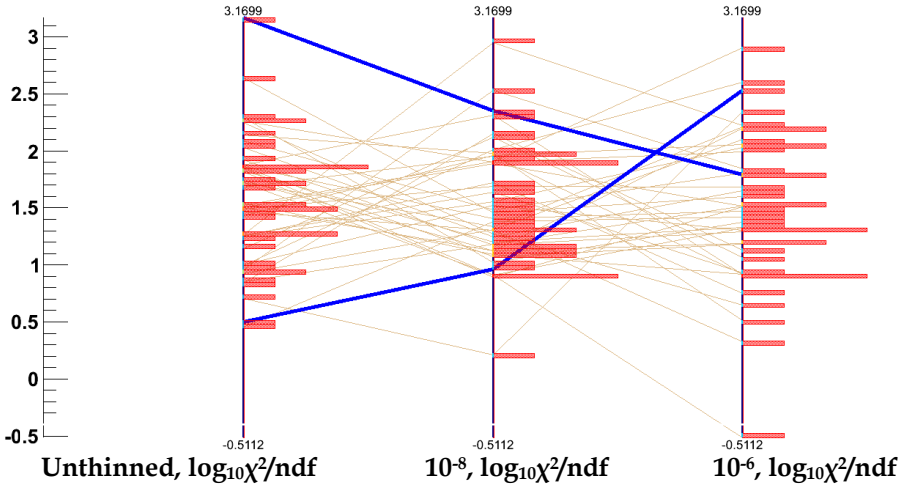


Figure 9.4-5 $\log_{10}\chi^2$/ndf of the same station from a shower at different thinning levels. The blue lines show the two stations with extreme values of $\log_{10}\chi^2$/ndf of the unthinned shower but the values are completely different in the case of thinned showers.

The mean of $\log_{10}\chi^2$/ndf agrees within uncertainties for each thinning level. The spread of the distribution is the largest for the 10^{-6} thinning level. However, in terms of the FADC trace at the individual station level, the results from thinned simulations should not be trusted. An example using station 91006, which has $\log_{10}\chi^2$/ndf ~ -0.5 for the 10^{-6} thinning level, is shown in Figure 9.4-6.

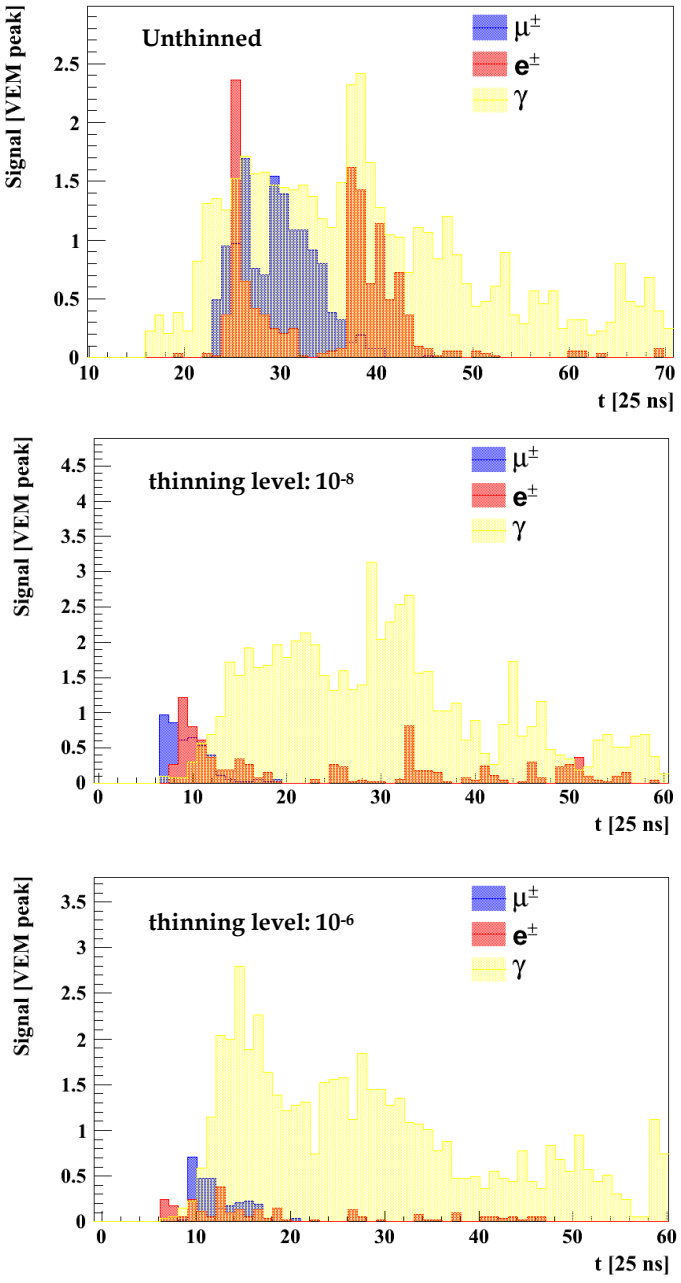


Figure 9.4-6 Component traces of detector 91006 recorded at different thinning levels. The signal and time information of particles are lost in the thinning showers

The distribution of the momentum of particles is shown in Figure 9.4-7. The y-axis shows the number of particles that contribute to the signals measured in the SD station.

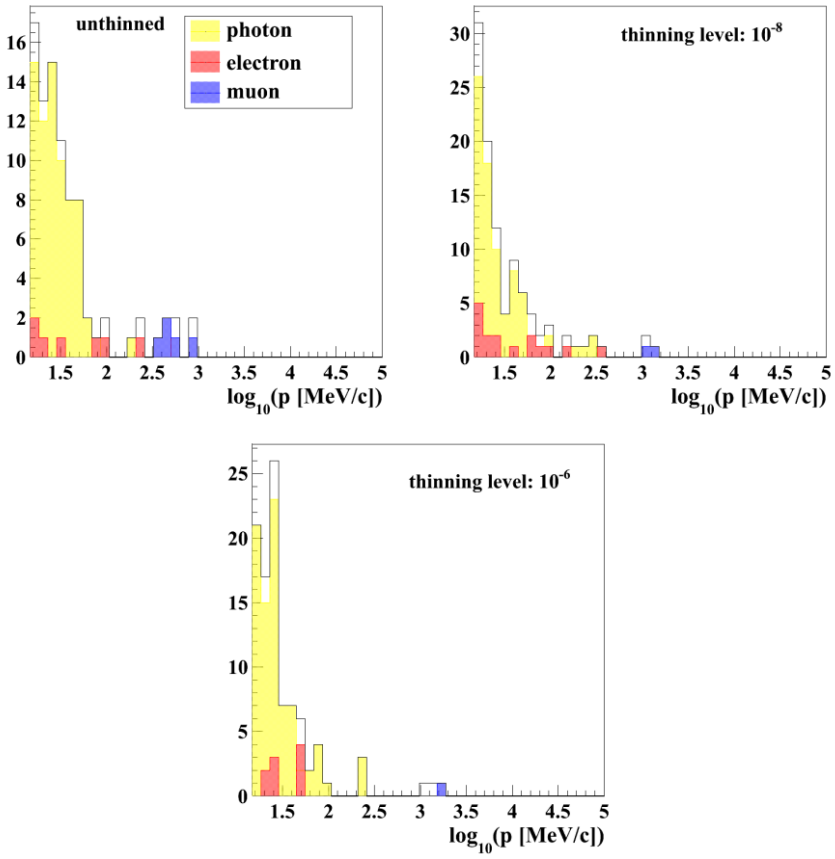


Figure 9.4-7 Momenta of particles in the station 91006 with different thinning levels. The y-axis shows the number of particles. For the thinned showers, the number of particles is estimated according to the weight the particle carries.

The distributions of momenta of particles are very different. During the de-thinning process, the weight of the particle is interpreted as the number of the particle with the same momentum. A large weight creates a huge number of particles with the same momentum instead of naturally distributed momenta. The shower with higher thinning level is likely to have particles piling up with the same momentum, which creates artificial signals in the station.

As a conclusion of the section, the thinning and de-thinning procedures affects the energy distribution and time information of particles arriving at a station. One should not in principle believe the FADC trace of a station from the thinned simulations. However for variables that are based on the mean of showers such as the mean of the FADC trace, we find agreements between thinned and unthinned shower. Next section is about comparing the $\log_{10}\chi^2$/ndf of the trace-fitting between the unthinned proton, photon and iron simulations.

9.4.2 $Log_{10}\chi^2$/ndf of the trace-fitting for unthinned simulations of different primaries

Three unthinned showers are used in this exercise: a proton, an iron and a photon shower all with $\theta = 0^\circ$ and Monte Carlo energy $E_{MC} = 10^{19}$ eV. Unlike the thinned showers, the sampling area of the unthinned shower is only the size of area that the shower projected in the station. This makes it possible to simulate more than 20 stations without worrying about overlapped sampling areas. Two-hundred stations at 650, 700 and 750 m were simulated at each distance respectively for all three primaries. The distribution of $\log_{10}\chi^2$/ndf from comparing the trace with the expectation from data is shown in Figure 9.4-8.

