

## Search for ultra-high energy photons with the AMIGA muon detector

Zur Erlangung des akademischen Grades eines  
DOKTORS DER NATURWISSENSCHAFTEN  
von der Fakultät für Physik des

*Karlsruher Institut für Technologie (KIT)*  
und der  
*Universidad Nacional de San Martín (UNSAM)*

genehmigte

DISSERTATION

von

Lic. Nicolás Martín González  
aus Buenos Aires

Tag der mündlichen Prüfung: 9 Februar 2018

- Referent: Prof. Dr. A. Etchegoyen,  
Instituto de Tecnologías en Detección y Astropartículas  
CNEA - CONICET - UNSAM
- Koreferent: Prof. Dr. J. Blümer,  
Bereich V  
Karlsruher Institut für Technologie
- Betreuer: Dr. F. A. Sánchez,  
Instituto de Tecnologías en Detección y Astropartículas  
CNEA - CONICET - UNSAM
- Betreuer: Dr. M. A. Roth,  
Institut für Kernphysik  
Karlsruher Institut für Technologie



UNIVERSIDAD  
NACIONAL DE  
SAN MARTÍN



Karlsruher Institut für Technologie

## Search for ultra-high energy photons with the AMIGA muon detector

Tesis presentada para optar por el título de  
DOCTOR EN ASTROFÍSICA  
del Instituto de Tecnología "Prof. Jorge A. Sabato" de la

*Universidad Nacional de San Martín (UNSAM)*  
y del  
*Karlsruher Institut für Technologie (KIT)*

por Lic. Nicolás Martín González

Fecha de la defensa oral: 9 de febrero de 2018

Director: Prof. Dr. A. Etchegoyen,  
Instituto de Tecnologías en Detección y Astropartículas  
CNEA - CONICET - UNSAM

Co-director: Prof. Dr. J. Blümer,  
Bereich V  
Karlsruher Institut für Technologie

Colaborador: Dr. F. A. Sánchez,  
Instituto de Tecnologías en Detección y Astropartículas  
CNEA - CONICET - UNSAM

Colaborador: Dr. M. A. Roth,  
Institut für Kernphysik  
Karlsruher Institut für Technologie

# Abstract

Cosmic rays ( $E > 10^{14}$  eV) have been known for more than 70 years, and yet their origin remains elusive. Similarly, the possibility of a very-high energy photon component of the cosmic radiation is one of the open problems in Astroparticle Physics. The search for high energy photons complements measurements of cosmic rays and neutrinos towards a multi-messenger understanding of the most energetic astrophysical phenomena. In particular, the discovery of photons with energies between  $10^{16.5}$  eV and  $10^{18}$  eV in the cosmic rays flux could be of particular interest not only for the field of Astroparticle Physics, but also for Astrophysics and fundamental Physics, since they are tracers of the highest-energy processes in the Universe.

In the search for ultra-high energy photons, it is crucial to define composition-sensitive parameters capable of adequately rejecting the hadronic cosmic-ray background. The muon content of the extensive air showers produced by primary cosmic rays as they enter the atmosphere is one of the most promising aspects that could lead to the best possible discrimination between photons and hadronic cosmic rays. The AMIGA underground muon detector, as a part of the upcoming AugerPrime upgrade for the Pierre Auger Observatory, offers a unique and straight-forward opportunity to directly measure high-energy muons of extensive air showers, and thus, enhance the sensibility of the Observatory to a primary photon signal.

The Pierre Auger Collaboration has proposed several parameters in order to study a possible ultra-high energy photon component in the hadronic cosmic-ray flux. However, the non-observation of photon candidates resulted in upper limits for energies above  $10^{18}$  eV. On the other hand, the energy window between  $10^{16.5}$  eV and  $10^{18}$  eV has only been explored by the KASCADE-Grande and the EAS-MSU experiments which impose upper limits to the photon flux. Currently, this energy domain is not explored by the Pierre Auger Observatory.

The main objective of this thesis is the extension of the ultra-high energy photon search down to  $\sim 10^{16.5}$  eV. The stringent theoretical and experimental upper limits to the photon flux at these energies make the search of a weak photon signal in the vast hadronic cosmic-ray background a challenging task. Therefore, parameters sensitive to the predominantly electromagnetic signal from photon primaries are of paramount importance. In this framework, we define and describe two new composition-sensitive observables designed for the photon/hadron discrimination quest in order to either detect photon primaries in the  $10^{16}$  eV energy domain or improve the upper limits established by previous experiments.

The observable  $M_b$  combines the muon densities measured by the AMIGA underground muon stations and their distance to the shower axis, similarly to the well-known  $S_b$  parameter used previously in photon searches by the Pierre Auger Collaboration. On the other hand, the observable  $Q$  exploits the difference in the slope of the lateral distribution of particles between photon- and hadron-initiated showers. In the latter case, the showers are expected to develop higher in the atmosphere and thus they arrive to the ground with a flatter distribution of particles.

In this work, we tune both observables to be applied in the surface and muon detector of the Auger Observatory for energies above  $E_{\text{rec}} = 10^{16.4}$  eV and  $\theta < 45^\circ$ . A multiparametric method, baptized  $M_b + Q$ , is extensively studied and its performance is addressed under numerous conditions. The background rejection and the signal efficiency for the compound method is proved to be suitable to impose the best upper limits of all the cosmic-ray experiments with only a few years of data, and particularly improve them by one order of magnitude by the end of Auger planned operation at December 31<sup>st</sup> 2025. Considering the high exposure of the Auger experiment in the direction of the Galactic center, the method described in this thesis represents a valuable tool with an unprecedented discovery potential in the detection of a minuscule photon signal in the cosmic-ray flux. Although a blind estimation of the sensitivity to a photon flux is presented in this research, the application to data is foreseen to be carried out in the near future for a dedicated full-author list paper of the Pierre Auger Collaboration.

**Keywords:** Pierre Auger Observatory, AMIGA, ultra-high energy photons, muon detector.

# Resumen

Los rayos cósmicos ( $E > 10^{14}$  eV) son conocidos desde hace más de 70 años y, sin embargo, su origen sigue siendo un misterio. Del mismo modo, la posibilidad de una componente fotónica de muy alta energía en el flujo de radiación cósmica es uno de los problemas abiertos en la Física de Astropartículas. La búsqueda de fotones de alta energía complementa las mediciones de los rayos cósmicos y neutrinos hacia una comprensión multi-canal de los fenómenos astrofísicos más energéticos. En particular, el descubrimiento de fotones con energías entre  $10^{16.5}$  eV y  $10^{18}$  eV en el flujo de los rayos cósmicos podría ser de particular interés no sólo para el campo de la Física de Astropartículas, sino también para la Astrofísica y la Física fundamental, ya que son rastreadores de los procesos de mayor energía en el Universo.

En la búsqueda de fotones de ultra-alta energía, es crucial definir parámetros sensibles a la composición capaces de rechazar adecuadamente el fondo hadrónico de los rayos cósmicos. El contenido muónico de las lluvias atmosféricas extensas producidas por los rayos cósmicos primarios al entrar en la atmósfera es uno de los aspectos más prometedores que podría conducir a la mejor discriminación posible entre los fotones y los rayos cósmicos hadrónicos. El detector de muones subterráneo AMIGA, como parte de la próxima mejora AugerPrime del Observatorio Pierre Auger, ofrece una oportunidad única y directa para medir muones de alta energía de lluvias atmosféricas extensas, y así, aumentar la sensibilidad del Observatorio a una señal fotónica primaria.

La Colaboración Pierre Auger ha propuesto varios parámetros para estudiar una posible componente fotónica de ultra-alta energía en el flujo hadrónico de rayos cósmicos. Sin embargo, su no-observación resultó en límites superiores para energías por encima de  $10^{18}$  eV. Por otra parte, la ventana energética entre  $10^{16.5}$  eV y  $10^{18}$  eV sólo ha sido explorada por los experimentos de KASCADE-Grande y EAS-MSU, que imponen límites superiores al flujo fotónico. Actualmente, este dominio energético no es explorado por el Observatorio Pierre Auger.

El objetivo principal de esta tesis es la extensión de la búsqueda de fotones de ultra alta energía hasta  $\sim 10^{16.5}$  eV. Los estrictos límites superiores teóricos y experimentales al flujo fotónico en estas energías hacen que la búsqueda de una débil señal fotónica en el vasto fondo de los rayos cósmicos hadrónicos sea una tarea desafiante. Por lo tanto, los parámetros sensibles a la señal predominantemente electromagnética de los fotones primarios son de suma importancia. En este marco, definimos y describimos dos nuevos observables sensibles a la composición diseñados para la discriminación fotón/hadrón con el fin de detectar fotones primarios a energías del orden de  $10^{16}$  eV o mejorar los límites superiores establecidos por experimentos anteriores.

El observable  $M_b$  combina las densidades de muón medidas por las estaciones de muón subterráneas AMIGA y su distancia al eje de la lluvia, de manera similar al conocido parámetro  $S_b$  utilizado anteriormente en las búsquedas fotónicas por la Colaboración Pierre Auger. Por otra parte, el observable  $Q$  explota la diferencia en la pendiente de la distribución lateral de partículas

entre las lluvias iniciadas por fotones y hadrones. En este último caso, se espera que las lluvias se desarrollen a mayores alturas en la atmósfera y lleguen al suelo con una distribución más plana de partículas.

En este trabajo, afinamos ambos observables para ser aplicados en el detector de superficie y detector de muones del Observatorio Auger para energías por encima de  $E_{\text{rec}} = 10^{16.4}$  eV y  $\theta < 45^\circ$ . Un método multiparamétrico, bautizado  $M_b + Q$ , es extensamente estudiado y su ejecución se aborda bajo numerosas condiciones. Se demuestra que el rechazo de fondo y la eficiencia de la señal para el método compuesto son adecuados para imponer los mejores límites superiores de todos los experimentos de rayos cósmicos con tan sólo unos pocos años de datos, y en particular mejorarlos por un orden de magnitud al final actualmente planificado de operación de Auger el 31 de diciembre 2025. Considerando la alta exposición del experimento Auger en dirección al centro galáctico, el método descrito en esta tesis representa una herramienta valiosa con un potencial de descubrimiento sin precedentes en la detección de una señal fotónica minúscula en el flujo de rayos cósmicos. Aunque en esta investigación se presenta una estimación ciega de la sensibilidad a un flujo fotónico, está previsto que la aplicación a los datos se lleve a cabo en un futuro próximo para una publicación oficial de la Colaboración Pierre Auger.

**Palabras clave:** Observatorio Pierre Auger, AMIGA, fotones de ultra-alta energía, detector de muones.

# Zusammenfassung

Kosmische Strahlung ( $E > 10^{14}$  eV) ist seit mehr als 70 Jahren bekannt, jedoch bleiben ihren Ursprung unerklärt. Ebenso ist das Vorhandensein einer hochenergetischen Photonenkomponente der kosmischen Strahlung weiterhin eine offene Frage der Astroteilchenphysik. Die Ergänzung von Messungen kosmischer Strahlung und Neutrinos durch hochenergetische Photonen verfeinert unsere Multi-Messenger-Suche von energetischsten astrophysikalischen Phänomenen. Insbesondere die Entdeckung von Photonen mit Energien zwischen  $10^{16.5}$  eV und  $10^{18}$  eV im Fluss der kosmischen Strahlung könnte nicht nur für die Astroteilchenphysik relevant sein, sondern auch für die Astrophysik und die Grundlagenphysik, da sie Indikatoren der energiereichsten Prozesse im Universum sind.

Bei der Suche nach ultrahochenergetischen Photonen ist es entscheidend, kompositionssensitive Parameter zu definieren, die in der Lage sind, den Untergrund der hadronischen kosmischen Strahlung adäquat zu eliminieren. Der Myonengehalt der Sekundarteilchen ausgedehnter Luftschauer gehört zu den vielversprechendsten Aspekten bei der Unterscheidung zwischen Photonen und hadronischer kosmischer Strahlung. AMIGA, ein sich in 2.5 m Bodentiefe befindlicher Myondetektor, ist Teil des AugerPrime-Upgrades für das Pierre-Auger-Observatorium. Es bietet eine einzigartige Lösung, wobei hochenergetische Myonen von ausgedehnten Luftschauern direkt gemessen werden. Damit wird die Sensivität des Observatoriums zum Nachweis von Photonen als Primärteilchen drastisch verbessert.

Es gab bereits mehrere Ansätze der Pierre-Auger-Kollaboration, um eine mögliche ultrahochenergetische Photonenkomponente in der kosmischen Strahlung zu untersuchen. Die Nichtbeobachtung von Photonenkandidaten führte jedoch zu Obergrenzen für Energien über  $10^{18}$  eV. Andererseits wurde das Energiefenster zwischen  $10^{16.5}$  eV und  $10^{18}$  eV nur von der KASCADE-Grande und den EAS-MSU-Experimenten untersucht, strikte Obergrenzen des Photonenflusses bestimmten. Derzeit wird dieser Energiebereich im Rahmen des Pierre-Auger-Observatoriums nicht erforscht.

Das Hauptziel dieser Arbeit ist die Erweiterung der ultrahochenergetischen Photonensuche bis zu Energien von  $\sim 10^{16.5}$  eV. Die strengen theoretischen und experimentellen Obergrenzen des Photonenflusses bei diesen Energien macht die Suche nach einem schwachen Photonsignal bei hohem hadronischen Strahlungsuntergrund zu einer herausfordernden Aufgabe. Daher sind Parameter, die auf das überwiegend elektromagnetische Signal empfindlich sind, wie im Falle von Photonen, von äußerster Wichtigkeit. In diesem Rahmen werden zwei neue photonsensitive Parameter definiert und beschrieben, die für die Suche nach der Photonen-/Hadronen-Diskriminierung entwickelt wurden, um entweder Photonen im Energiebereich  $10^{16}$  eV zu detektieren oder die Obergrenzen zu verbessern, die durch bisherige Experimente ermittelt wurden.

So wie mit dem  $S_b$  Parameter, welcher zuvor bei der Photonensuche der Pierre-Auger-Kollaboration verwendet wurde, wird bei  $M_b$  die Myondichte der AMIGA-Stationen und deren Abstand zur

Schauerachse kombiniert. Die Observable  $Q$  nutzt den Unterschied in der Steilheit der lateralen Verteilung der Teilchen zwischen photon- und hadroninduzierten Luftschauern. Im letzteren Fall wird erwartet, dass sich die Luftschauer in der Atmosphäre höher entwickeln und deshalb mit einer flacheren Teilchenverteilung auf den Boden gelangen.

In dieser Arbeit werden die Observablen der Oberflächen- und AMIGA-Detektoren so optimiert, dass sie im Energiebereich über  $E_{\text{rec}} = 10^{16.4}$  eV und  $\theta < 45^\circ$  verwendet werden können. Eine multiparametrische Methode,  $M_b + Q$  genannt, wird ausführlich untersucht und ihre Leistungsfähigkeit wird unter zahlreichen Bedingungen untersucht. Die Untergrundunterdrückung und der Photonnachweis für das multivariate Verfahren sind geeignet zur Bestimmung der striktesten Obergrenzen aller Experimente in diesem Energiebereich innerhalb nur weniger Jahre Messzeit. Bis zum Ende der geplanten Laufzeit des Observatoriums wird eine um eine Größenordnung verbesserte Obergrenze erwartet. In Anbetracht der hohen Exposition des Auger-Experiments in Richtung des galaktischen Zentrums, stellt die in dieser Arbeit beschriebene Methode ein wertvolles Werkzeug dar eine winziges Photonensignal im galaktischen Strahlungsfluss zu finden. Obwohl im wesentlichen nur eine Blindstudie zur Empfindlichkeit eines Photonenflusses in dieser Studie vorgestellt wird, ist die Anwendung auf Daten in naher Zukunft im Rahmen einer Auger-Publikation in Arbeit. Erst dann dürfen die bisher unbenutzen Daten untersucht werden.