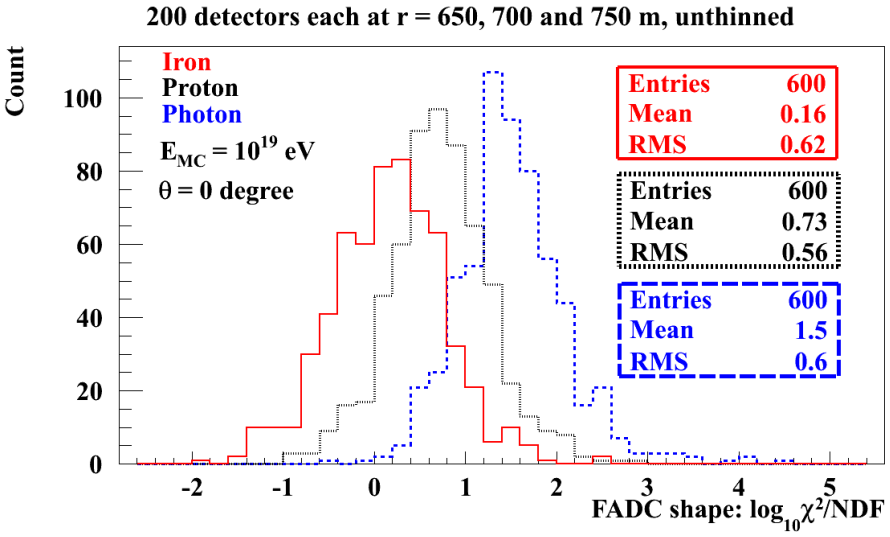


Figure 9.4-8 $\log_{10}\chi^2$/ndf from comparing the FADC traces to mean traces of data for unthinned proton, photon and iron simulations

To make fair comparisons with data, 638 stations were selected from Auger events with $0 < \theta < 10^\circ$ and reconstructed energy on the hadronic scale $10^{18.5} < E_{rec} < 10^{19.1}$ eV. Photon simulations which have been used to set the limit (section 6.2) are with thinning level 10^{-6}. 87 stations were selected from photon showers that have $10^{18.8} < E_{MC} < 10^{19.2}$ eV and $0 < \theta < 20^\circ$. All stations are located between 650 – 750 m from the core position. Values of $\log_{10}\chi^2$/ndf are shown in Figure 9.4-9.

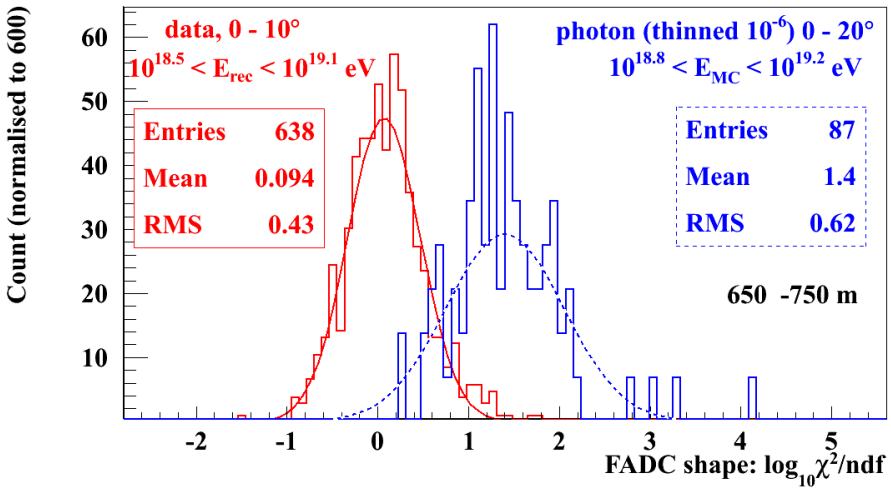


Figure 9.4-9 $\log_{10}\chi^2$/ndf from the FADC trace of data and photon simulations (thinned)

The unthinned iron shower is most data-like and the proton shower is between the iron and photon. FADC traces of stations from unthinned simulations with $\log_{10}\chi^2$/ndf > 1.5 were checked and an example of the detector from iron shower is shown in Figure 9.4-10, from the proton shower in Figure 9.4-11. Three examples were taken for the photon shower because the different signature of the trace as shown in Figure 9.4-12, Figure 9.4-13 and Figure 9.4-14.

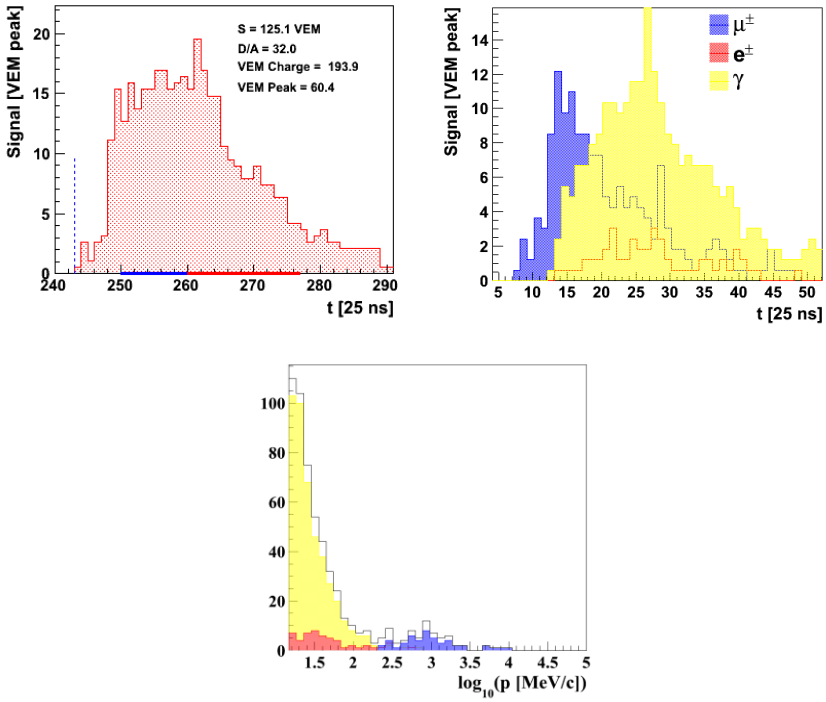


Figure 9.4-10 Example event from the iron unthinned shower that have $\log_{10}\chi^2$/ndf > 1.5 . The FADC trace is shown on the top left corner. The component trace is shown on the right top corner. The distribution of momenta of different particles is shown in the bottom with the y-axis as the number of particles. The small signal at the beginning of the trace is caused by muons.

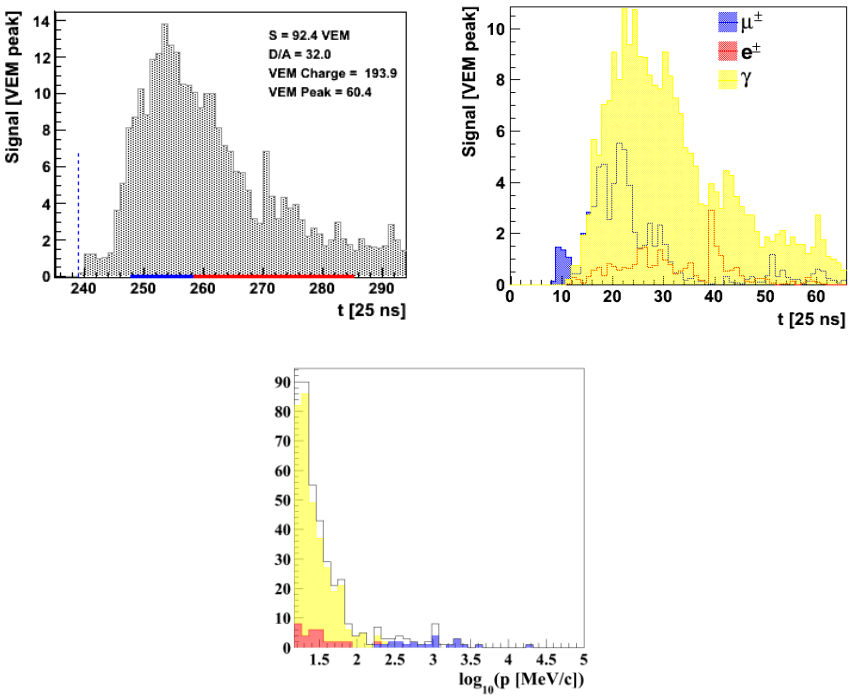


Figure 9.4-11 Example event from proton unthinned shower that has $\log_{10}\chi^2/\text{ndf} > 1.5$. The pre-pulse is due to the muon signal.