**Keywords:** Pierre Auger Observatorium, AMIGA, ultrahochenergetische Photonen, Myondetektor.

# Acknowledgments

This research would not have been possible without the help and support from many people throughout my career. It is impossible to order all of them in a hierarchy, so I will try to mention them in a chronological order.

I want to thank my parents, Miguel and Graciela, for the continuous encouragement to improve myself and face the difficulties, since I was a teenager and my personal issues took over me quite a few times. To my family, specially my grandparents, Angel, Argentina, Jose and Josefina, for countless memories that bring a smile to me each time I remember them. To my friends, the present, the past and the gone, for staying by me even when I preferred to study for an exam rather than meeting them during my undergraduate studies.

To my teachers at the Universidad de Buenos Aires, from whom I learnt the solid foundations of fundamental Physics that I needed to carry out my Master and now this Doctoral thesis.

To Brian Wundheiler and Federico Sánchez who offered me the possibility to take my first steps in the world of science back in 2010, for my last experimental laboratory course of my undergraduate studies. To Prof. Alberto Etchegoyen for giving me the chance to develop as a physicist, being the director of my home institution. To my colleagues from ITeDA, Federico, Brian, Juan Manuel Figueira and Diego Ravnani, from whom I learn day after day how to improve my research and how to improve as a researcher. To my colleagues from “the office”, Belén, Botti and Perlin, for making a nice working environment during my time in Buenos Aires.

To Prof. Alberto Etchegoyen and Prof. Johannes Blümer for giving me the opportunity to take part in the Double Doctorate Degree in Astrophysics. To my friends from Karlsruhe, where I stayed one year and lived an unforgettable experience, with a large number of adventures. Specially to Ewa, Sarah, David, Darko, Alvaro, Alex (x3), Dani and Ariel with whom I made quite strong bonds. To Marie-Christine Kauffmann and Frank Schröder, for taking care of any issues I had during my two stays in Karlsruhe.

To Federico Sánchez and Markus Roth, for being excellent researchers and examples to follow, giving valuable advice regarding the analyses themselves, but most importantly, for being open with me to talk about other topics and personal issues as well. I really appreciate that.

To Mercedes, my life partner, the most important person to me, for staying by my side in the saddest and happiest moments of my life, for supporting me when I decided to spend one year abroad and then coming once during each stay despite the enormous effort that a trip like that means, for giving me strength when I needed it and for forgiving me each time I spent long nights and entire weekends working.

*A mi país, que ya verá tiempos mejores.*

# Contents

<b>1</b>	<b>Cosmic Rays</b>	<b>1</b>
1.1	The History of the Cosmic Rays . . . . .	1
1.2	Acceleration of Cosmic Rays . . . . .	4
1.2.1	Bottom-up scenarios . . . . .	5
1.2.2	Top-down scenarios . . . . .	11
1.3	Propagation of Cosmic Rays . . . . .	13
1.3.1	Magnetic deflection . . . . .	13
1.3.2	GZK effect . . . . .	14
1.3.3	Energy losses for nuclei . . . . .	16
1.4	Energy spectrum of Cosmic Rays . . . . .	17
1.4.1	Features of the energy spectrum . . . . .	18
1.5	Ultra-high energy photons . . . . .	20
1.5.1	Sources and generation mechanisms . . . . .	20
1.5.2	Propagation . . . . .	25
1.5.3	Upper limits to the photon flux . . . . .	27
1.5.4	Motivations for the search at ultra-high energies . . . . .	28
<b>2</b>	<b>Physics of extensive air showers</b>	<b>31</b>
2.1	Components of the EAS . . . . .	31
2.2	Topology of an EAS . . . . .	33
2.3	The Heitler model . . . . .	35
2.3.1	Extension to hadronic showers . . . . .	37
2.3.2	Superposition model . . . . .	38
2.4	Properties of EAS produced by primary photons . . . . .	39
2.4.1	LPM and Preshower effects . . . . .	40
<b>3</b>	<b>The Pierre Auger Observatory</b>	<b>43</b>
3.1	Site location and base design . . . . .	43
3.2	The Fluorescence Detector . . . . .	45
3.2.1	Telescope hardware . . . . .	45
3.2.2	Atmospheric monitoring . . . . .	47
3.2.3	FD trigger . . . . .	47
3.2.4	FD reconstruction . . . . .	48
3.3	The Surface Detector . . . . .	50
3.3.1	Calibration of the WCD . . . . .	51
3.3.2	WCD local triggers . . . . .	53
3.3.3	SD array triggers . . . . .	54
3.3.4	SD reconstruction . . . . .	57
3.3.5	The SD-750 array . . . . .	61
3.3.6	The SD-433 array . . . . .	61
3.4	The AMIGA Muon Detector . . . . .	63
3.4.1	Layout of the MD Unitary Cell . . . . .	63
3.4.2	Scintillation bar - fiber tandem of a MD module . . . . .	65
3.4.3	Optic system of the MD . . . . .	65
3.4.4	Electronics and data acquisition of the MD . . . . .	67
3.4.5	Muon counting strategy . . . . .	68

3.4.6	Muon LDF reconstruction . . . . .	70
3.5	Upgrades of the Pierre Auger Observatory . . . . .	70
3.5.1	Current enhancements . . . . .	70
3.5.2	AugerPrime . . . . .	71
<b>4</b>	<b>Simulation of the detector responses</b>	<b>75</b>
4.1	AugerOffline framework . . . . .	75
4.2	Event simulation within AugerOffline . . . . .	77
4.2.1	Production of EAS . . . . .	77
4.2.2	Monte Carlo procedure . . . . .	78
4.2.3	Tuning the MD simulation . . . . .	79
4.3	WCD and array triggers . . . . .	82
4.3.1	Array trigger efficiency . . . . .	83
4.4	Features of the SD reconstructed events . . . . .	89
4.4.1	Geometry reconstruction . . . . .	90
4.4.2	LDF reconstruction . . . . .	91
4.4.3	Energy reconstruction . . . . .	93
4.5	Features of the MD reconstructed events . . . . .	97
<b>5</b>	<b>Photon-hadron discrimination with the observable <math>M_b</math></b>	<b>101</b>
5.1	The observable $M_b$ . . . . .	102
5.2	Modelling the $\rho_\mu^{\text{ref}}(E, \theta)$ for $M_b$ . . . . .	103
5.3	Separation power of $M_b$ . . . . .	106
5.4	Saturation effects on $M_b$ . . . . .	111
5.5	Systematic effect of the reconstructed energy . . . . .	114
5.5.1	Background rejection and signal efficiency . . . . .	114
5.6	Application of $M_b$ to non-vertical events . . . . .	117
5.6.1	Background rejection and signal efficiency . . . . .	121
<b>6</b>	<b>Multivariate analysis combining SD and MD</b>	<b>125</b>
6.1	Extracting information from the LDF . . . . .	127
6.1.1	Parameterization of the photon LDF slope . . . . .	127
6.1.2	Testing the photon LDF model . . . . .	130
6.1.3	Performance of the observable $Q$ . . . . .	132
6.2	Linear discriminant analysis . . . . .	133
6.2.1	Performance of the Fisher discriminant . . . . .	135
6.2.2	Background rejection and signal efficiency . . . . .	136
6.3	Parameterization of the Fisher coefficients . . . . .	138
6.4	Non-linear multivariate analysis . . . . .	142
6.5	Discussion on possible systematic effects . . . . .	143
<b>7</b>	<b>Towards the search for ultra-high energy photons in data</b>	<b>147</b>
7.1	Data selection . . . . .	148
7.1.1	General features of the data . . . . .	148
7.1.2	Impact of the quality cuts . . . . .	151
7.2	Exposure calculation . . . . .	152
7.2.1	Uptime of the detectors . . . . .	155
7.2.2	Correction by quality cuts . . . . .	157
7.2.3	Correction by trigger efficiency . . . . .	158
7.2.4	Exposure estimation . . . . .	159
7.3	Sensitivity and upper limits to the photon flux . . . . .	160
7.3.1	Comparison with other experiments . . . . .	163
7.4	Outlook and future prospects for the photon search . . . . .	164
<b>8</b>	<b>Conclusions</b>	<b>167</b>
	<b>Bibliography</b>	<b>169</b>

<b>A Sequences of <u>Offline</u> modules</b>	<b>185</b>
<b>B Model for the <math>M_b</math> distribution</b>	<b>189</b>
<b>C A thought experiment: adding <math>X_{\max}</math> to the MVA</b>	<b>191</b>



# Nomenclature

AERA	Auger Engineering Radio Array
AGNs	Active Galactic Nuclei
AMIGA	Auger Muons and Infill for the Ground Array
Auger	Pierre Auger Observatory
BDT	Boosted Decision Trees
CDAS	Central Data Acquisition System
CIC	Constant Intensity Cut
CL	Confidence Level
CLF	Central Laser Facilities
CMB	Cosmic Microwave Background
CORSIKA	Cosmic Ray Simulations for Cascade
EA	Engineering Array
EAS	Extensive Air Showers
EM	Electromagnetic
FADC	Flash Analog to Digital Converter
FD	Fluorescence detector
FLT	First Level Trigger
FoV	Field of View
FPGA	Field Programmable Gate Array
FR II	Fanaroff-Riley class II
GDAS	Global Data Assimilation System
GPS	Global Positioning System
GRBs	Gamma-ray Bursts
GZK	Greisen-Zatsepin-Kuzmin
HEAT	High Elevation Auger Telescopes
HESS	High Energy Stereoscopic System
HiRes	High Resolution Fly's Eye
ICS	Inverse Compton Scattering

IRB	Infrared Radio Background
KT	Kathy Turner
LAT	Large Area Telescope
LDF	Lateral Distribution Function
LPM	Landau-Pomeranchuk-Migdal
MAGIC	Major Atmospheric Gamma Imaging Cherenkov
MC	Monte Carlo
MD	Muon detector
MLDF	Muonic lateral distribution function
MVA	Multivariate Analysis
pc	Parsec
PE	Photoelectrons
PMT	Photomultiplier tube
PP	Pair Production
PVC	Polyvinyl Chloride
SD	Surface detector
SDEU	Surface Detector Electronics Upgrade
SDP	Shower Detector Plane
SHDM	Super-Heavy Dark Matter
SiPM	Silicon Photomultiplier
SLT	Second Level Trigger
SNRs	Supernova Remnants
SSD	Scintillator Surface Detector
Thr	Threshold
TLT	Third Level Trigger
ToT	Time over Threshold
TPP	Triplet Pair Production
UC	Unitary Cell
UMD	Underground Muon Detector
URB	Universal Radio Background
UTM	Universal Mercator System
VEM	Vertical Equivalent Muon
WCD	Water-Cherenkov detector
WLS	Wavelength shifter
XLF	Extreme Laser Facilities

# Chapter 1

## Cosmic Rays

A mind stretched by a new idea can never go back to its original dimensions.

---

Oliver W. Holmes, Jr.

### 1.1 The History of the Cosmic Rays

Cosmic rays are relativistic particles originated in outer space that impinge the planet permanently. This source of inexhaustible ionizing radiation consists mainly of protons and ionized atomic nuclei (mostly hydrogen), with a small fraction of heavier nuclei and photons. It is a phenomenon that extends over a wide range of energies: from  $10^9$  eV to  $10^{20}$  eV. Although the nomenclature *Cosmic Rays* can be deceiving, since these particles are mainly ionized nuclei, is the one that the community has adopted to talk about this phenomenon.

Following the discovery of radioactivity by Henry Becquerel in 1896, the ionization of atmospheric molecules was believed to be caused only by terrestrial radioactive elements. Subsequent studies by Marie and Pierre Curie indicated that radioactive material emits charged particles that would discharge an electroscope<sup>1</sup>. Its discharge rate was then used as a measure of the air radioactivity. It was observed that electroscopes spontaneously discharged even in isolated environment. Therefore, the existence of background radiation was postulated, whose terrestrial origin was taken for granted.

At the beginning of the 20<sup>th</sup> century, numerous experiments were carried out to discover the origin and nature of this radiation. In 1909, Theodor Wulf developed the first electrometer<sup>2</sup>, a device that was used to measure the production of ions inside a hermetic chamber. His experiments in the Eiffel tower resulted in higher levels of radiation at the top of the tower than at its base. However, the publication of his results was not widely accepted. In 1911, Domenico Pacini compared the discharge velocity of an electroscope on a lake and 3 m below its surface. His results indicated a sharp decrease in the discharge rate of the electroscope when it was placed under water, so that Pacini concluded that much of the ionizing radiation came from the atmosphere [1].

---

<sup>1</sup>The electroscope was the first instrument designed to qualitatively detect electrical charges.

<sup>2</sup> An electrometer is an electroscope that can be used to quantify the electric charge of an object.



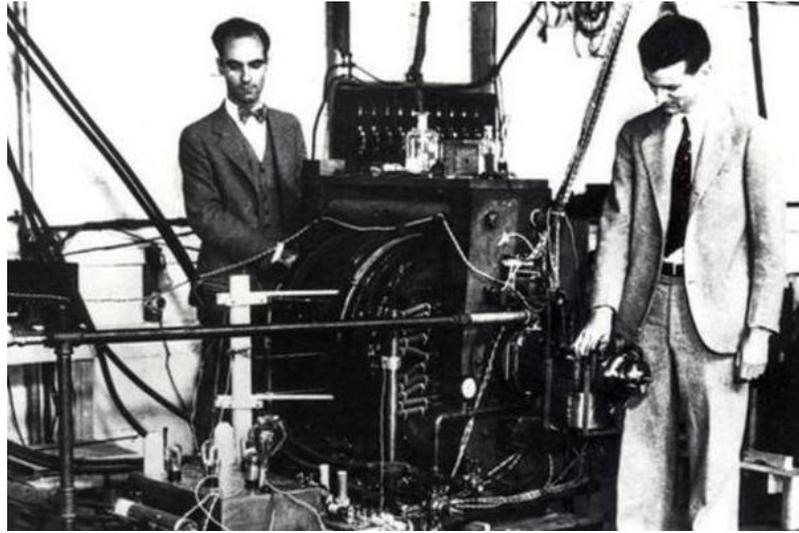
**Figure 1.1:** Between 1911 and 1912, Victor Franz Hess made ten balloon trips carrying electroscopes to determine the source of the atmospheric radiation. This photograph shows him in the ascent to nearly 4900 m realized in August of 1912. Credits to 'Victor Franz Hess Society', Schloss Pöllau, Austria.

Parallel to Pacini, Victor Hess performed experiments on hot air balloons to explain why radiation was detected in ionization chambers even though they were heavily shielded. The Fig. 1.1 shows Hess in one of these ascents. He found that the radiation level at 5000 m from the surface was almost 4 times greater than that detected on Earth. Hess ruled out the possibility that this effect was due to solar radiation since he made this ascension during a total solar eclipse. These results together with the reports of Wulf and Pacini led to the conclusion that the radiation came from outer space. In this way, Hess is credited with the discovery of what would later be known as *Cosmic Rays*, for which he received the Nobel Prize in Physics in 1936.

In the next 20 years, experiments were carried out in order to unravel the composition of this ionizing radiation. In 1925, Robert Millikan confirmed the extraterrestrial nature and baptized this radiation as *cosmic rays*. Millikan believed that cosmic rays were ultra-energetic photons with some secondary electrons produced during their propagation [2]. Consequently, cosmic radiation had to be mostly neutral and could not interact with the geomagnetic field. A worldwide study by Arthur Compton in 1932 [3] showed that cosmic radiation depended on geomagnetic latitude<sup>3</sup>, and thus it had to be predominantly composed by charged particles<sup>4</sup>. During the 1930s, a large variety of experiments supported this conclusion and further indicated that the particles observed on the ground were composed of an electromagnetic component and a component of "penetrating particles".

<sup>3</sup>The geomagnetic latitude is a parameter analogous to geographical latitude, except that it is measured with respect to the magnetic poles.

<sup>4</sup>Nowadays, it is known that this effect is efficient for particles with energies up to  $10^9$  eV to  $10^{11}$  eV (in vertical incidence). In particular, the deflection angle is inversely proportional to the magnetic rigidity  $R$  defined as  $R = pc/Ze$ , where  $pc$  is the relativistic kinetic energy of the particle and  $Ze$  its electric charge.



**Figure 1.2:** Carl Anderson (left) and Seth Neddermeyer (right) with the magnet cloud chamber Anderson used to discover the positron in 1932, and together, discover the muon in 1937. Photo from Caltech Archives, Office of Public Relations.

While the secondary particles of the cosmic rays were studied, new subatomic particles were discovered. In 1932, Carl Anderson discovered the first antimatter particle in the form of antielectron using a cloud chamber that he surrounded with an electromagnet which caused the ionized particles to move on circular trajectories [4]. He received the Nobel Prize in Physics in 1936 for this discovery. In 1937, together with Seth Neddermeyer and using the same cloud chamber as shown in Fig. 1.2, they discovered the muon [5]. These findings encouraged the use of cosmic rays to develop Particle Physics until the advent of particle accelerators in the 1950s. In 1937, Homi J. Bhabha and Walter Heitler described how primary cosmic rays interact with the upper atmosphere to produce electromagnetic particles. The Heitler model predicts many of the main characteristics of these type of processes and it is described in Sec. 2.3.

In 1938, Pierre Auger and Robert Maze showed that secondary particles arrived simultaneously at surface detectors separated by distances of up to 20 m [6]. This temporal correlation indicated that the particles came from the same physical event which was suggested to be a high-energy cosmic ray that interacted with the molecules of the atmosphere. The process would be as follows: after the initial interaction, a cascade of electrons, photons and muons would develop and some of them would eventually reach the ground to simultaneously fire the surface detectors. Auger and Maze named this phenomenon of particle multiplication from a primary agent as *Extensive Air Showers* (EAS). This phenomenon is described in detail in Chapter 2. Further experiments at the Jungfrauoch region in the Swiss Alps showed that coincidences continued to be observed even at distances of 300 m above the surface. Pierre Auger and his group estimated the primary energy to be around  $10^{15}$  eV. This estimate was based on the number of particles in the shower [7]. It is worth quoting the final remarks of Auger at the 1939 Symposium held in Chicago:

*One of the consequences of the extension of the energy spectrum of cosmic rays up to  $10^{15}$  eV is that it is actually impossible to imagine a single process able to give to a particle such an energy. It seems much more likely that the charged particles which constitute the primary cosmic radiation acquire their energy along electric fields of a very great extension [6].*

Auger and his colleagues had discovered particles with energies of  $10^{15}$  eV at a time when the highest observed energies were of a few MeV in radioactive phenomena. In 1949, Enrico Fermi proposed a mechanism of acceleration based on the interaction with the interstellar magnetic fields to explain how these particles acquired such energy [8]. According to this mechanism, particles would gain energy by colliding with a moving magnetized plasma cloud. The original Fermi theory was modified in 1954 [9], proposing a more efficient acceleration process in which particles would be accelerated in successive encounters with large astrophysical shock waves, reaching energies up to  $10^{15}$  eV. In Sec. 1.2 both models are described.

In 1948, Melvin B. Gottlieb and James Van Allen observed through photographic plates carried in aerostatic balloons to the upper atmosphere that the primary cosmic rays are mostly protons and helium nuclei with a small fraction of heavy nuclei. In the 1950s, the use of group of detectors for the study of EAS was popularized and larger and more complex arrays were built, replacing the Geiger-Müller tubes with scintillation detectors. The first measurements of energy and arrival direction of cosmic rays above  $10^{15}$  eV were carried out in 1954 by the Rossi Cosmic Ray Group of the University of Massachusetts using an array of 11 scintillation detectors arranged in a circle of 230 m radius. Between 1954 and 1957, the cosmic-ray energy spectrum was extended to  $10^{18}$  eV with data provided by the Agassiz Station at Harvard [10]. In 1962, Volcano Ranch in Albuquerque, New Mexico, would be the first of these giant arrangements to detect an ultra-energetic event of  $1.4 \times 10^{20}$  eV [11]. From these and similar works, it was concluded that the energy spectrum of the primary cosmic rays extended beyond  $10^{20}$  eV. The energy spectrum and its features are described in Sec. 1.4.

After the discovery of the cosmic microwave background (CMB) by Arno A. Penzias and Robert W. Wilson in 1964 [12], Kenneth Greisen, Georgiy Zatsepin and Vadim Kuzmin demonstrated that at energies above  $5 \times 10^{19}$  eV protons traveling through outer space begin to interact with the CMB photons, producing secondary pions [13]. This interaction degrades the proton energy, preventing particles with energies greater than  $\sim 5 \times 10^{19}$  eV to be observed at distances greater than  $\sim 100$  Mpc<sup>5</sup>. The suppression of cosmic rays with higher energies is known as the *GZK cut-off* and it is discussed in Sec. 1.3.2.

Since that Volcano Ranch event, more events with energies higher than the GZK cut-off were detected by various experiments, including the "Oh My God" particle ( $3 \times 10^{20}$  eV) detected by Fly's Eye<sup>6</sup> in 1991. The inconclusive experimental proof about the existence of the GZK cut-off was a main reason for the design of a new Observatory that would be able to gather enough statistics at the highest energies where the cosmic-ray flux is  $\sim 1 \text{ km}^{-2} \text{ century}^{-1}$ . This piece of the cosmic-ray puzzle led, among other factors, to the construction of the Pierre Auger Observatory.

Nowadays, more than a century after the discovery of cosmic rays, the most energetic end of this phenomenon remains a subject of study and debate because the astrophysical mechanisms that accelerate particles at energies  $\sim 10^{20}$  eV have not been fully explained, nor the details of the observed spectrum that would be generated in its propagation, nor the mass composition and the arrival direction of this form of radiation have been established.

## 1.2 Acceleration of Cosmic Rays

Ever since the reports of the detection of two cosmic-ray showers of energies well above  $10^{20}$  eV

<sup>5</sup>One parsec (pc) is equivalent to 3.2616 light years or  $3.0857 \times 10^{16}$  m.

<sup>6</sup>Fly's Eye was a fluorescence detector that operated from 1981 to 1993 in Utah. It was later replaced by the *High Resolution Fly's Eye* (HiRes) detector.

[14], the origin and the nature of the highest energies known in nature have been the subject of strong interest and intense discussion. The possible explanations range from conventional shock acceleration to particle physics beyond the Standard Model. The motivation for some of the more exotic scenarios may have diminished by newest data. On the contrary, conventional shock have been favored by the large-scale anisotropy in the cosmic-ray arrival directions above  $8 \times 10^{18}$  eV reported by the Pierre Auger Collaboration [15]. Although such acceleration mechanisms in astrophysical objects pushes the present theoretical ideas to their extreme, two classes of processes can be distinguished: the so called *bottom-up* and *top-down* scenarios. The origin of the particles remains unknown.

### 1.2.1 Bottom-up scenarios

The conventional or *bottom-up* astrophysical models focus the explanation on the existence of active objects to accelerate particles and thus, generating the cosmic rays. Within these models, a distinction can be made depending on whether it is a direct acceleration or a diffusive process. In the first case, the process takes place in intense electric fields, such as those found in compact objects like neutron stars or black hole accretion discs. In the second case, nuclei are accelerated in magnetized plasma, such as shock wave systems in supernova remnants (SNRs).

**Direct acceleration** The idea behind the direct or "non-stochastic" acceleration mechanism can be traced back to the early 1930s when Willam F. G. Swann pointed out that particle acceleration may take place in the increasing magnetic field of a sunspot [16]. More recently, similar "one-shot" mechanisms have been extended to astrophysical objects such as the rapidly-rotating highly-magnetized neutron stars (pulsars), active Galactic nuclei (AGNs) or radio-active galaxies.

Pulsars are formed when the core of a massive star collapses to a neutron star during a supernova. Young pulsars, such as the Crab or the Vela [17], with surface magnetic field in the range  $10^{12} - 10^{14}$  G could accelerate iron nuclei through relativistic hydrodynamical winds up to  $\sim 10^{20}$  eV [18]. The electric fields are located in regions of high energy densities, and thus the cross section for synchrotron radiation is high. Therefore, the accelerated particles would interact and subsequent formation of cascades of electron-positron pairs would reduce the efficiency of such acceleration processes. The maximum attainable energy for charged particles in this environment is in the order of  $10^{15}$  eV [19]. It has also been proposed that supermassive black holes (with mass  $\sim 10^8 - 10^{10} M_{\odot}$ , where  $M_{\odot} = 1.98 \times 10^{30}$  kg is the mass of the Sun) with magnetic fields  $B \sim 10^4 - 10^6$  G in the AGNs might also be accelerators up to  $\sim 10^{20}$  eV [20].

The direct acceleration mechanisms are, however, not widely favored these days as the cosmic-ray acceleration mechanism. A major disadvantage of the mechanism is that it is difficult to obtain the characteristic power-law spectrum of the cosmic rays in any natural way [21].

**Stochastic acceleration** The diffusive or "stochastic" processes are based on the Fermi acceleration theory. In 1949, Enrico Fermi proposed a first model by which particles are accelerated by the interaction with moving magnetized plasma inside the Galaxy. The particles would acquire energy due to collisionless scatterings with randomly moving inhomogeneities of the turbulent magnetic field [8], a process which resembles the acceleration of a ball bouncing elastically between two rigid walls that approach each other. In this case, although in each individual encounter the particle may either gain or lose energy, there is on average a net gain of energy  $\Delta E$  after many

encounters, which depends on the velocity of the cloud  $v$  through:

$$\frac{\Delta E}{E} = \frac{4}{3}\beta^2 \quad (1.1)$$

where  $\beta = v/c \ll 1$ . Then, the time scale of this process can be calculated as:

$$\tau = \left( \frac{1}{E} \frac{dE}{dt} \right)^{-1} = \left( \frac{\Delta E}{E} \frac{1}{\tau_{\text{col}}} \right)^{-1} = \frac{4}{3} \frac{\tau_{\text{col}}}{\beta^2} \quad (1.2)$$

being  $\tau_{\text{col}} = c/\lambda_{\text{col}}$  the mean time between collisions (see the full calculation in [22]). Due to the quadratic dependence on  $\beta$ , this model is known as *second-order* Fermi mechanism. Even though the average magnetic field may vanish, there can still be a net transfer of the macroscopic kinetic energy from the moving cloud to the particle. Considering typical cloud velocities  $v \simeq 10 \text{ km seg}^{-1}$  and  $\lambda_{\text{col}} \simeq 1 \text{ pc}$ , the time scale results to be  $\tau \sim 2 \times 10^9 \text{ yr}$ , which makes the process inefficient. In addition, although the resulting spectrum of particles happens to be a power-law in energy, the power-law index depends on the cloud velocity, and so the superimposed spectrum due to many different sources with widely different cloud velocities would not in general have a power-law shape. It should be noted that for a given acceleration site, there is a maximum achievable energy,  $E_{\text{max}}$ , which is defined by the size of the shock and by the time scale of the acceleration process (see Sec. 1.2).

A more efficient version of Fermi acceleration was proposed in the late 1970s [23]. In this model (now called *first-order*) particles can be accelerated by a shock wave front propagating through interstellar space [9]. In this case, the average fractional energy gain of a particle per encounter (defined as a cycle of one crossing and then a re-crossing of the shock after the particle is turned back by the magnetic field) is proportional to the relative velocity between the shock front and the isotropic cosmic-ray frame:

$$\frac{\Delta E}{E} = \frac{4}{3} \left( \frac{\zeta - 1}{\zeta} \right) \beta \quad (1.3)$$

where  $\beta$  corresponds to the shock wave speed and  $\zeta$  is the compression factor that relates the velocity of the collision front to the velocity of the plasma in the post-shock region. The mechanism is more efficient for greater confinements of the particles, which can be easily achieved if the magnetic fields are increasingly intense. In particular, this situation is favored when the direction of the magnetic field is parallel to the front of the shock wave. The complementary situation, where the magnetic field is perpendicular to the front of the collision, allows the particles to diffuse away from the region thus reducing confinement and decreasing the efficiency of the acceleration mechanism.

Assuming a typical value for the shock wave speed of  $\beta = 0.1$  and particles with energies  $\sim 10^{18} \text{ eV}$ , the acceleration time is  $\sim 10^6 \text{ yr}$ . Consequently, this process of acceleration by shock waves is several orders of magnitude faster than the second-order model<sup>7</sup>. An important feature of the first-order Fermi acceleration is that particles emerge out of the acceleration site with a characteristic power-law spectrum with an index that depends only on the shock compression ratio  $\zeta$ , and not on the shock velocity  $v$ .

In any Fermi acceleration mechanism, the net energy gain of a particle is proportional to its energy. Therefore, when considering  $n$  interactions with the accelerator object, the acquired

---

<sup>7</sup>The basic ideas of the first-order Fermi acceleration have received impressive confirmation from in-situ observations in the Solar System, in particular, from observations of high energy particles accelerated at the Earth's bow shock generated by collision of the solar wind with the Earth's magnetosphere [24].

energy is:

$$E = E_0(1 + \alpha)^n \quad (1.4)$$

where  $\alpha$  verifies that  $\Delta E = \alpha E$ . The particles must escape from this acceleration region to begin their propagation. Assuming that a particle has an escape probability  $\epsilon$ , the number of particles  $N$  escaping after  $n$  interactions is:

$$N = N_0(1 - \epsilon)^n \quad (1.5)$$

Combining Eqs. 1.4 and 1.5, it can be seen that the flux of particles follows a power-law with the energy (see Eq. 1.6). However, it should be noted that a power-law spectrum does not necessarily point to Fermi acceleration, but results whenever a fractional gain in energy of a few particles is accompanied by a significantly larger fractional loss in the number of remaining particles [25].