The pre-pulse signature has been discussed in section 6.5 and an example of the FADC trace from data with such feature is shown in Figure 6.4-7. One needs to estimate the frequency that the pre-pulses occur at the event level when only 2 or 3 stations are selected per event to calculate the hadronic background for the photon search.

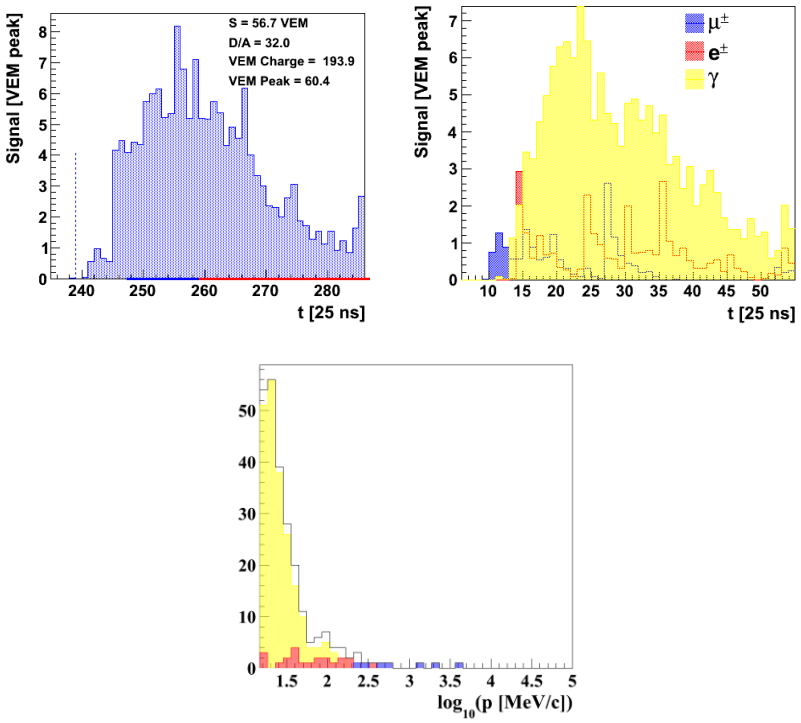


Figure 9.4-12 Example event from photon unthinned shower that has $\log_{10}\chi^2/\text{ndf} > 1.5$. The pre-pulse is due to the muon signal.

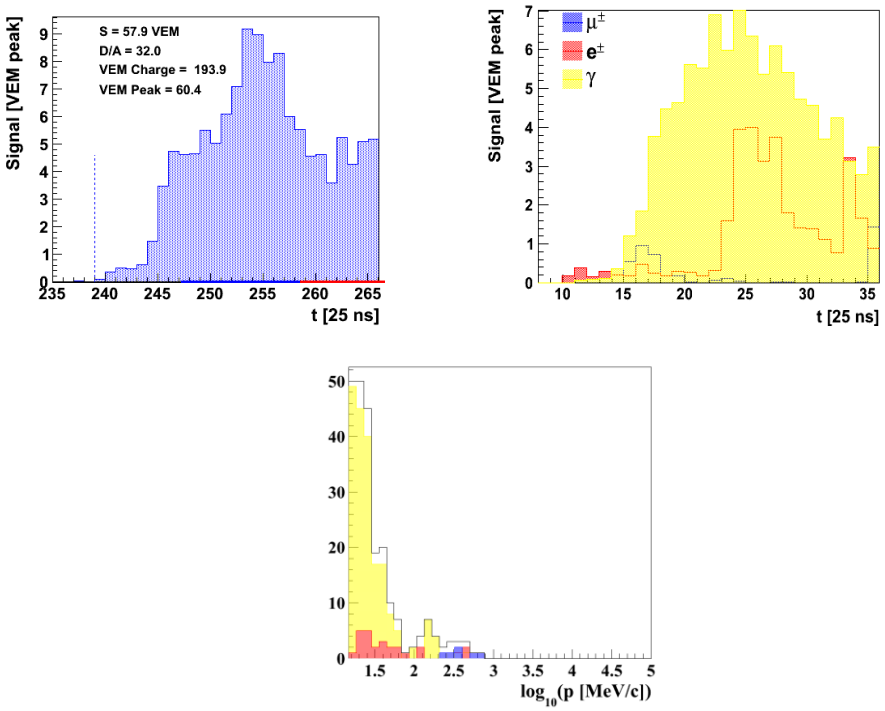


Figure 9.4-13 Example event from photon unthinned shower that has $\log_{10}\chi^2/\mathrm{ndf} > 1.5$. The pre-pulse is due to the electron signal.

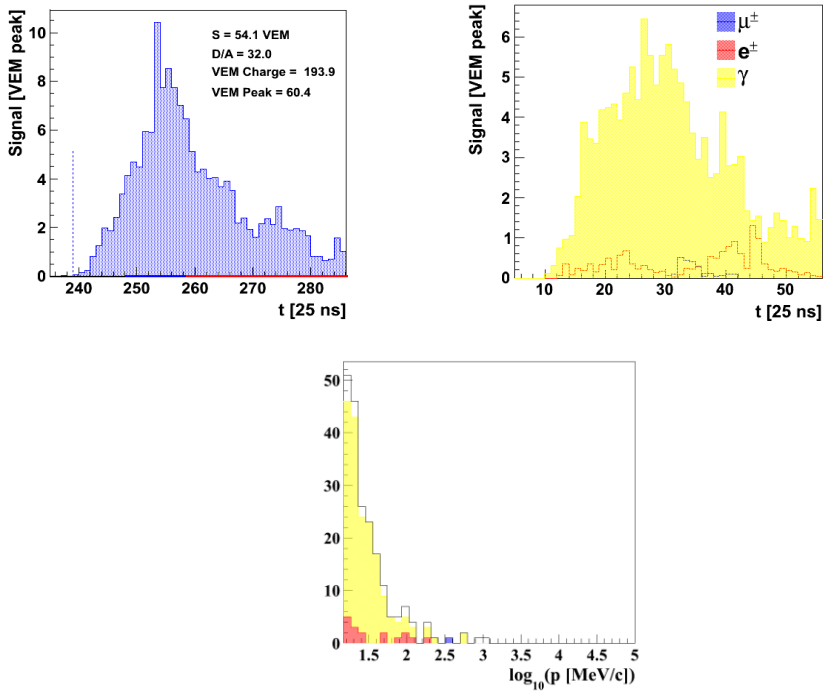


Figure 9.4-14 Example event from photon unthinned shower that has $\log_{10}\chi^2/\mathrm{ndf} > 1.5$

Pre-pulses due to the muon signal could be due to an early muon that were produced at the top of the atmosphere and arrive earlier. Unfortunately the storage space needed for an unthinned shower is too huge that information of the source of the muon was not saved.

Interestingly the last example from the photon shower does not have a pre-pulse but just a slowly rised FADC trace, which has not been seen in hadronic simulations.

One thing we should keep in mind is that the $\log_{10}\chi^2/\mathrm{ndf}$ is a station variable. The event variable – the entity likelihood ratio combines all values of $\log_{10}\chi^2/\mathrm{ndf}$ of stations into one variable, which increases the chance of identifying photons. As there are only three showers, it is not feasible to compare the entity likelihood ratio for unthinned simulations at the moment. It is important to produce a library with a decent number of unthinned showers to make conclusions. One perhaps needs to rely on grid simulations, which is at the moment scheduled with productions for beyond 2015 (Auger upgrading).

This chapter summarises the outlook of the entity method. The method has an important role in photon-hadron separations but is also useful for searching for exotic particles. The effects of thinning are problematic for FADC traces of stations, however do not shift the mean of distribution of variables such as the risetime and the $\log_{10}\chi^2/\mathrm{ndf}$ of the entity method. In the future more work is needed to explore the full potential of the entity method.

References

*ICRC stands for International Cosmic Ray Conference
*GAP note is the internal paper in the Pierre Auger Collaboration

[1] V. Hess, Physik. Zeitschr. **13** 1084 (1912)

[2] P. Auger et al., Rev. Mod. Phys. **11** 288 (1939)

[3] C. D. Anderson, Phys. Rev. **43** 491 (1933)

[4] J. C. Street and E. C. Stevenson, Phy. Rev. **52** 1003 (1937)

[5] D. H. Perkins, Nature **159** 126 (1947)

[6] R. Engel, D. Heck and T. Pierog, Ann. Rev. Nucl. Part. Sci. **61** 467 (2011)

[7] T. Antoni et al., Astropart. Phys. **24** 1(2005)

[8] B. Bartoli et al., Phys. Rev. D **85** 092005 (2012)

[9] W. Apel et al., Phys. Rev. Lett. **107** 171104 (2011)

[10] J. Hoerandel, Astropart. Phys. **21** 241 (2004)

[11] J. Linsley, Proc. 8th ICRC Jaipur **4** 77 (1963)

[12] R. Aloisio, V. Berezinsky and A. Gazizov, Astropart. Phys. **34** 620 (2011)

[13] N. Yanasak et al., Astrophys. J. **563** 768 (2001)

[14] V. Ginzburg and S. Syrovatskii, "The Origin of Cosmic Rays", Pergamon Press and Macmillan Company (1964)

[15] M. Ackermann et al., Science **339** 807 (2013)

[16] E. Fermi, Phys. Rev. **75** 1169 (1949)

[17] R. Protheroe, "Acceleration and interaction of ultra high energy cosmic rays," in Topics in Cosmic-Ray Astrophysics, M. Duvernois, ed., Nova Science Publishing p247 (1999)

[18] G. Krymskii, Dokl. Akad. Nauk. SSSR. **234** 1306 (1977)

[19] A. R. Bell, Mon. Not. R. Astron. Soc. **182** 147 (1978)

[20] A. M. Hillas, Ann. Rev. Astron. Astrophys. **22** 425 (1984)

[21] Kotera and A. Olinto, Ann. Rev. Astron. Astrophys. **49** 119 (2011)

[22] M. S. Longair, "High Energy Astrophysics", 3rd edition, Cambridge University Press p531 (2011)

[23] K. Greisen, Phys. Rev. Lett. **16** 748 (1966)

[24] G. T. Zatsepin and V. A. Kuzmin, JETP. Lett. **4** 78 (1966)

[25] D. Allard et al., JCAP **09** 005 (2006)

[26] M. Ave et al., Astropart. Phys. **23** 19 (2005)

[27] A. Schulz for the Pierre Auger Collaboration, Proc. 33rd ICRC Rio (2013)

[28] K.-H. Kampert et al., Astropart. Phys. **42** 41 (2013)

[29] R. Aloisio et al., JCAP **10** 007 (2012)

[30] W. B. Fretter, Proc. Echo Lake Cosmic Ray Symposium (1949)

[31] F. Schmidt and J. Knapp in Leeds (2005)

http://www-zeuthen.desy.de/~jknapp/fs/photon-showers.html

[32] M. S. Longair, "High Energy Astrophysics", 3rd edition, Cambridge University Press p167 (2011)

[33] W. Heitler, "Quantum theory of radiation", 2nd edition, Oxford University Press (1944)

[34] J. Linsley Proc. 15th ICRC Plovdiv **12** 89 (1977)

[35] K. Greisen, Prog. in Cosmic Ray Phys. **3** 1 (1956)

[36] J. Matthews, Astropart. Phys. **22** 387 (2005)

[37] K.-H.Kampert and M. Unger, Astropart. Phys. **35** 660 (2012)

[38] V. de Souza for the Pierre Auger Collaboration, Proc. 33rd ICRC Rio (2013)

[39] T. Pierog et al., arXiv:1306.0121 (2013)

[40] A. Aab et al., arXiv:astro-ph/1409.5083, accepted for publication in Phys. Rev. D (2014)

[41] B. Kegl for the Pierre Auger Collaboration, Proc. 33rd ICRC Rio (2013)

[42] Z. Fodor, S. D. Katz and A. Ringwald, Jour. High Energy Phys. **6** 46 (2012)

[43] P. Bhattacharjee and G. Sigl, Phys. Rep. **327** 109 (2000)

[44] A. Taylor, Lecture note for the Astropart. Phys. school in Peru (2009)

[45] M. Risse and P. Homola, Mod. Phys. Lett., A **22** 749 (2007).

[46] P. Coppi and F. Aharonian, Astrophys. J. **487** L9 (1997)

[47] N. Hayashida et al., Astron.J. **120** 2190 (2000)

[48] V. P. Egorova et al., Nucl. Phys. Proc. Suppl. **136** 3 (2004)

[49] M. Ave, et al., Phys. Rev. Lett. **85** 2244 (2000)

[50] G. Rubtsov et al., Phys. Rev. D **73** 063009 (2006)

[51] J. Abraham et al., Astropart. Phys. **31** 399 (2009)

[52] D. Hooper, A. M. Taylor and S. Sarkar, Astropart. Phys. **34** 340 (2011)

[53] A. Etchegoyen for the Pierre Auger Collaboration, Proc. 30th ICRC Merida (2007)

[54] P. Abreu, et al., JINST **7** 10011 (2012)

[55] C. Williams, et al., Phys. Procedia **37** 1341 (2012)

[56] T. Suomijarvi for the Pierre Auger Collaboration, Proc. 31st ICRC Łódź (2009)

[57] I. Allekotte et al., Nucl. Instrum. Meth. A **586** 409 (2008)

[58] X. Bertou et al., Nucl. Instrum. Meth. A **568** 839 (2006)

[59] M. Aglietta for the Pierre Auger Collaboration., Proc. 29th ICRC Pune **7** 83 (2005)

[60] J. Abraham, et al., Nucl. Instrum. Meth. A **523** 50 (2004)

[61] J. Abraham, et al., Nucl. Instrum. Meth. A **613** 29 (2010)

[62] D. Veberic and M. Roth, GAP 2005-035 (2005)

[63] I. Maris, PhD thesis, Karlsruher Institut für Technologie (2008)

[64] K. Kamata and J. Nishimura, Prog. Theoret. Phys. Suppl. **6** 93 (1958)

[65] M. Ave et al., Nucl. Instrum. Meth. A **578** 180 (2007)

[66] A. M. Hillas, Proc. 11th ICRC Budapest **8** 353 (1969)

[67] D. Newton, J. Knapp and A. A. Watson, Astropart. Phys. **26** 414 (2007)

[68] T. Schmidt, I. C. Maris and M. Roth, GAP 2007-106 (2007)

[69] T. Waldenmaier, J. Blümer and H. Klages, Astropart. Phys. **29** 205 (2008)

[70] M. Nagono et al., Astropart. Phys. **22** 235 (2004)

[71] B. Fick et al., JINST **1** 11003 (2006)

[72] P. Abreu et al., Astropart. Phys. **35** 591 (2012)

[73] C. Meurer et al., Astrophys. Space Sci. Trans. **7** 183 (2011)

[74] D. Küempel, K.-H. Kampert and M. Risse, Astropart. Phys. **30** 167 (2008)

[75] J. Abraham et al., Nucl. Instrum. Meth. A **620** 227 (2010)

[76] D. Heck et al., report FZKA 6019, Forschungszentrum Karlsruhe (1998)

[77] T. Pierog et al., Nucl. Phys. Proc. Suppl. **151** 159 (2006)

[78] T. Bergmann et al., Astropart. Phys. **26** 420 (2007)

[79] T. Pierog et al., Proc. 30th ICRC Merida (2007)

[80] K. Werner, F. M. Liu and T. Pierog, Phys. Rev. C **74** 044902 (2006)

[81] N. N. Kalmykov, S. S. Ostapchenko and A. I. Pavlov, Bull. Russ. Acad. Sci. Phys. **58** 1966 (1994)

[82] N. N. Kalmykov, S. S. Ostapchenko and A.I. Pavlov, Nucl. Phys. B **52** 17 (1997)

[83] V. N. Gribov, Sov. Phys. JETP **26** 414 (1968)

[84] V. N. Gribov, Sov. Phys. JETP **29** 483 (1969)

[85] A. B. Kaidalov and K. A. Ter-Martirosyan, Sov. J. Nucl. Phys. **39** 979 (1984)

[86] A. M. Hillas, Proc. 17th ICRC Paris **8** 193 (1981)

[87] R. Bruijn, J. Knapp and I. Valino, Proc. 32nd ICRC Beijing (2011)

[88] M. Kobal, Astropart. Phys. **15** 259 (2001)

[89] P. Billior, Astropart. Phys. **30** 270 (2008)

[90] R. Bruijn and J. Knapp, GAP 2009-011 (2011)

[91] J. Abraham et al., Astropart. Phys. **27** 155 (2007)

[92] V. Scherini for the Pierre Auger Collaboration, EPJ Web of Conferences **53** 05002 (2013)

[93] L. D. Landau and I. Ya. Pomeranchuk, Dokl. Akad. Nauk. SSSR. **92** 535 (1953)

[94] A. B. Migdal, Phys. Rev. **103** 1811 (1956)

[95] J. Abraham et al., Astropart. Phys. **31** 399 (2009)

[96] M. Settimo for the Pierre Auger Collaboration, Proc. International Conference on the Structure and the Interactions of the Photon Paris (2013)

[97] C. Wileman, PhD thesis, University of Leeds (2008)

[98] A. A. Watson and J. G. Wilson, J. Phys. A: Math. Nucl. Gen. **7** 1199 (1974)

[99] R. Walker and A. A. Watson, J. Phys. G: Nucl. Phys. **8** 1131 (1982)

[100] K. S. C. Mora, PhD thesis, Karlsruher Institut für Technologie (2010)

[101] B. Smith, PhD thesis, University of Leeds (2008)