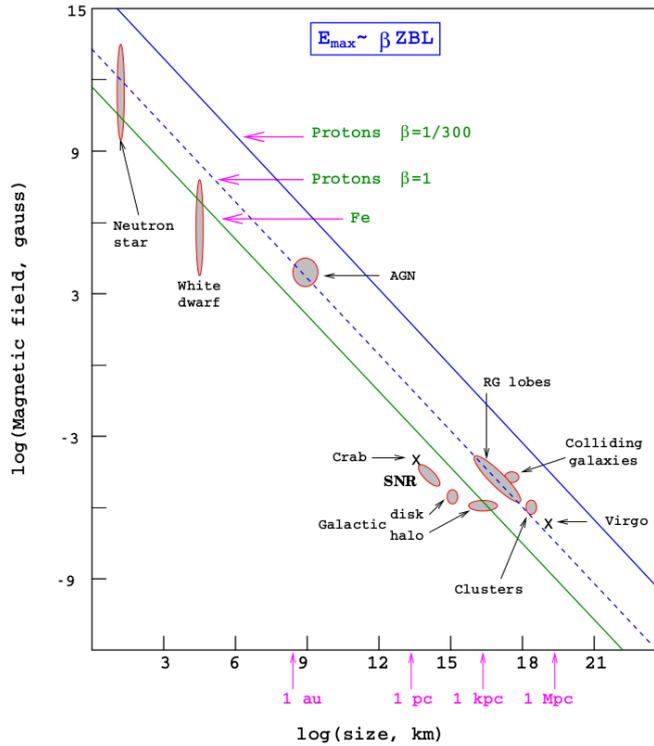
$$\frac{N}{N_0} = \left( \frac{E}{E_0} \right)^{\ln(1-\epsilon)/\ln(1+\alpha)} \quad (1.6)$$

**Possible regions of acceleration** The cosmic rays must be accelerated to ultra-high energies by electromagnetic processes at extragalactic sites, both because there are no known sites in our Galaxy which can accelerate and magnetically contain them and also because most of the observed cosmic rays arrive from directions outside of the Galactic plane. So far, no astrophysical object has been unambiguously identified as a source for ultra-high energy cosmic rays. Despite the exact acceleration mechanism, there is an argument proposed by Anthony M. Hillas [26] that limits the source candidates. Hillas noted that any object that accelerates charged particles must have a magnetic field intense enough to confine them for a minimum amount of time, similar to the situation in man-made accelerators such as at CERN. Sometimes the acceleration region itself only exists for a limited period of time; for example, supernovae shock waves dissipate after about  $10^4$  yr [27]. Otherwise, if the plasma disturbances persist for much longer periods, the maximum energy may be limited by an increased likelihood of escape from the region. If the gyroradius of the particle is required to be less than the characteristic dimension of the object, it is possible to relate the maximum energy  $E_{\max}$  that the particle could gain with the characteristics of the accelerator object. For the case of shock waves and relativistic particles, it can be obtained that [28]:

$$E_{\max} = \beta c Z e B L \approx \beta Z \left( \frac{B}{\mu\text{G}} \right) \times \left( \frac{L}{\text{pc}} \right) \times 10^{15} \text{ eV} \quad (1.7)$$

where  $\beta = v/c$  is the shock wave speed or the efficiency of the acceleration mechanism,  $eZ$  is the electric charge of the particle,  $B$  is the magnetic field inside the acceleration volume and  $L$  is the characteristic size of the object. The dimensional argument expressed by Eq. 1.7 is often presented in the form of the famous ‘‘Hillas diagram’’ in Fig. 1.3, which shows candidates for possible accelerator objects depending on the strength of their magnetic field and their characteristic size. It can be seen in the diagram that there are few astrophysical objects capable of transferring an energy of the order of  $10^{20}$  eV to a charged particle, such as pulsars ( $B \sim 10^{13}$  G,  $L \sim 10$  km), AGNs ( $B \sim 10^4$  G,  $L \sim 10$  au) and giant radiogalactic lobes ( $B \sim 0.1 \mu\text{G}$ ,  $L \sim 10$  kpc).

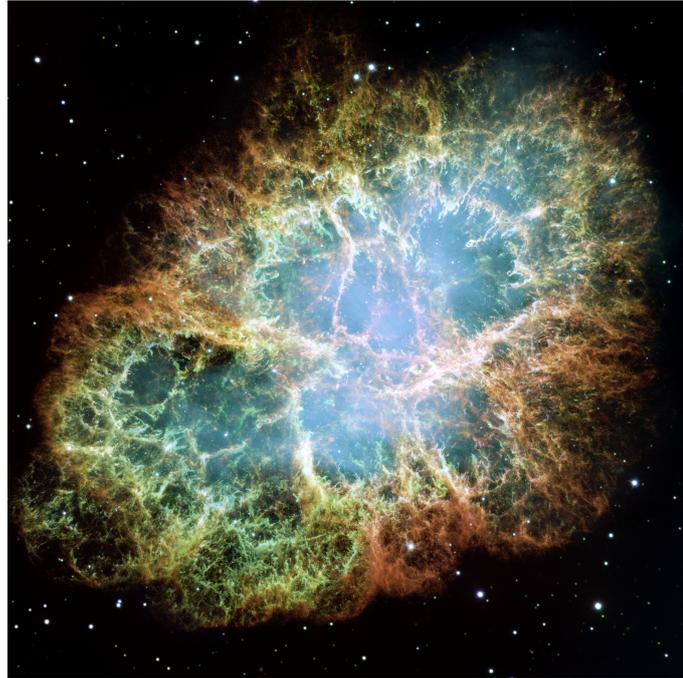
The remnants associated with Galactic supernovae have sizes up to  $L \sim \text{pc}$  with magnetic fields up to the  $\mu\text{G}$  range. According to Eq. 1.7 they should thus be able to accelerate cosmic rays at least



**Figure 1.3:** The Hillas diagram shows the size and intensity of magnetic field of possible accelerators of ultra-high energy cosmic rays [21]. The objects below the diagonal lines (shown for proton and iron primaries) cannot accelerate particles above  $10^{20}$  eV.  $\beta c$  is the characteristic velocity of the magnetic scattering centers.

up to the  $10^{15}$  eV, although more recent calculations show that iron nuclei can be accelerated up to  $\sim 10^{18}$  eV by interacting with multiple SNRs as they propagate through the interstellar medium. This and the fact that the power required to maintain the cosmic-ray density in the Galaxy is comparable to the kinetic energy output rate of Galactic supernovae suggests that they are the predominant sources of cosmic rays below  $\sim 10^{15}$  eV [29]. Until recently, a direct evidence for the acceleration of protons in SNRs was lacking. In 2013, the *Fermi Large Area Telescope* (Fermi LAT) Collaboration identified a characteristic feature in the gamma-ray spectra of two different SNRs, which can be related to the decay of neutral pions produced in interactions of the accelerated protons with interstellar matter, thus providing evidence for the acceleration of protons in SNRs [30]. As an example of an SNR, the Crab Nebula is shown in Fig. 1.4. The observations provided by Fermi-LAT strongly support the assumption that SNRs are sources of Galactic cosmic rays. Additionally, evidence has been recently reported for a source in the Galactic center capable of accelerating particles to  $\sim 10^{15}$  eV [31].

Extremely powerful radio galaxies, such as Fanaroff-Riley class II (FR II) objects, are likely astrophysical accelerators [32]. The Cygnus A radio Galaxy (3C 405), situated at 232 Mpc with a supermassive black hole of  $\sim 25 \times 10^8 M_{\odot}$  at its center [33], is shown in Fig. 1.5. Jets from the central black hole of the active Galaxy and perpendicular to the accretion disk end at a termination shock where the interaction of the jet with the intergalactic medium forms radio lobes and "hot spots". There, charged particles may also be accelerated via first-order Fermi mechanisms, reaching energies around  $\sim 10^{19}$  eV [34]. In contrast to particles accelerated in the inner region of an AGN, particles accelerated at the end of the jets can leave the acceleration region without large energy losses, since the radiation field is much less dense than in the inner region of the AGN. However,

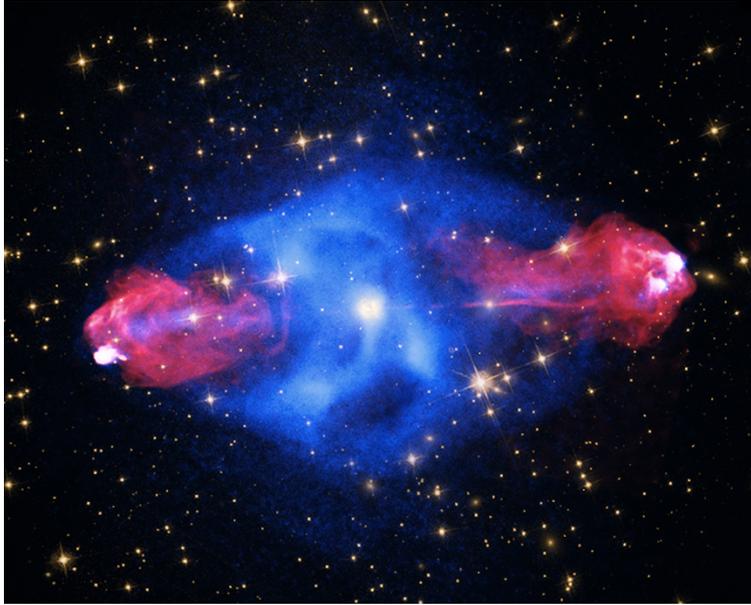


**Figure 1.4:** The Crab Nebula and pulsar. This SNR is particularly interesting since it is one of the brightest persistent sources of gamma-rays in the sky. This image is the largest image ever taken with Hubble’s WFPC2 camera. Credits to NASA, ESA and Allison Loll/Jeff Hester (Arizona State University).

there is considerable debate as to what values have to be considered for  $B$  and  $L$ . The magnetic field within the hot spots of the lobes varies from source to source. Additionally, the observed energy spectrum of the cosmic rays (discussed in Sec. 1.4) can be obtained assuming a magnetic field intensity in the local supercluster of about  $0.1 \mu\text{G}$  [35]. Whereas Galactic magnetic fields are reasonably well studied, extragalactic fields are still poorly understood [36].

In general, when these sites are considered more carefully, great difficulties are found due to either energy losses in the acceleration region or the great distances of known sources from our Galaxy (discussed in Sec. 1.3). The mechanisms of energy loss and gain compete in any acceleration site. With Fermi shock acceleration, the acceleration time is proportional to the mean free path for scattering in the shock wave (see Eq. 1.2), which is itself approximately inversely proportional to the magnetic field strength. Therefore, a certain magnitude of  $B$  is required, not only to confine the particles within the site, but also to accelerate them quickly. However, if the magnetic field is too strong it can cause charged particles to lose their energy via synchrotron radiation. Other strong energy losses are caused by collisions with photons and/or matter at the acceleration site. This leads to the additional requirement that the site must have sufficiently low densities of radiation and matter. For example, the core regions of AGNs are ruled out because of this reason. On the other hand, in the regions of space where the magnetic fields are weak the main mechanism that produces energy loss is the interaction with the CMB. This last process is manifested essentially through the photoproduction of pions and pair-production, which are feasible due to the high energy of the protons<sup>8</sup>. As a consequence, a fraction of the energy gained by the acceleration is lost through synchrotron emission and CMB interactions, further limiting the possible candidates for ultra-high energy cosmic-ray sources.

<sup>8</sup>In the case of heavier composition, nuclei can first interact with infrared background fields that dissociate it, breaking it into lighter particles.



**Figure 1.5:** The Cygnus A radio-active AGN. Each color represents a frequency band in the electromagnetic spectrum: X-ray in blue, radio in red and visible in yellow. Radio emission extends to either side along the same axis for nearly  $\sim 3 \times 10^5$  light-years powered by jets of relativistic particles emanating from the AGN. Hot spots likely mark the ends of the jets impacting surrounding cool, dense material [37].

The FR II hot spots models avoid the energy losses faced by acceleration models in AGN central regions. However, the location of possible sources is problematic for both types of mechanisms. Extremely powerful AGNs with radio lobes and hot spots are rare and far apart. The closest known object is M87 in the Virgo cluster ( $\sim 18$  Mpc away) and could be a main source of ultra-high energy cosmic rays. Although a single nearby source may be able to fit the spectrum for a given strength and structure of the intergalactic magnetic field [38], it is unlikely to match the observed arrival direction distribution. After M87, the next known nearby source is NGC315 which is already too far at a distance of  $\sim 80$  Mpc.

There is one possible accelerator object not included in Fig. 1.3. The *Gamma-Ray Bursts* (GRBs) are transient high-energy phenomena that last from 30 ms to hundreds of seconds and may act as ultra-high energy accelerators. The bursts present an amazing variety of temporal profiles, spectra, and timescales that have puzzled astrophysicists for more than four decades [39]. The angular distribution of these bursts is isotropic within the statistical limits [40]. The average rate of gamma-ray energy emitted by GRBs is  $\sim 10^{44}$  erg Mpc $^{-3}$  yr $^{-1}$ , which is comparable to the integrated cosmic-ray energy above  $10^{19}$  eV in a redshift independent cosmological distribution of sources [41].

The most popular interpretation of the GRB phenomenology is that the observable effects are due to the dissipation of the kinetic energy of a relativistic expanding plasma wind, a “fireball”, whose primal cause is not yet known [42]. Cosmological fireballs may generate ultra-high energy cosmic rays through Fermi acceleration by internal shocks. In this model, the generation spectrum is estimated to be  $dN/dE \propto E^{-2}$  which is consistent with observations provided the efficiency with which the kinetic energy is converted to gamma-rays and ultra-high energy cosmic rays are similar. There are several problems with the “GRB - cosmic rays” common origin proposal (see [42] for details). One of the problems is that the origin of the highest energy end of the cosmic-ray spectrum has to be explained with GRBs closer than  $\sim 50$  Mpc. However, only one GRB is expected

to have occurred within this region over a period of 100 yr. Therefore, a very large dispersion of  $\gtrsim 100$  yr in the arrival time of protons produced in a single burst would be a necessary condition.

### 1.2.2 Top-down scenarios

The basic idea of a top-down origin of cosmic rays can be traced back to Georges Lemaître [43] and his theory of the *Primeval Atom*, the precursor to the Big Bang model of the expanding Universe. The entire material content of the Universe and its expansion, according to Lemaître, originated from the "super radioactive disintegration" of a single atom of extremely large atomic weight, the Primeval Atom. The cosmic rays were envisaged as the energetic particles produced in intermediate stages of its decay. They were thus "glimpses of the primeval fireworks" [43]. Indeed, Lemaître regarded cosmic rays as the main evidential relics of the Primeval Atom in the present Universe.

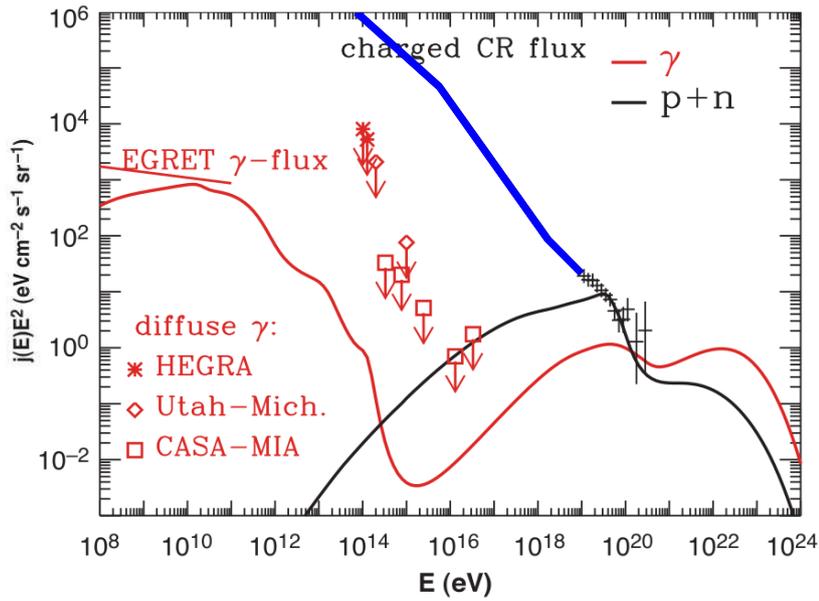
In modern versions of the Primeval Atom, some cosmologists have related the origin of ultra-high energy cosmic rays to other basic ideas beyond the Standard Model. Most top-down models were formulated to avoid the energy loss of cosmic rays due to the interaction with the CMB, the so-called GZK effect (see Sec. 1.3.2), after the AGASA Collaboration claimed to have measured the cosmic-ray spectrum without a sharp suppression at the highest energies [44] (see Sec. 1.4). Even more exotic models were proposed to that end. For example, some theories predict a Lorentz invariance violation that suppresses the cross section for inelastic collision between nucleons and CMB photons [45].