[102] Heather Cook, MPhil thesis, University of Leeds (2012)

[103] J. Abraham et al., Astropart. Phys. **29** 243 (2008)

[104] N. Krohm and C. Bleve et al., GAP 2012-059 (2012)

[105] J. Linsley and L. Scarsi, Phys. Rev. **128** 2384 (1962)

[106] T. Abu-Zayyad et al., Phys. Rev. D **88** 112005 (2013)

[107] M. Settimo for the Pierre Auger Collaboration, Proc. 32nd ICRC Beijing (2011)

[108] P. Bassi, G. Clark and B. Rossi Phys. Rev. **92** 441 (1953)

[109] B. Kegl for the Pierre Auger Collaboration, Proc 33rd ICRC Rio (2013)

[110] L. Lu and A. A. Watson, GAP 2013-007 (2013)

[111] P. Billor, GAP 2005-074 (2005)

[112] L. C. Boado, PhD thesis, Universidade de Santiago de Compostela (2004)

[113] D. J. L. Bailey, PhD thesis, University of Oxford (2002)

[114] D. Maurel, PhD thesis, Karlsruher Institut für Technologie (2013)

[115] P. Billoir, C. Roucelle and J. Hamilton, arXiv:astro-ph/0701583 (2007)

[116] B. Sarkar et al., Proc. 32nd ICRC Beijing (2011)

[117] G. Gelmini, O. Kalashev, and D. V. Semikoz, JCAP **11** 002 (2007)

[118] G. Maier and J. Knapp, Astropart. Phys. **28** 72 (2007)

[119] D. Heck and T. Pierog, CORSIKA User's Guide, Karlsruher Institut für Technologie (2013)

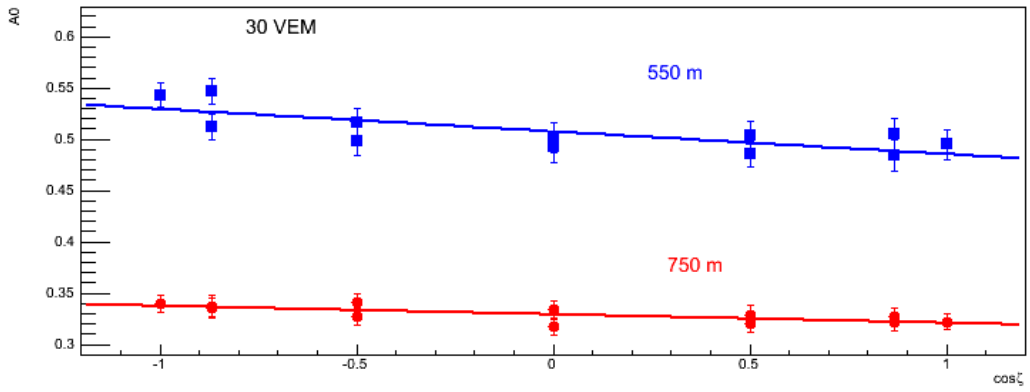
[120] C. Baus et al., Proc. 32nd ICRC Beijing (2011)

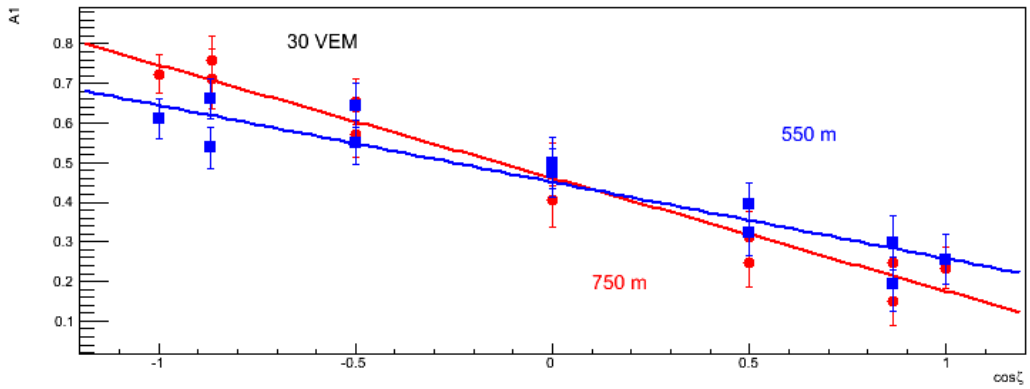
[121] C. E. Santo and A. A. Watson, private communications (2014)

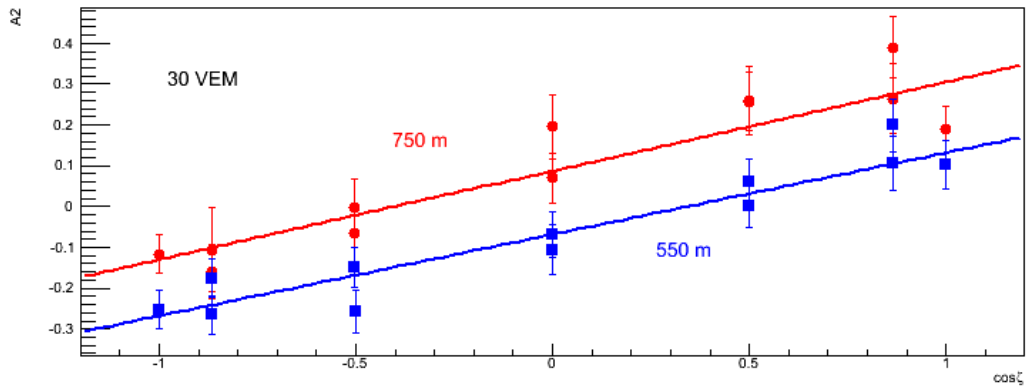
Appendix A

Details of the parameterisation of the mean time-trace:

A_0, A_1 and A_2 vs cos ζ



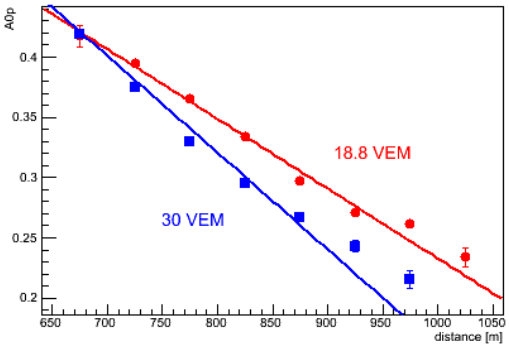


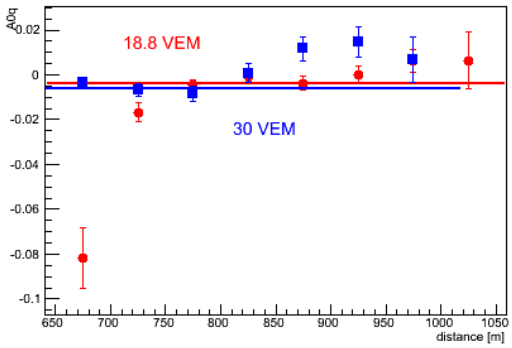


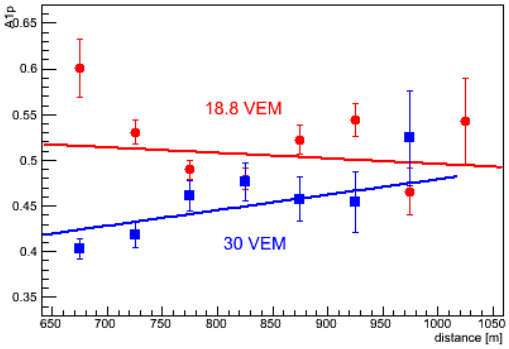
The fitted intercepts and slopes are

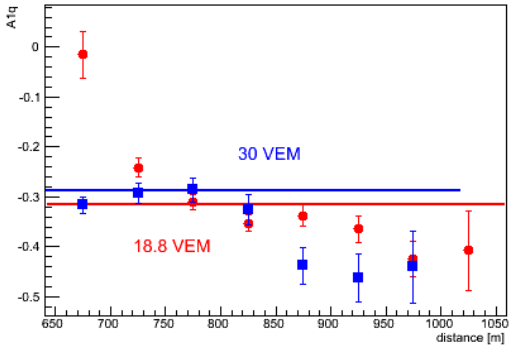
A0p, A0q. A1p, A1q. A2p, A2q

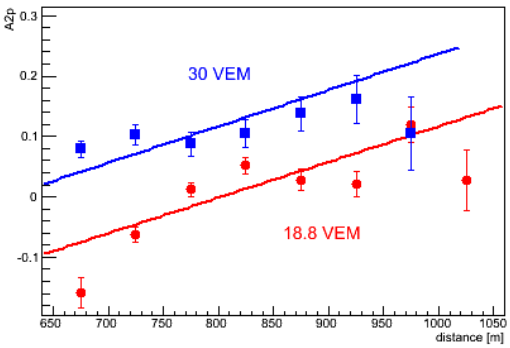
A0p, A0q, A1p, A1q, A2p, A2q vs distance