These models avoid the problems of particle acceleration by postulating supermassive *X particles* with an energy of  $\sim 10^{24}$  eV. The spontaneous rupture of symmetry between fermionic and bosonic matter in the early stages of the Universe would have generated the X particles. Other models that explain the X production include the annihilation of ultra-energetic neutrinos with a background of neutrinos inherited from the early Universe almost as dense as the CMB or the decay of "topological defects", i.e. regions of space-time where the density of matter and energy remains as high as at the origin of the Universe.

The decay of any of these relics produces leptons and quarks. The leptons rapidly cascade down to photons and electrons. The quarks decay into jets of  $\sim 10^4 - 10^5$  hadrons (pions and a small fraction of baryons). In turn, the charged pions would decay in electrons and neutrinos and the neutral pions into photons [28, 46]. In order for the decay products of the X particles to be observed as cosmic rays today, they must decay in a recent cosmological epoch, or equivalently at non-cosmological distances ( $\lesssim 100$  Mpc) from Earth. Otherwise the decay products of the X particles lose all energy by interacting with the background radiation fields<sup>9</sup> (see Sec. 1.3).

Typical features of these scenarios are a predominant contribution of photons to the cosmic-ray flux at sufficiently high energies (from  $\sim 10\%$  at  $10^{19}$  eV up to  $50\%$  at  $10^{20}$  eV [47]) as shown in Fig. 1.6 and a notorious presence of neutrinos in the observed spectrum above  $\gtrsim 10^{19}$  eV [21]. However, these predictions strongly depend on the *Universal Radio Background* (URB), whose strength is poorly determined in the relevant MHz regime [48], and extragalactic magnetic fields. Although a variety of physical parameters, e.g., the mass of the supermassive X particles, are

<sup>9</sup>A possible exception is the case of neutrinos of sufficiently high energy originated by the decay of an X particle at large cosmological distances  $\gg 100$  Mpc. These neutrinos may then produce high-energy nucleons and/or photons within 100 Mpc from Earth [21].

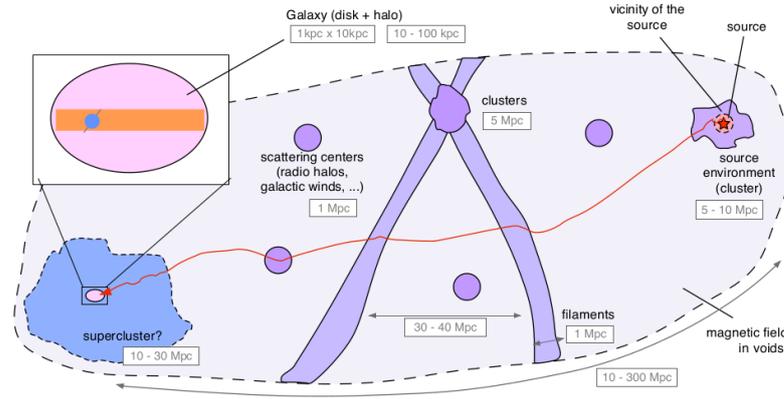


**Figure 1.6:** The predicted spectra of nucleons (black line) and photons (red line) from a top-down model involving the decay of  $X$  particles of mass  $10^{25}$  eV into two quarks, released from homogeneously distributed topological defects, with an extragalactic magnetic field  $\ll 10^{-11}$  G. Also shown are piecewise power-law fits to the observed cosmic-ray flux below  $10^{19}$  eV (blue line), the measurement of the diffuse gamma-ray flux between 30 MeV and 100 GeV by the EGRET instrument [52] (red straight line) as well as upper limits on the diffuse gamma-ray flux from various experiments at higher energies (see [21] for more details). The points with  $1\sigma$  error bars show the combined data from the Haverah Park, the Fly’s Eye, and AGASA above  $10^{19}$  eV. Figure modified from [53].

loosely subject to observational and theoretical constraints, they are bounded by the current non-observation of photons above  $10^{18}$  eV (see Sec. 1.5.3). However, recent models concerning *Super Heavy Dark Matter* (SHDM) particles have been re-proposed to accommodate the existing photon limits and constrain the SHDM particles lifetime [49, 50].

In addition, top-down models predict an energy spectrum which is considerably harder than the case of shock acceleration,  $dN/dE \propto E^{-1.9}$  above  $\sim 7 \times 10^{19}$  eV [51] and without a cut-off at the highest energies. Some top-down models may also produce distinctive features on the cosmic-ray spectrum at lower energies. The secondary particles produced in the decay of the  $X$  particle may induce electromagnetic cascades by interactions with the low-energy photon backgrounds, contributing to the diffuse gamma-ray flux between 30 MeV and 100 GeV. This contribution is close to the flux measured by the EGRET detector on board the Compton Gamma-ray Observatory satellite [52]. These features are summarized in Fig. 1.6 for a specific top-down model based on the decay of topological defects.

In summary, these models are disfavoured by recent experimental observations whereas they can not be completely ruled out. There are also hybrid models that include elements from both bottom-up and top-down models. The most successful one is the Z-burst model [54] which is based on the idea that somewhere in the Universe neutrinos of ultra-high energy are generated. These neutrinos annihilate with cosmological neutrinos and generate  $Z_0$  bosons which decay and generate a local flux of nucleons, pions, photons and neutrinos. In any case, the accurate knowledge of the mass composition and shape of the observed energy spectrum would serve to distinguish between different acceleration and decay scenarios.



**Figure 1.7:** A schematic representation of magnetized regions intervening in ultra-high energy cosmic rays propagation. Their approximative characteristic length scales are indicated in grey [55].

## 1.3 Propagation of Cosmic Rays

While propagating from their sources to the observer, cosmic rays interact with background radiation fields. These interactions may affect their energy, composition and their propagation direction. Each process may leave a variety of signatures on the observables and generate secondary neutrinos and photons.

### 1.3.1 Magnetic deflection

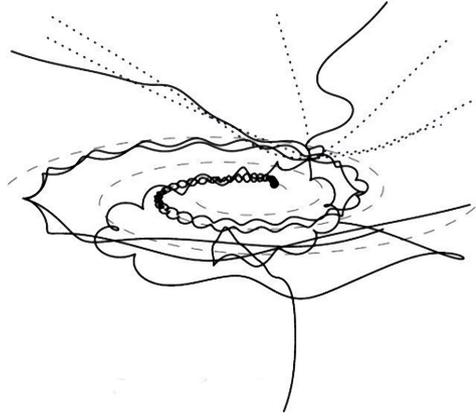
The absence of powerful astrophysical counterparts in the arrival directions of ultra-high energy cosmic rays is probably related to the effect of cosmic magnetic fields that deflect and delay particles during their propagation. Charged particles are subject to the influence of magnetic fields in the source environment, in the intergalactic medium, and in the Galaxy, as depicted in Fig. 1.7. Since very little is known about cosmic magnetic fields, the parameter space for an accurate description is quite large.

It is known that lower energy cosmic rays, those which are thought to originate from within the Galaxy, must propagate an average distance of  $\sim 1$  Mpc. This implies that Galactic cosmic rays diffuse through the Galaxy and so arrive isotropically at the Earth. On the other hand, at ultra-high energies, cosmic rays are most likely extragalactic in origin since they cannot be effectively contained within the Galaxy. For example, the gyroradius for a proton at  $10^{18}$  eV in a field of  $\sim 3 \mu\text{G}$  is about 300 pc, as thick as the Galactic disc.

The angular deflection from the incoming cosmic-ray path, as a function of the traveled distance and primary energy, is given by the Larmor radius  $r_L$ :

$$r_L \sim \frac{E}{Z \cdot B_{\perp}} \quad (1.8)$$

where  $Z$  is the charge of the particle and  $B_{\perp}$  is the magnetic field component perpendicular to the particle momentum. If the magnetic field is constant over the traveled distance  $d$ :



**Figure 1.8:** Examples of possible trajectories of nuclei with  $E/Z = 10^{18}$  eV (solid lines) and  $E/Z = 10^{19}$  eV (dotted lines) in the Galactic magnetic fields. Dashed lines indicate the Galactic spiral arms [57].

$$d\theta(E, d) \sim \frac{d}{r_L} \propto Z \cdot \left( \frac{E}{10^{20} \text{ eV}} \right) \cdot \left( \frac{B_{\perp}}{\text{nG}} \right) \cdot \left( \frac{d}{\text{Mpc}} \right) \quad (1.9)$$

This means a deviation of less than  $1^\circ$  for a proton of energy  $10^{20}$  eV in a magnetic field of  $\mu\text{G}$  and on a distance of few kpc, or equivalently in a field of nG over a distance of the order of the  $\text{Mpc}^{10}$ . In Fig. 1.8 the trajectories of nuclei with  $E/Z = 10^{18}$  eV (solid lines) and  $E/Z = 10^{19}$  eV (dotted lines) in the Galactic magnetic field model are drawn. Therefore, the existence of a light component in the cosmic rays flux at the highest energies is crucial for Astronomy with charged particles to be feasible. In this case, the information on the incoming direction is conserved and the correlation with sources may be established. In particular, the regular component of the Galactic magnetic field can distort the angular images of cosmic-ray sources: the flux may appear dispersed around the source or globally translated in the sky with a small dispersion. Since Galactic magnetic fields are not uniform in the sky, angular deflections also depend on the observed direction [56].

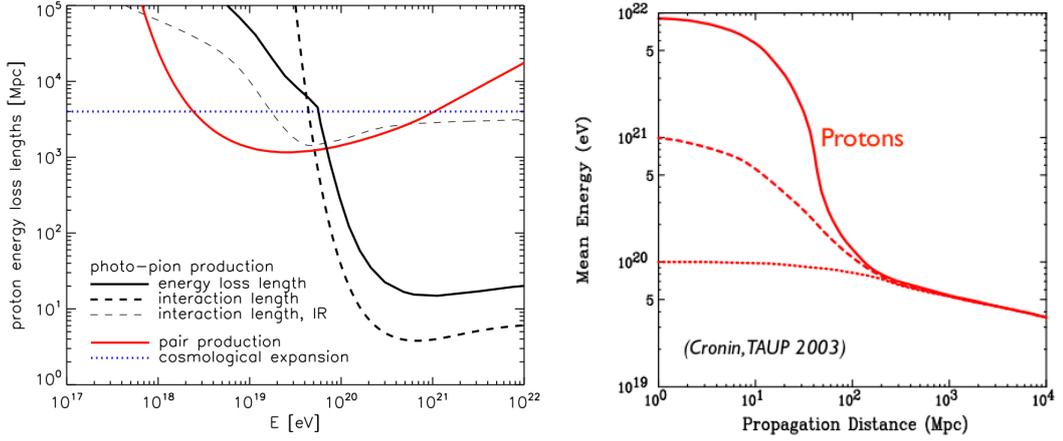
### 1.3.2 GZK effect

In the intergalactic medium, cosmic rays primarily interact with the CMB photons at the highest energies, and with infrared (IRB) and radio background photons at slightly lower energies [58]. According to the GZK effect, cosmic rays would interact with the CMB during their propagation via photopion production. The process for protons, which involves an intermediate  $\Delta$ -resonance, is:



The energy threshold for this process given a background photon with energy  $\epsilon$  is:

<sup>10</sup>Employing a more realistic model, the dependence of the distance  $d$  is found to be weaker and the coherence length of interaction must be taken into account [21]. However, the dependence of  $d\theta$  with the energy and the magnetic field is correctly described by this simple model.



**Figure 1.9:** (Left) Proton energy loss lengths in terms of its energy. The black solid line represents the energy loss length for photopion production on CMB and IRB photons; red solid line for pair production on CMB photons. Dashed lines represent the interaction length (or mean free path to interaction) for photopion production on CMB photons (thick) and IRB photons (thin). The dotted line indicates the losses due to cosmological expansion [55]. (Right) The mean energy of protons interacting with the CMB as a function of the traveled distance. The curves correspond to three different initial energies [59].

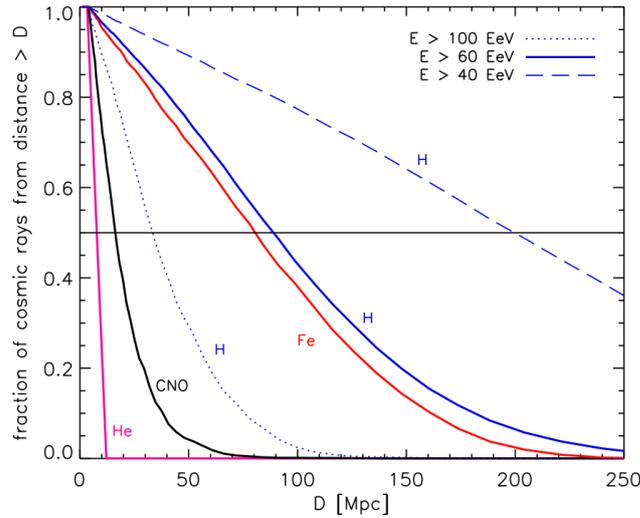
$$E_{\text{th}} = \frac{m_{\pi}}{4\epsilon} \cdot (2m_p + m_{\pi}) \sim 2 \times 10^{20} \left( \frac{\epsilon_{\text{CMB}}}{\epsilon} \right) \text{ eV} \quad (1.11)$$

with  $\epsilon_{\text{CMB}} \simeq 2.7k_{\text{B}}T_{\text{CMB}} \simeq 6 \times 10^{-4} \text{ eV}$ , the mean energy of a CMB photon, leading to  $E_{\text{thr}} \sim 10^{20} \text{ eV}$ . The energy loss length  $x_{\text{loss}} \equiv |E^{-1}dE/cdt|^{-1}$  is shown in Fig. 1.9, left. Above  $E \sim 6 \times 10^{19} \text{ eV}$  the distance that particles can travel without losing their energy shortens considerably. If cosmic rays originate from cosmological distances, their flux above this energy should be consequently suppressed, producing the well-known GZK feature in the cosmic rays spectrum (see Sec. 1.4). This property further imposes that the sources of the observed ultra-high energy cosmic rays at a given energy should be located in our local Universe, within a distance  $l \lesssim x_{\text{loss}}(E)$ . The energy loss per interaction with the CMB is  $\sim 20\%$ . Due to the mean free path of this interaction, extragalactic protons traveling farther than distances on the order of 100 Mpc, and with energies above the threshold, will never be observed on Earth. This defines the *GZK horizon*, as shown in Fig. 1.9, right.

At energies below  $E \sim 7 \times 10^{19} \text{ eV}$  cosmic-ray protons are affected mainly by the Bethe-Heitler pair production mechanism (see Eq. 1.12). The same interactions can occur with IR, optical, and UV backgrounds in intergalactic space, but this contribution is almost irrelevant over the entire energy range. The threshold energy for this process is shown in Eq. 1.13. Since its inelasticity is only 0.1%, it is not dominant at higher energies compared to the energy loss for the photopion production, as it can be seen in Fig. 1.9, left.



$$E_{\text{th}} = \frac{m_e(m_p + m_e)}{\epsilon} \simeq 8 \times 10^{17} \left( \frac{\epsilon_{\text{CMB}}}{\epsilon} \right) \text{ eV} \quad (1.13)$$



**Figure 1.10:** The fraction of cosmic rays that survives propagation over a distance larger than  $D$ , for protons above  $4 \times 10^{19}$  eV,  $6 \times 10^{19}$  eV and  $10^{20}$  eV and for He, the CNO group and Fe above  $6 \times 10^{19}$  eV. The black solid line shows where 50% of a given species can originate for a given atomic mass and energy. At trans-GZK energies ( $E > 6 \times 10^{19}$  eV), only protons and iron survive the propagation over  $D > 50$  Mpc [55].

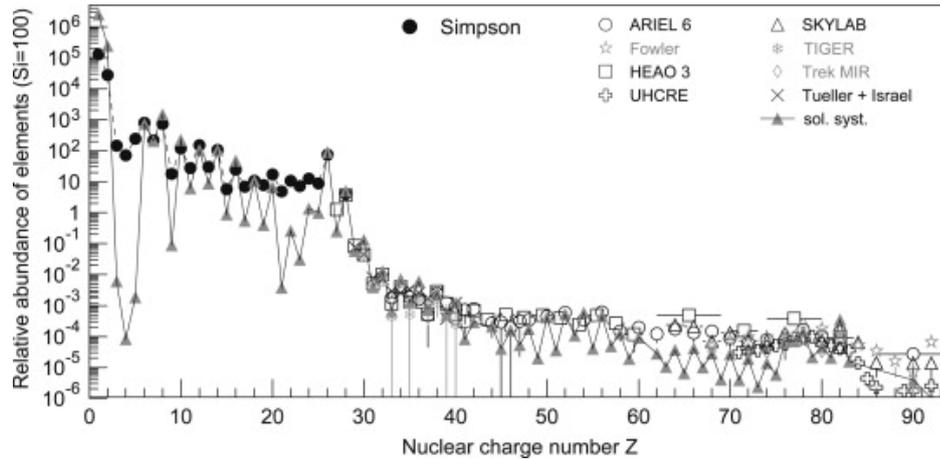
### 1.3.3 Energy losses for nuclei

For primary cosmic rays with mass number  $A > 1$ , the photo-disintegration processes come into play, both with the CMB and IRB, according to the following relations:

$$\begin{aligned}
 A + \gamma_{\text{cmb,irb}} &\rightarrow (A - 1) + N \\
 A + \gamma_{\text{cmb,irb}} &\rightarrow (A - 2) + 2N \\
 A + \gamma_{\text{cmb,irb}} &\rightarrow A + e^+ + e^-
 \end{aligned} \tag{1.14}$$

Given that the energy of the primary nuclei is shared between nucleons, the threshold energy for these processes is typically higher than for proton. The photo-disintegration process leads to the ejection of one or several nucleons  $N$  from the nucleus. At ultra-high energies, nuclei photo-disintegrate on CMB and IR-UV photons through three main types of processes that contribute at increasing energy ranges: the *Giant Dipolar Resonance*, the *Quasi Deuteron* process and the *Baryonic Resonance*. In a first approximation, the Lorentz factor of the primary nucleus can be considered unchanged through these interactions. Nuclei also experience photopair production that decreases the Lorentz factor without affecting the number of nucleons. At larger energies (above  $10^{21}$  eV for iron) the photo-pion production starts to become relevant and dominates on the photo-disintegration processes for energies larger than  $3 \times 10^{21}$  eV. The photo-disintegration processes with CMB dominate through most of the energy range, while the pair production, as well as the interaction on IRB, Optical and UV backgrounds are negligible.

One remarkable effect of the propagation of nuclei is that nuclei with mass number  $A < 20$  cannot travel farther than few tens of Mpc without disintegrating (see Fig. 1.10). In particular, heavy nuclei could be found in abundance at trans-GZK energies only if the composition were essentially dominated by the iron nuclei group.



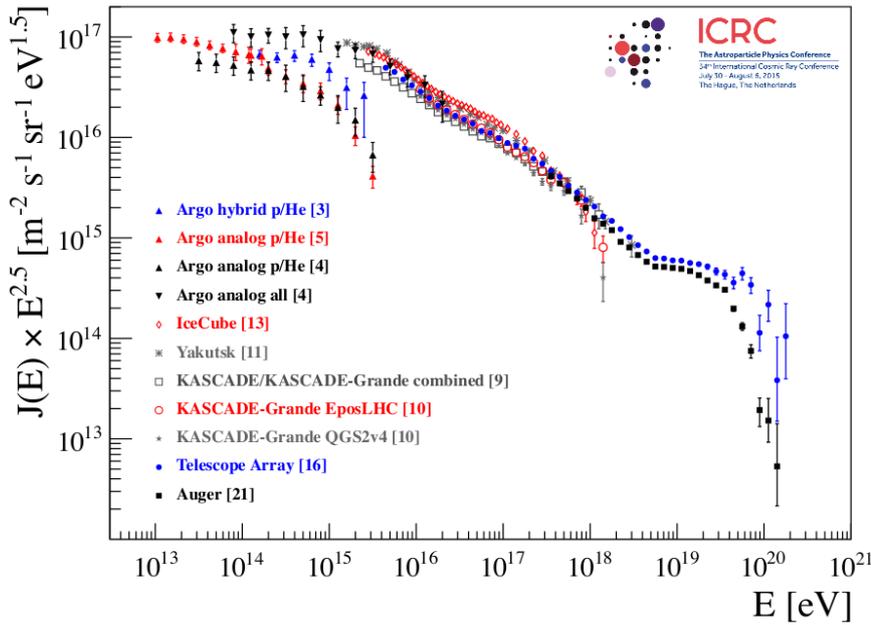
**Figure 1.11:** The abundance of elements in cosmic rays as a function of their nuclear charge number  $Z$  at energies around 1 GeV per nucleon, normalized to Si = 100 [61]. The abundance for light nuclei is shown in black dots [62], while heavy nuclei, as measured by several experiments, are shown with the quoted marker styles [63]. In addition, the abundance of elements in the Solar System is represented by gray triangles [64].

## 1.4 Energy spectrum of Cosmic Rays

The energy spectrum of the cosmic rays spans more than 12 orders of magnitude in energy, between  $10^9$  eV to more than  $10^{20}$  eV, and more than 30 orders of magnitude in flux. The observed differential flux follows a power law  $E^{-\gamma}$  above  $10^{15}$  eV with  $\gamma \simeq 3$ , implying that the number of cosmic rays decreases by a factor  $\sim 10^3$  for each decade increase in energy, being  $\sim 1$  particle  $\text{m}^{-2} \text{yr}^{-1}$  above  $10^{15}$  eV and only  $\sim 1$  particle  $\text{km}^{-2} \text{century}^{-1}$  at  $10^{20}$  eV.

The solar wind prevents charged particles from the interstellar medium from entering the Solar System below  $\sim 10^8$  eV. The Sun itself is a sporadic source of nuclei and electrons that are accelerated by shock waves at its corona or by magnetic energy released during solar eruptions. In the latter case, the flux of particles impinging the Earth increases by a factor between  $10^2$  and  $10^6$  over a period of hours to days<sup>11</sup>. The maximum energy that solar cosmic rays can possess is typically between  $10^7$  eV and  $10^8$  eV, reaching a maximum value of  $10^9$  eV (once per year) and  $10^{10}$  eV (once per decade) [60]. Above  $\sim 10^{10}$  eV, the effect of the solar wind becomes negligible and the cosmic rays of the Galactic medium begin to enter the Solar System.

Between  $10^{11}$  eV and  $10^{14}$  eV, the cosmic-ray composition can be measured directly from balloons or with detectors carried on satellites. These measurements point towards a predominantly light composition, mainly protons and  $\alpha$  particles [60]. A more detailed comparison with the relative abundances of the components of the Solar System, as shown in Fig. 1.11, shows that the presence of H and He in the cosmic rays is smaller than in the Solar System, which is not completely understood. It could indicate that the heavier elements are more easily accelerated or could be a consequence of the mass composition at the sources. On the other hand, some nuclei such as Li, Be, Ti and Cr have higher abundances than in the Solar System, which is a direct



**Figure 1.12:** The cosmic-ray spectrum multiplied by  $E^{2.5}$  measured by several experiments [65]. The breaks in the energy spectrum denote changes of the mass composition or acceleration mechanisms and origins of the primary particles (see text for details).

consequence of the nuclear spallation of C, O and Fe nuclei.

Above  $10^{14}$  eV, the flux becomes so low that only ground-based experiments with large apertures and long exposure times can hope to acquire a significant number of events. Such experiments exploit the atmosphere as a giant calorimeter. The incident cosmic radiation interacts with the atomic nuclei of air molecules and produces air showers which spread out over large areas (see Chapter 2). There is a general consensus about the Galactic nature of cosmic rays up to  $10^{15}$  eV, since the shock waves associated with the SNRs are efficient accelerators up to these energies (see Sec. 1.2).

From  $10^{14}$  eV up to the highest energies, the energy spectrum follows a power law with spectral index  $\gamma \simeq 3$ . The features of the spectrum are best described as small changes in the spectral slope or *breaks*. The ultra-high energy cosmic-ray spectrum is shown, multiplied by a factor of  $E^{2.5}$ , in Fig. 1.12. The equivalent LHC energy in proton-proton collisions with a fixed target is  $\sim 10^{17}$  eV. Above this energy, particle physics is no longer directly constrained by data from accelerator experiments. Because of differences between the energy calibration of different experiments, it is often difficult to compare the absolute energy of the various features. It is much easier to compare the spectral slopes. The breaks in the spectrum could be explained in terms of changes in the propagation, the location of the sources (fundamentally inside or outside our Galaxy) or the mass composition of the primary flux.

### 1.4.1 Features of the energy spectrum

**The Knee** The first of these breaks, known as the *knee*, occurs at  $\sim 10^{15.5}$  eV where the energy dependence changes from  $E^{-2.7}$  to  $E^{-3.0}$  [66]. Above the knee, the light particles are not confined

<sup>11</sup> These events are more frequent during the phases of greater activity of the solar cycle.

in the regions of acceleration since their Larmor radius is similar to the characteristic length of the shock waves and, therefore, they escape before being accelerated. Accordingly, the composition would tend to become heavier. At energies around the knee, the results of the KASCADE Collaboration show that there is a gradual change in the composition from light to heavy elements [67]. The cosmic rays flux in this region could also be explained by a superposition of power laws (corresponding to the different types of primaries) with dedicated breaks (knees) at different energies<sup>12</sup>.

**The Second Knee** The next change, less clear than the first and still under discussion, has been observed at  $\sim 10^{17.6}$  eV by detectors like HiRes [72] and Akeno [73], but has not been seen at these energies by more recent experiments like Auger and KASCADE-Grande. Moreover, KASCADE has reported the indication of this inflection at  $\sim 10^{16.9}$  eV assuming a heavier mass composition [74]. The break would correspond to a further softening to an index of  $\sim -3.3$ . It receives the name of *second knee*. If confirmed, the second knee could represent a limit for the acceleration of Galactic heavy elements since between  $10^{17}$  eV and  $10^{18}$  eV the Galactic SNRs would cease to be effective accelerators. Otherwise it could indicate the energy above which the extragalactic cosmic-ray component becomes dominant. In any case, the mass composition of the primary cosmic rays is the key to describe this transition and rule out acceleration and propagation models.

**The Ankle and the suppression** Around  $\sim 10^{18.5}$  eV, the spectrum retreats by retaking the exponent  $-2.7$  in a break that the community baptized as the *ankle* of the spectrum. Beyond the second knee, the particles have enough energy to travel from extragalactic sources in less than the Hubble time [75] and the Galactic magnetic fields would not be able to confine heavy nuclei. Therefore, the extragalactic component should begin to dominate somewhere in the spectrum between the second knee and the ankle.

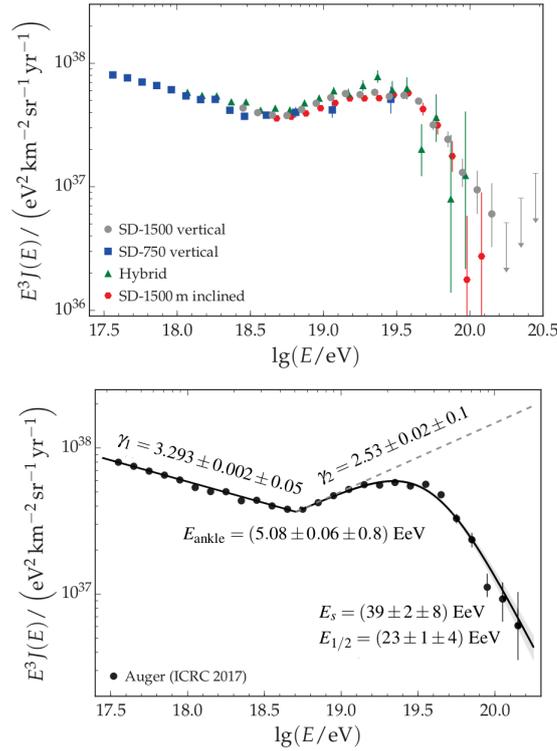
Above the ankle, the cosmic rays would necessarily be extragalactic. At  $\sim 5 \times 10^{19}$  eV, an abrupt suppression of the flux is observed. The cutoff interpretation is still unclear. If the particles were protons, their interaction with the CMB would induce energy losses, as already explained in Sec. 1.3.2. On the other hand, if the flux was composed by heavy nuclei a suppression at such high energies could be due to the limit of acceleration in the extragalactic sources.

The energy spectrum at the highest energies is measured with an unprecedented precision at the Pierre Auger Observatory with an exposure exceeding  $50000 \text{ km}^2 \text{ sr yr}$  [76]. The spectrum is obtained from the data of the  $3000 \text{ km}^2$  surface detector array, the hybrid events detected simultaneously by the fluorescence telescopes and the surface array, and the data from a denser array. All the detectors of the Pierre Auger Observatory are extensively described in Chapter 3. The individual spectra are shown in the top panel of Fig. 1.13, while the energy spectrum obtained combining all measurements is shown in the bottom panel.

In summary, the cosmic-ray flux and its spectral distribution show the astrophysical richness and complexity of the region from  $\sim 10^{17}$  eV onwards. The different changes in the spectrum are mainly related to changes in the nature of the arriving particles and the energy spectrum in, and the type of, the sources.

---

<sup>12</sup>The energy spectrum manifests a flattening at about  $10^{16}$  eV, a feature often called *low energy ankle*, firstly reported by the KASCADE-Grande Collaboration [68], and recently confirmed independently by the Yakutsk [69], IceCube [70] and Telescope Array [71] Collaborations. Its interpretation is still unclear.



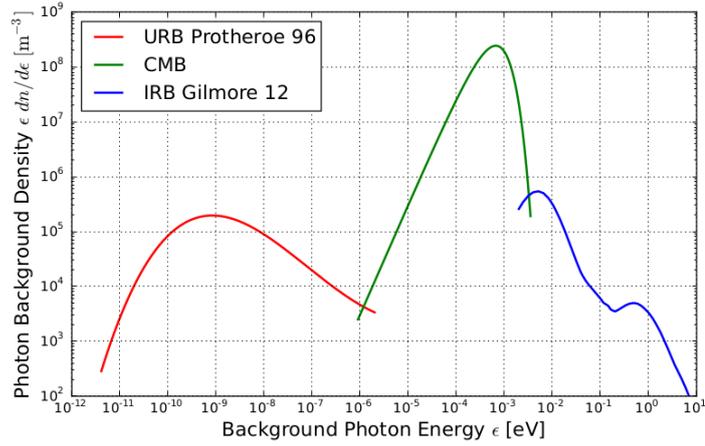
**Figure 1.13:** The energy spectra measured with individual detectors (top) and combining all detectors (bottom) by the Pierre Auger Collaboration [77].