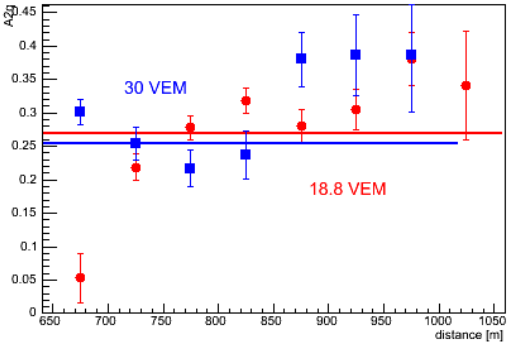




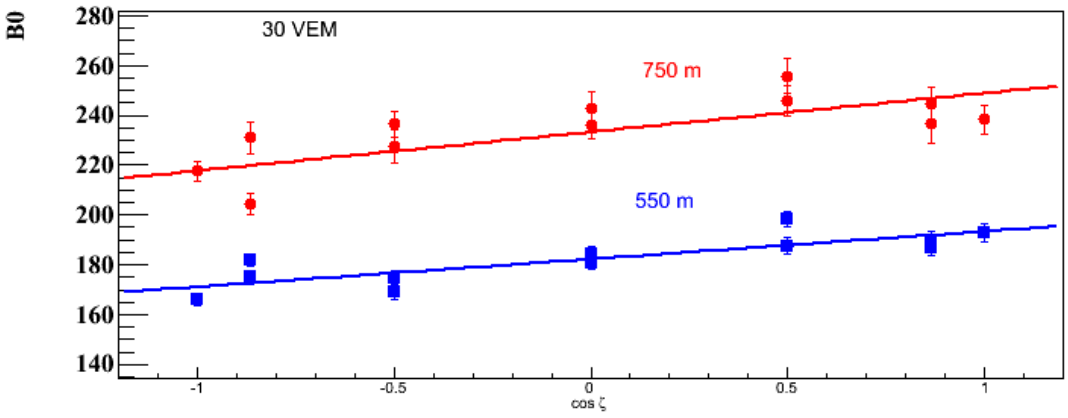


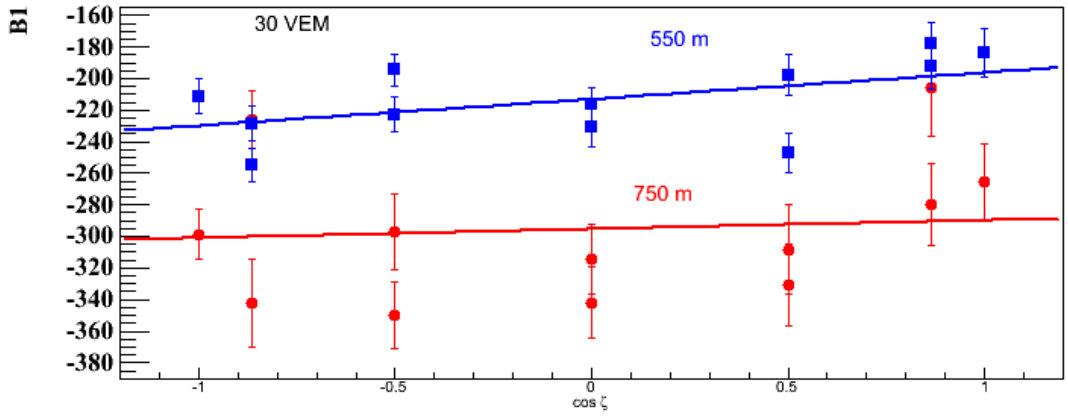


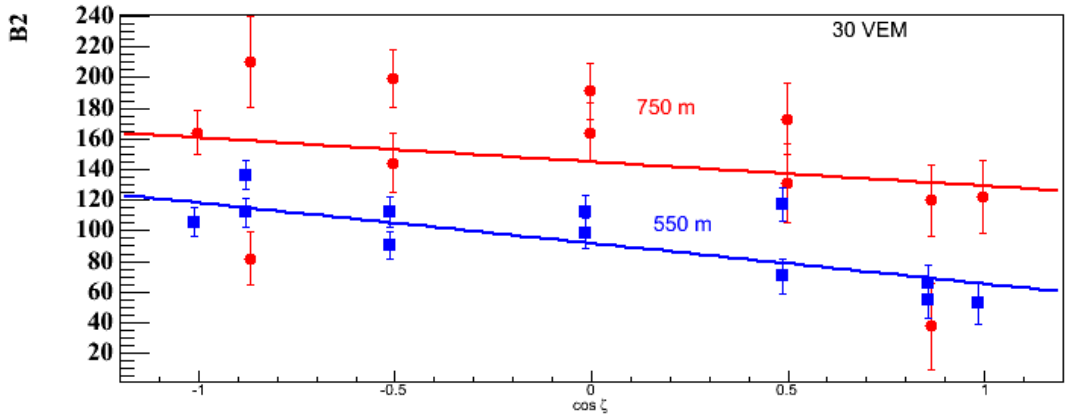




B_0, B_1 and B_2 vs cos ζ







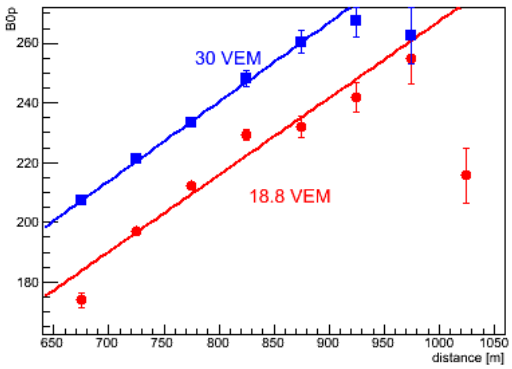
The fitted intercept and slopes are

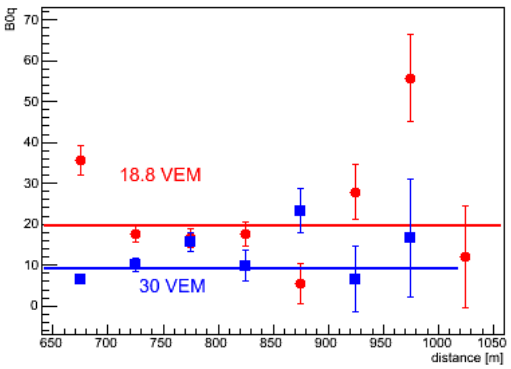
B0p, B0q

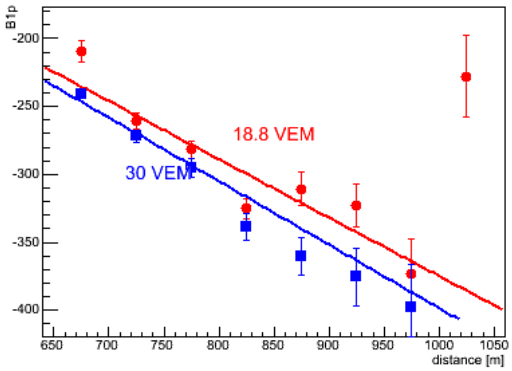
B1p, B1q

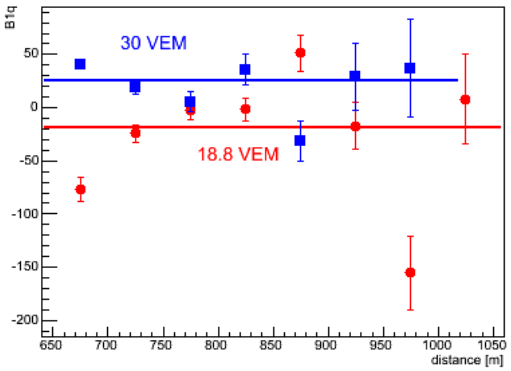
B2p, B2q

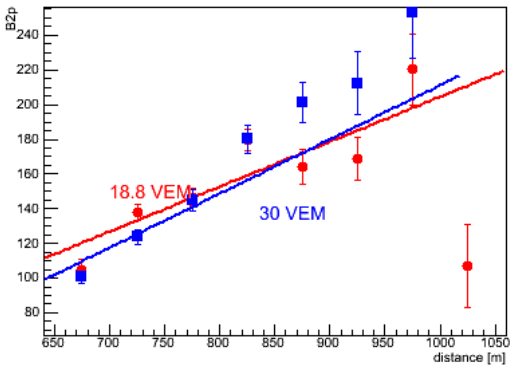
B0p, B0q, B1p, B1q, B2p, B2q vs distance

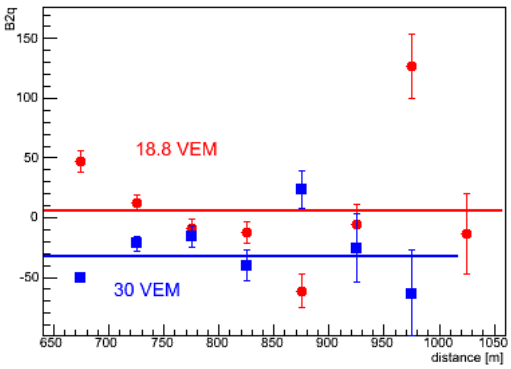












The fitted intercept and slopes are

B0pm, B0pn, B0qm

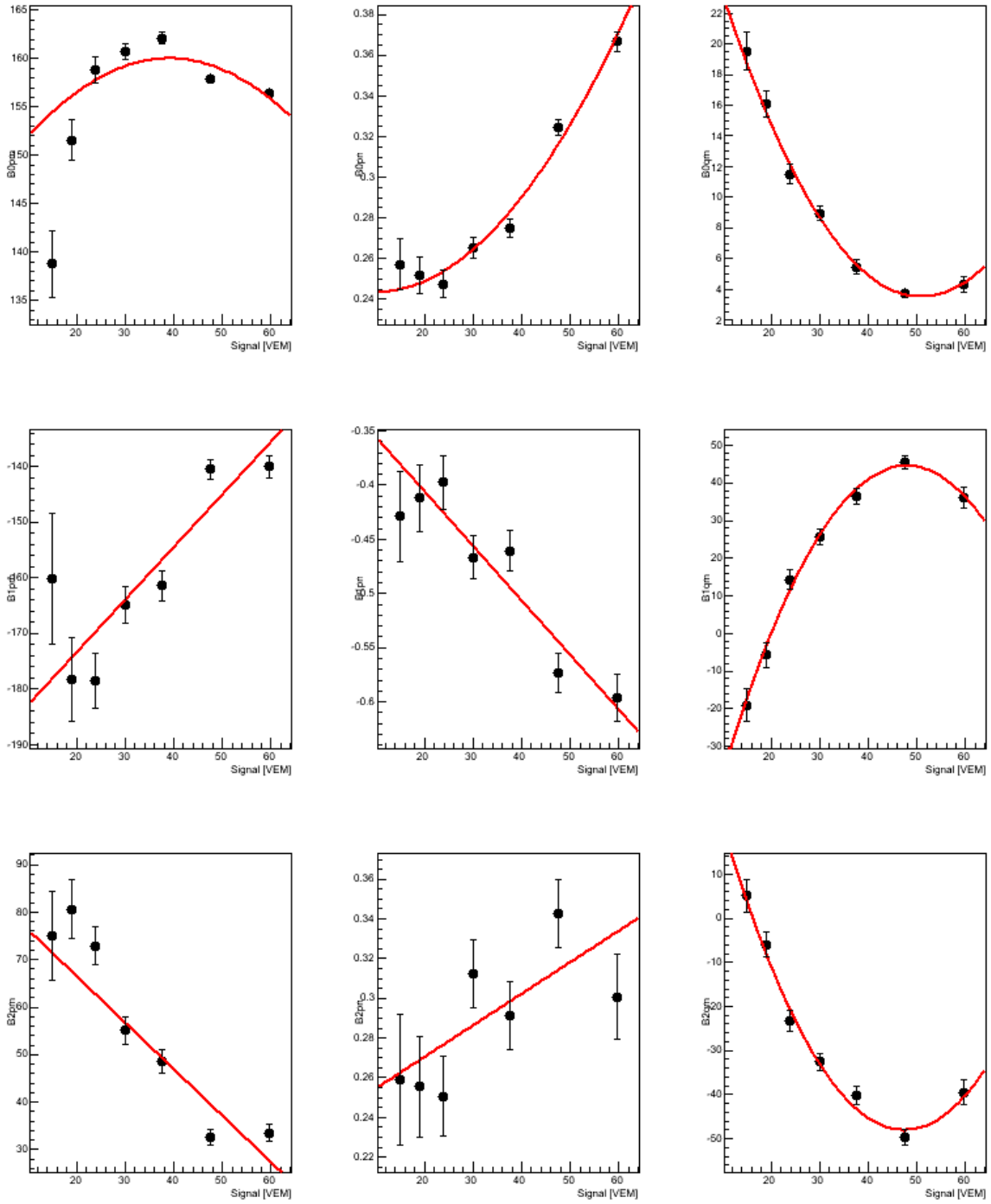
B1pm, B1pn, B1qm

B2pm, B2pn, B2qm

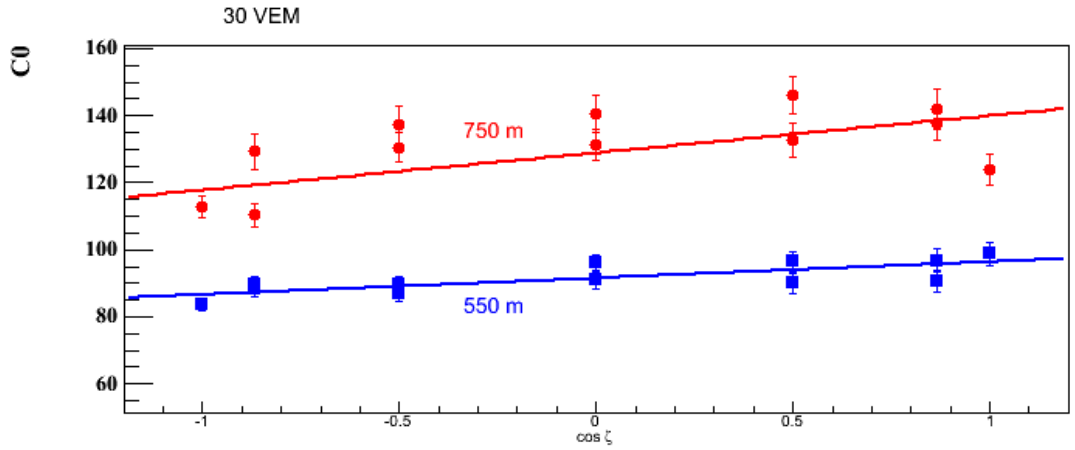
B0pm, B0pn, B0qm

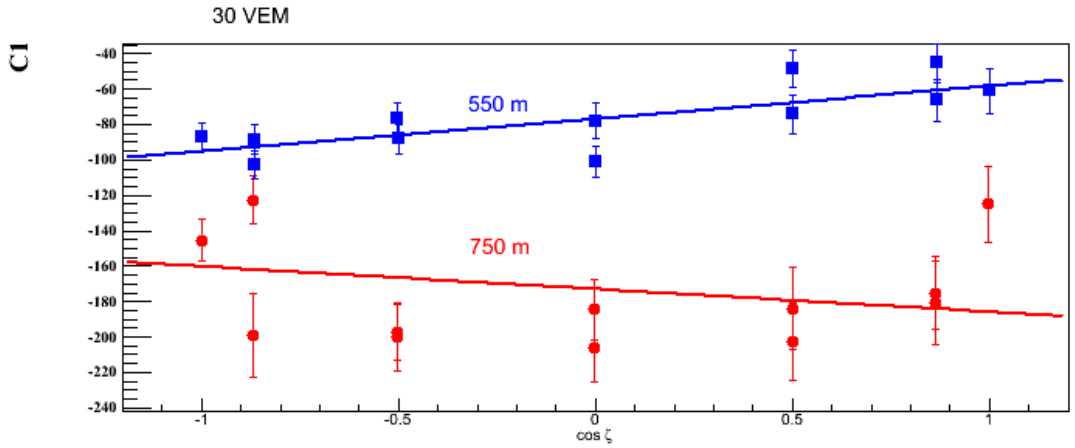
B1pm, B1pn, B1qm vs signal

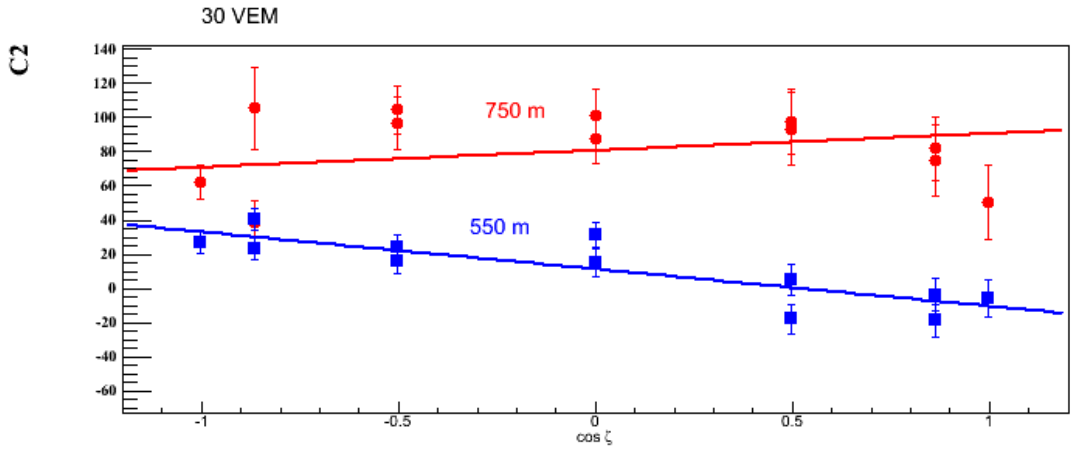
B2pm, B2pn, B2qm



C_0, C_1 and C_2 vs cos ζ







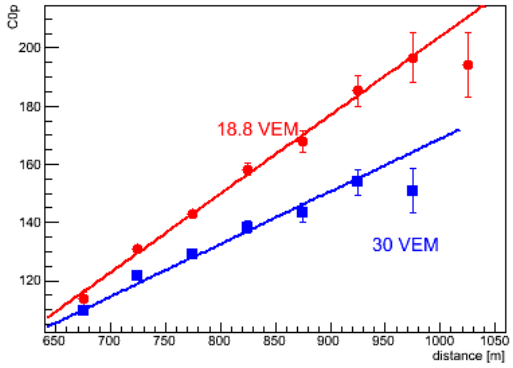
The fitted intercept and slopes are

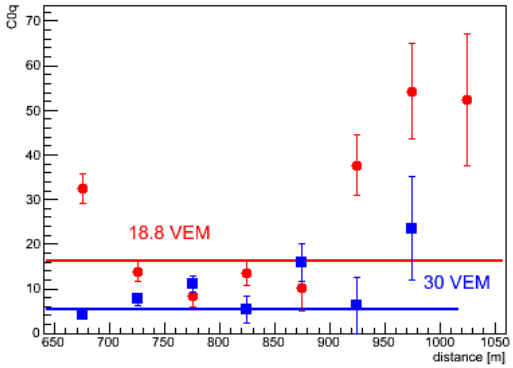
C0p, C0q