## 1.5 Ultra-high energy photons

The ultra-high energy cosmic rays have been known for more than 70 years, and yet their origin remains elusive. Similarly, the possibility of an ultra-high energy photon component of the cosmic radiation is one of the open problems of Astroparticle Physics. The most energetic and violent, but also less understood, astrophysical objects in the Universe are expected to produce cosmic rays, with an accompanied flux of ultra-high energy photons and neutrinos. The simultaneous observation of these particles, the so-called *multi-messenger approach*, is a key ingredient for discovering the sources themselves and for better understanding the underlying mechanisms responsible for their violent activity.

### 1.5.1 Sources and generation mechanisms

Ultra-high energy photons and neutrinos are expected from the decay of  $\pi^0$  and  $\pi^\pm$ , which are produced by any primary process. One possible generation channel is the GZK effect, which involves the CMB, as discussed in Sec. 1.3.2. In this case, the photons and neutrinos are usually called *cosmogenic*. Since they are produced during the propagation of the hadronic cosmic rays, they may conform a *diffuse* isotropic flux. Another possible generation channel involves direct interaction between matter and radiation at the cosmic-ray sources. In this second case, the photons and neutrinos are usually called *astrophysical*. Since they point back to their sources (photons and neutrinos are not deflected by magnetic fields), they could carry valuable information about the acceleration regions.



**Figure 1.14:** The energy spectrum of the different components of the cosmic photon background: CMB in green, URB in red [78] and IR in blue [79].

**Cosmogenic photons** Ultra-high energy particles interact with the cosmic photon background. This background can be separated into several parts, three of them are relevant for cosmic-ray interactions. The URB has the largest differential photon density in the range between  $10^{-12}$  eV to  $10^{-6}$  eV. The CMB dominates up to about 3 meV when the IRB and *Optical Background* effects begin to be non-negligible until approximately 10 eV. Therefore, the photon background covers about 13 orders of magnitude of the electromagnetic spectrum as shown in Fig. 1.14.

A guaranteed flux of cosmogenic photons is expected from the decay of the neutral pions produced by the GZK effect. The complete generation channel, extended from Eq. 1.10, is:

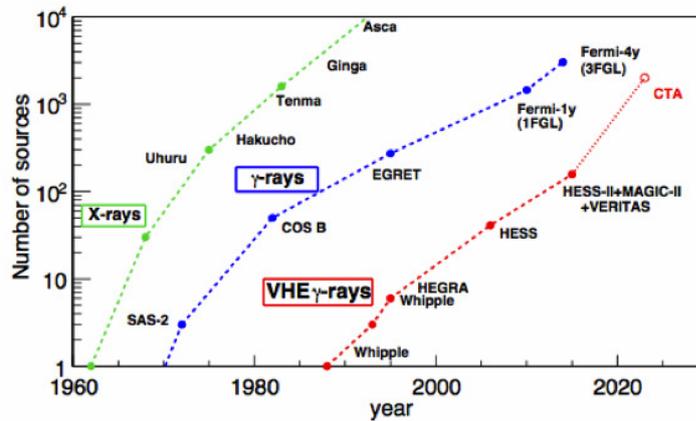


If ultra-high energy cosmic rays are mostly nuclei, they would interact with radiation backgrounds primarily through photo-disintegration, breaking up into lighter nuclei and nucleons [80] (see Sec. 1.3.3). As these nucleons would often be below the energy threshold for pion-production, fewer ultra-high energy photons and neutrinos would be produced [81]. The flux of ultra-high energy photons in this scenario would be up to two orders of magnitude smaller than the flux calculated in the pure-proton scenario. Thus, the detection of ultra-high energy photons provides an additional probe of the composition of ultra-high energy cosmic rays.

Photons emerging from the pion decay carry away  $\sim 20\%$  of the original nucleon energy [82]. Since the energy threshold for the GZK process is  $\sim 3 \times 10^{19}$  eV, photons emerging from this process would have a minimum energy of  $\sim 3 \times 10^{18}$  eV. However, the energy of the photons observed on Earth would be lower due to farther interactions with radiation fields during their propagation to Earth (see Sec. 1.5.2). In this scenario, the observed flux would be composed of secondary photons produced by the cascading of the primary photons.

In the theoretical models that follow the top-down approach, the primary process is directly related to the decay or the annihilation of the postulated super-massive particles. In the decay or annihilation processes, typically two or more quarks and gluons are produced, which initiate QCD cascades [47]. Eventually, the partons in the cascades hadronize, and thus cosmic rays are produced along with a large number of pions which leads to a predicted flux of photons that is, for some models, two orders of magnitude larger than in the bottom-up models [82].

While the LHC failed to discover any dark-matter candidate, models of dark matter which are



**Figure 1.15:** The number of sources detected over time for various energy domains and different experiments [87].

beyond the reach of man-made accelerators are becoming more and more popular. In particular, the SHDM model, originally put forward to explain the apparent excess of  $E \gtrsim 10^{20}$  eV cosmic rays (presently disfavoured), has its own cosmological motivation [83]. Its important prediction is a significant fraction of secondary photons among the decay products of these super-massive particles. Though this scenario is largely constrained by the upper limits on the photon flux (see Sec. 1.5.3), it is not completely ruled out [84].

**Astrophysical photons** Galactic photons may be produced during the propagation of ultra-high energy cosmic rays through the central region of the Milky Way. In this case, the interaction with starlight and infrared photons emitted by dust re-radiation, similarly to the GZK process, may result in production of pions which generate a secondary flux of photons and neutrinos [85].

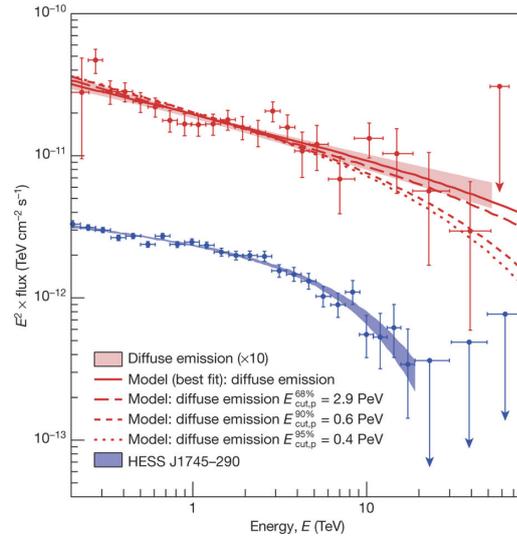
Ultra-high energy photons may be produced in the vicinity of acceleration regions as well through cosmic-ray interactions with surrounding radiation and matter [86]. However, it is presently unknown whether the acceleration of particles up to  $\sim 10^{17} - 10^{18}$  eV may happen in any single object or region in the Galaxy. In any case, these objects are not expected to be numerous. Therefore a certain degree of clustering of the arrival directions of photons would be expected in this scenario.

On the other hand, many sources of photons have been identified at lower energies by both space- and ground-based detection systems. The success of the field can be illustrated by the number of sources detected at the  $10^9$  eV and  $10^{12}$  eV energy regimes, which are commonly known as the high-energy and very high-energy domains, as depicted in Fig. 1.15 by blue and red lines respectively. In the following, a few selected experiments are presented and some of their major achievements are described.

The *High Energy Stereoscopic System* (HESS), consisting of four imaging atmospheric Cherenkov telescopes in Namibia<sup>13</sup>, has observed many extragalactic objects in the search for very high-energy gamma-ray emission [89, 90]. These objects include AGNs, notably blazars and various kinds of galaxies such as seyfert<sup>14</sup>, radio and starburst [91].

<sup>13</sup>An upgrade of this experiment through the installation of a 28 m telescope at the centre of the original array in 2012, marked the onset of the so-called HESS II era [88].

<sup>14</sup>Seyfert galaxies have quasar-like nuclei (very luminous, distant and bright sources of electromagnetic radiation) with very high surface brightnesses but unlike single quasars, their host galaxies are clearly detectable.



**Figure 1.16:** The gamma-ray spectrum in the direction of the Galactic center measured by HESS (red dots). The error bars represent the  $1\sigma$  statistical error and the arrows represent the upper limits at 90% CL. The shaded areas represent the  $1\sigma$  confidence bands of the best-fit spectra. The measured spectrum from the unidentified point source HESS J1745-290 is also shown for comparison (blue dots) [31].

The HESS Collaboration measured a gamma-ray spectrum at energies around  $10^{12}$  eV in the direction of the Galactic center region [31]. The measured diffuse gamma-ray emission follows a power law according to  $dN/dE \propto E^{-\gamma}$  with a spectral index  $\gamma = 2.32 \pm 0.05_{\text{stat}} \pm 0.11_{\text{sys}}$  without any observation of a cutoff or a spectral break up to  $\sim 10^{13}$  eV. Since these photons result from the decay of  $\pi^0$  produced by hadronic interactions, the derivation of such a hard power-law spectrum implies that the spectrum of the parent protons should extend to energies close to  $10^{15}$  eV. The hadronic interactions between  $10^{15}$  eV protons could also be studied by the observation of emitted neutrinos or X-rays from the synchrotron emission of secondary electrons and positrons. However, the expected fluxes of neutrinos and X-rays are below or at best close to the sensitivities of the current instruments. This was the first robust detection of a very high-energy cosmic hadronic accelerator which operates as a source of  $10^{15}$  eV particles (a *PeVatron*) [31]. The measurements from HESS are displayed in Fig. 1.16.

The *Major Atmospheric Gamma Imaging Cherenkov* (MAGIC) telescope is currently the largest stand-alone operating telescope of its kind. Located in the Canary Island of La Palma, Spain, it is based on the same detection technique as the HESS telescopes. The site has been chosen previously for the HEGRA experiment, and it also hosts several optical telescopes and solar observatories, including the 10.4 m diameter telescope CANARIAS. MAGIC was designed to achieve an energy threshold as low as possible allowing to detect gamma-rays in an unexplored energy range: the gap between  $10^{10}$  eV and  $\sim 2 \times 10^{11}$  eV, where the particle flux is too low to allow significant detections with the EGRET satellite and the Cherenkov light produced in air showers is insufficient for a clear detection by the previous generation of ground-based Cherenkov telescopes. A second 17 m diameter telescope, named MAGIC-II, has been commissioned at the same site as the first telescope during 2009. It is located at a distance of 80 m from the first telescope and can be operated independently or in stereoscopic mode with its predecessor.

The MAGIC telescopes have been employed to discover several Galactic and extragalactic sources including AGNs at  $z > 0.2$  [93], a binary system LSI+61 303 [94] and to detect the



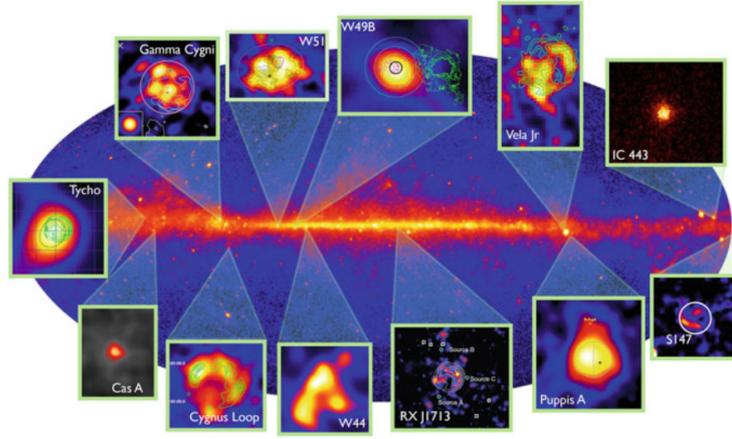
**Figure 1.17:** The two 17 m diameter MAGIC telescope system operating at the “Roque de los Muchachos” observatory in La Palma. The front telescope is the MAGIC-II [92].

gamma-ray emission at energies above  $10^{11}$  eV from the SNR IC443 [95]. Perhaps one of the most important contributions to the field by the MAGIC Collaboration was the detection of pulsed gamma-ray emission from the Crab pulsar at energies as low as  $2.5 \times 10^{10}$  eV [96]. This indicates that the emission occurs far out in the magnetosphere, hence excluding the polar-cap scenario as a possible explanation.

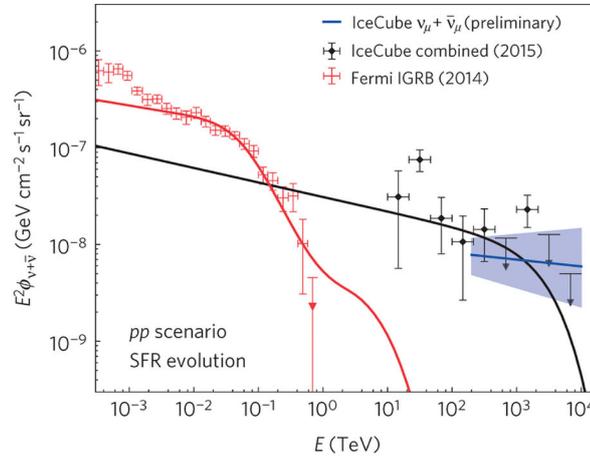
The *Fermi-Large Area Telescope* (Fermi-LAT) experiment, commissioned in 2008, is mainly devoted to gamma-ray astronomy in the energy range from  $2 \times 10^7$  eV to more than  $3 \times 10^{11}$  eV, over the large background of energetic charged particles at the 565 km altitude orbit of the Fermi satellite. For each gamma-ray, Fermi-LAT measures its arrival time, direction, and energy. The key improvements of Fermi-LAT from its predecessor, the EGRET experiment, have been obtained because of the newer technologies, principally in particle detection and in electronics. These improvements provide a larger effective area ( $\sim 8000$  cm<sup>2</sup> compared to  $\sim 1500$  cm<sup>2</sup> of EGRET) over a much larger field-of-view [97]. The Fermi-LAT Collaboration has released data on transient sources and light curves for more than 20 regularly monitored sources, and continuously adds more sources to the list as they show significant brightening. For example, the 2FGL catalog contains 1873 sources detected and characterized in the  $10^8$  eV -  $10^{11}$  eV range, some of which can be seen in Fig. 1.18, thanks to its unprecedented angular resolution. In general, all sources of gamma-rays at these energies are either locations of high-energy (nonthermal) processes or indicate locations of new physics in the Universe beyond the Standard Model of particle physics (see [98] for a review).

The IceCube Collaboration has recently reported evidence for a high-energy neutrino flux. During two years of operation 28 events with energies between  $3 \times 10^{13}$  eV and  $1.2 \times 10^{15}$  eV were observed while only 10.6 events were expected from conventional atmospheric background [100]. The hadronic interactions responsible for this IceCube excess would have also produced a flux of high-energy photons that can serve as a probe of source direction and distance. The existing limits for the diffuse photon flux at these energies (see Sec. 1.5.3) support the interpretation that the IceCube excess is mostly of extragalactic origin.

The measurement of the extragalactic isotropic diffuse gamma-ray background measured by Fermi-LAT reinforces the extragalactic origin hypothesis for the IceCube neutrinos. In Fig. 1.19, both spectra are shown after accounting for the cascading of the  $10^{15}$  eV photons in radiation backgrounds between source and observation, assuming that the secondary particles are by-products



**Figure 1.18:** The Fermi-LAT sky map in Galactic coordinates. Superimposed, the position of the principal SNRs in the gamma-ray sky [99].