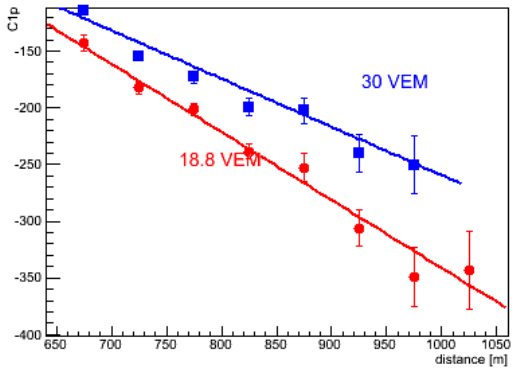
C1p, C1q

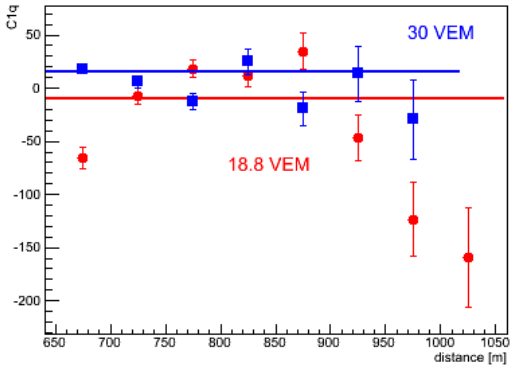
C2p, C2q

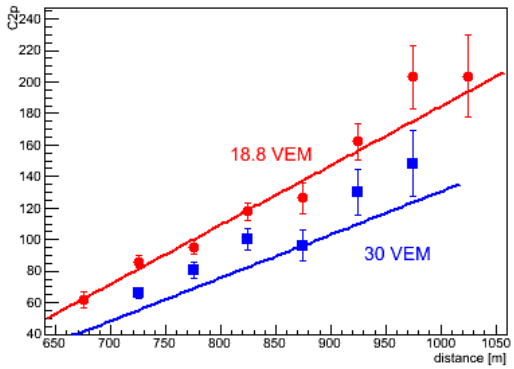
C0p, C0q, C1p, C1q, C2p, C2q vs distance

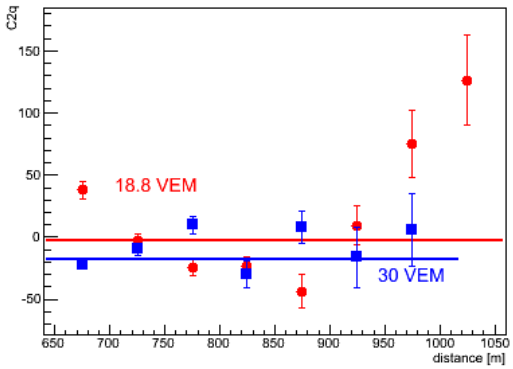












The fitted intercept and slopes are

C0pm, C0pn, C0qm

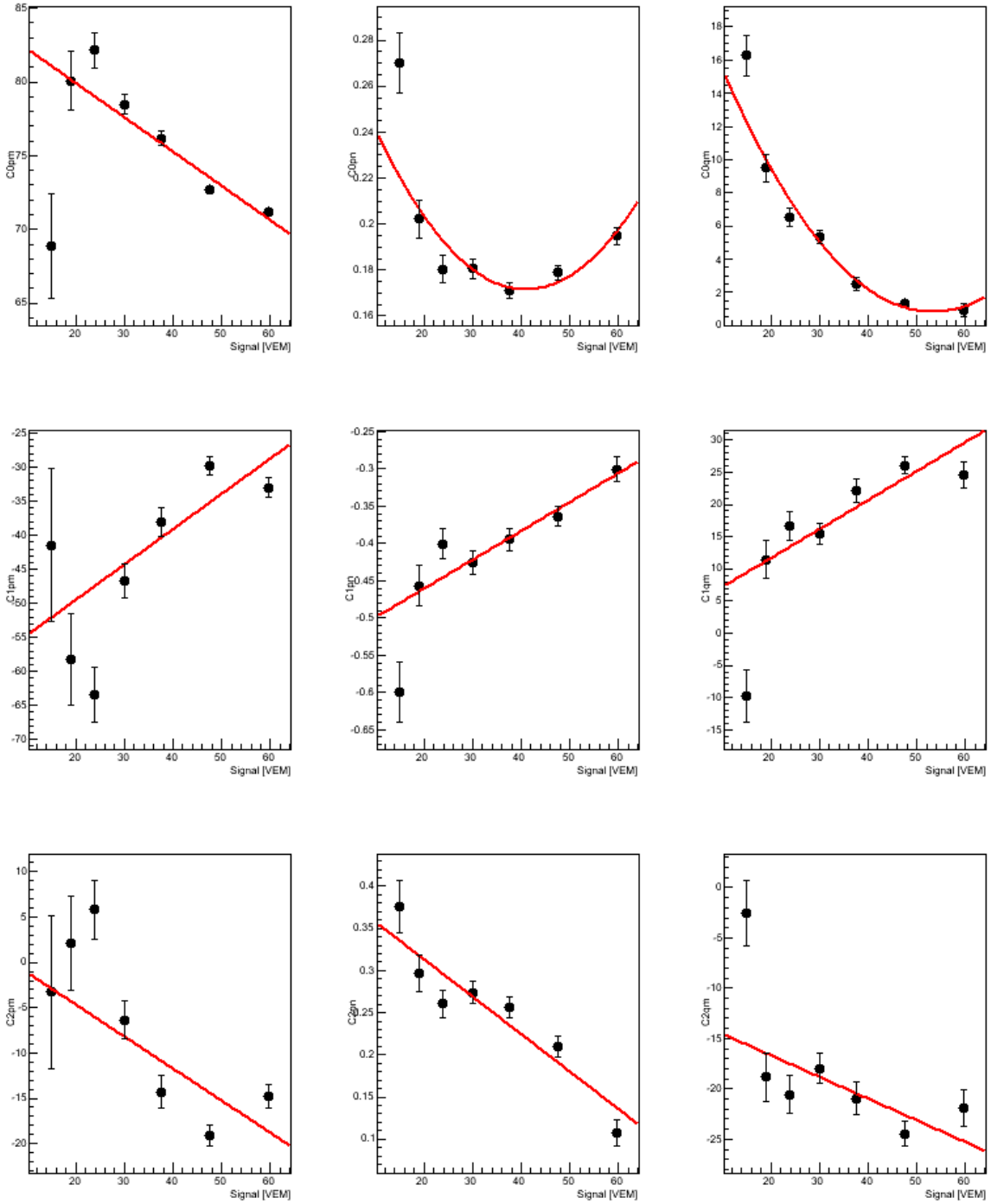
C1pm, C1pn, C1qm

C2pm, C2pn, C2qm

C0pm, C0pn, C0qm

C1pm, C1pn, C1qm vs signal

C2pm, C2pn, C2qm



Appendix B

Use Δ(entity likelihood ratio) for photon-hadron separation in integrated energy bins