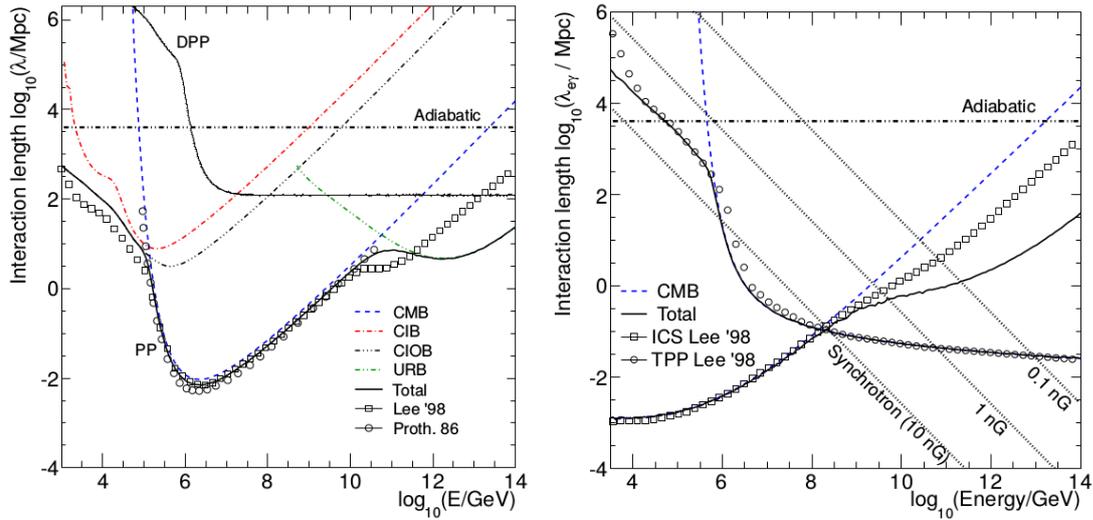


**Figure 1.19:** The astrophysical neutrino flux (black line) observed by IceCube matching the corresponding cascaded gamma-ray flux (red line) observed by Fermi-LAT [101]. The black data points are combined IceCube results [104]. Figure extracted from [103].

of proton-proton interactions. The black line in Fig. 1.19 shows an  $E^{-2.15}$  neutrino spectrum with an exponential cutoff around  $10^{15}$  eV. This scenario actually matches the extragalactic isotropic diffuse gamma-ray background measured by Fermi-LAT [101]. This indicates that the contribution of gamma-rays accompanying IceCube neutrinos to the extragalactic flux measured by Fermi-LAT is significant, suggesting a common origin at some level [102]. Although no definite identification of the sources of cosmic neutrinos has yet emerged, it is rather clear that a multiwavelength path to the neutrino sources looks very promising [103].

## 1.5.2 Propagation

The predictions of the ultra-high energy photon flux at Earth not only rely on theoretical models at the production sites, but depend as well on the interactions that photons undergo during their propagation. Although photons are not subjected to magnetic deflections, they interact with the extragalactic background photons  $\gamma_b$  inducing electromagnetic cascades. The main interaction



**Figure 1.20:** The average interaction length for photons (left) and PP electrons (right) in terms of their energy for interactions with different background radiation fields [107]. The dashed line gives the energy loss length for adiabatic energy losses due to the expansion of the Universe [108].

channel is through *Pair Production* (PP)<sup>15</sup> [105]:

$$\gamma_{\text{UHE}} + \gamma_b \rightarrow e^+ + e^- \quad (1.16)$$

The threshold energy  $E_{\text{thr}}$  for PP with a background photon of energy  $\varepsilon$  is:

$$E_{\text{thr}} = \frac{m_e^2}{\varepsilon} \simeq 2.6 \times 10^{11} \left( \frac{\varepsilon}{\text{eV}} \right)^{-1} \text{ eV} \quad (1.17)$$

where  $m_e$  denotes the electron mass [21]. Hence, background photons with an energy  $\varepsilon \lesssim 4 \times 10^{-7} \text{ eV}$  are the most efficient targets for primary photons with energies around  $10^{19} \text{ eV}$ . These photons belong to the URB, as seen in Fig. 1.14. Unfortunately, the URB is not very well known mostly because it is difficult to disentangle the Galactic and extragalactic components [48, 106].

The absorption of ultra-high energy photons by PP with the low energy photons of the extragalactic background radiation, specially the CMB, limits their observability to the local Universe. The effect is energy-dependent, with photons at energies around  $10^{16} \text{ eV}$  having a horizon distance of order of the Milky Way size, while photons at  $10^{18} \text{ eV}$  are detectable from distances of  $\sim 300 \text{ kpc}$ . Typical interaction lengths in terms of the photon energy are shown in Fig. 1.20, left. Photons at energies between  $10^{16} \text{ eV}$  and  $10^{17} \text{ eV}$  are thus created in the Galaxy unless some new physics is assumed.

The energy distribution between the electron and the positron produced in the PP process according to Eq. 1.16 is not symmetric due to the very high center-of-mass energy. One of the particles carries away  $\sim 90\%$  of the energy of the primary photon. This leading particle can then undergo *Inverse Compton Scattering* (ICS)<sup>16</sup> or *Triplet Pair Production* (TPP) with background photons:

<sup>15</sup>The Double Pair-Production, in which four electrons and positrons are generated, is also possible but at energies below  $10^{12} \text{ eV}$ .

<sup>16</sup>In contrast to other types of interactions, the ICS is a well known process occurring without threshold.

$$e^{\pm} + \gamma_b \rightarrow e^{\pm} + \gamma_{\text{UHE}} \quad (1.18)$$

$$e^{\pm} + \gamma_b \rightarrow e^{\pm} + e^{-} + e^{+} \quad (1.19)$$

In the ICS, most of the energy of the electron or positron is transferred to the upscattered background photon, which can now be considered a secondary ultra-high energy photon [21]. This process is dominant if the electron has an energy below  $\sim 10^{17}$  eV, as seen in Fig. 1.20, right. At higher energies, the most probable interaction process is the TPP or the synchrotron emission. The development of the electromagnetic cascade depends on the strength of the intervening extragalactic magnetic fields. For sufficiently intense field, electrons in the cascade lose most of their energy by synchrotron radiation and the cascade development stops when the synchrotron cooling time scale becomes smaller than the mean ICS interaction path [107]. The energy-loss lengths corresponding to the average synchrotron radiation emission for  $e^{\pm}$  [46], is also shown in Fig. 1.20, right, for three different intensities of the magnetic field. Finally, electron and photons adiabatically lose their energy because of the expansion of the Universe.

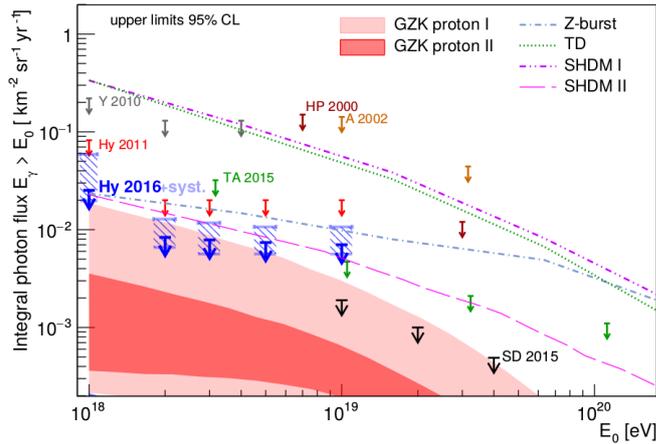
Through repeated cycles of PP and ICS<sup>17</sup>, an electromagnetic cascade develops and thus the presence of a weak intergalactic magnetic field can propagate photons over significant distances [109]. This cascade accelerates at lower energies due to the decreasing interaction lengths until most of the photons fall below the PP threshold (see Eq. 1.17). At this point they pile-up with a characteristic  $E^{-1.5}$  spectrum [110]. Most of the energy of fully developed electromagnetic cascades ends up below  $\sim 100$  GeV where it is constrained by measurements of the diffuse gamma-ray flux by the EGRET experiment [111].

### 1.5.3 Upper limits to the photon flux

Although it is theoretically possible that photons contribute to the flux of ultra-high energy cosmic rays, no direct observation has been confirmed above  $8.1 \times 10^{13}$  eV [112]. Instead, upper limits to the photon flux in higher energy domains have been established by several experiments [113]. The Pierre Auger Observatory has conducted several photon searches for energies above  $10^{18}$  eV by combining parameters related to the longitudinal development of the EAS and to the particles arriving to ground (see Chapter 2). No photon candidates have been univoquely identified and thus upper limits on the integral flux of ultra-high energy photons of 0.027, 0.009, 0.008, 0.008 and  $0.007 \text{ km}^{-2} \text{ sr}^{-1} \text{ yr}^{-1}$  at 95% confidence level have been established above 1, 2, 3, 5 and  $10 \times 10^{18}$  eV. These limits bound the fractions of photons in the all-particle integral flux below 0.1%, 0.15%, 0.33%, 0.85% and 2.7% (see [114] and references therein). The upper limits established by Auger are compared with results published by previous experiments, as well as predictions of different cosmic-ray propagation and top-down models, in Fig. 1.21.

The first evidence in favor of the presence of ultra-high energy photons has been found by re-analysing the data of the EAS-MSU experiment [121]. An estimated ultra-high energy photon flux of  $(1.55_{-0.67}^{+0.75}) \times 10^{-16} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$  above  $5.2 \times 10^{16}$  eV and  $(0.45_{-0.21}^{+0.36}) \times 10^{-16} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$  above  $9.1 \times 10^{16}$  eV was reported by searching for an excess in muonless events with a  $36.4 \text{ m}^2$  buried muon detector [122]. Although photons at these energies are expected to have a Galactic origin, as discussed before, no significant anisotropy in the arrival directions has been found. Moreover, the theoretical origin mechanism for photon primaries in this energy domain has not been yet clearly

<sup>17</sup>In the high energy limit, the total cross sections for PP and ICS are related through  $\sigma_{\text{PP}} \simeq 2\sigma_{\text{ICS}}$  [21].



**Figure 1.21:** The upper limits on the integral photon flux derived by the Pierre Auger Collaboration (blue arrows) for a photon flux  $E^{-2}$ . The limits obtained when the detector systematic uncertainties are taken into account are shown as horizontal segments (light blue) delimiting a dashed-filled box at each energy threshold. Previous limits are also reported: for Auger (SD 2015 [115] and Hy 2011 [116]), Telescope Array (TA [117]), Yakutsk (Y [118]) and Haverah Park (HP [119]). The shaded regions and the lines give the predictions for the GZK photon flux [82] and for top-down models (TD, Z-Burst, SHDM I [120] and SHDM II [49]).

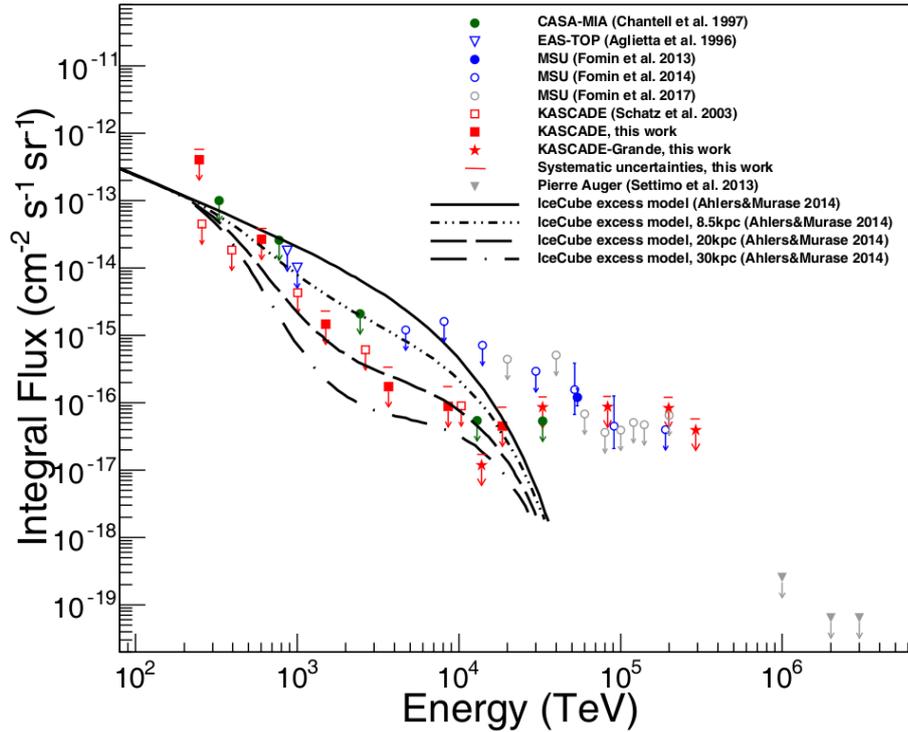
established. Nevertheless, a later re-analysis did not confirm this evidence [123]. The existence of primary photons in the cosmic-ray flux between  $10^{16.5}$  eV and  $10^{18}$  eV have only been studied by this experiment and KASCADE-Grande [124], as shown in Fig. 1.22.

#### 1.5.4 Motivations for the search at ultra-high energies

The search for ultra-high energy ( $E > 10^{14}$  eV) photons complements measurements of cosmic rays and neutrinos towards a multi-messenger understanding of the most energetic astrophysical phenomena. In particular, the discovery of photons with energies between  $10^{16}$  eV and  $10^{18}$  eV in the cosmic rays flux could be of particular interest not only for the field of Astroparticle Physics, but also for Astrophysics and fundamental Physics, since they are tracers of the highest energy processes in the Universe [127].

Therefore, the detection of a photon flux between  $10^{16.5}$  eV and  $10^{18}$  eV or the upper limit estimation following a non-observation of a clear flux would represent an important tool to answer many open questions. The detection of ultra-high energy photons in this energy domain would also open the possibility for exploring the Universe in a novel approach, complementing recent neutrino and gravitational waves observations. Following the discussion of the previous sections, a short summary of the motivations for the search of primary photons between  $10^{16.5}$  eV and  $10^{18}$  eV is presented in this section.

**Point sources identification** Photons point back to their astrophysical sources as they are not deflected by magnetic fields. Thus, they could reveal details about the cosmic-ray acceleration in their sources as well as deliver information about their propagation to Earth. This makes them suitable candidates to extend astronomical observations to unprecedented energy ranges. Moreover, if the sources of cosmic rays are transient, typically involving compact objects (such as GRBs), an observation of photons, neutrinos and gravitational waves could in fact be the only path that might lead to a full understanding of the underlying processes.



**Figure 1.22:** The upper limits to the integral photon flux published by several experiments compared with theoretical curves by an IceCube excess model [100], see Sec. 1.5.1. The lines represent the unattenuated flux (solid) and the flux from sources at 8.5 kpc, 20 kpc, and 30 kpc (dotted lines), respectively. The marker colors represent the limits published by different experiments: EAS-MSU in blue [121, 123, 122], KASCADE and KASCADE-Grande in red [125, 126] and Pierre Auger in gray [116], among others. Figure taken from [126].

**Mass composition of cosmic rays** The ultra-high energy cosmic rays are expected to produce pions by the GZK effect in the conventional models of propagation [13]. These pions in turn decay in secondary photons that would comprise less than 1% of the observed cosmic-ray flux depending on the mass of the primary cosmic ray: a light primary composition would produce a larger flux of secondary photons. Therefore, a hypothetical ultra-high energy photon flux could serve as a tool to describe the mass composition of cosmic rays near their sources. Since still today the cosmic rays composition is a controversial subject, the search for an ultra-high energy photon flux becomes relevant [128].

**Dark matter and new physics** The ultra-high energy photons coming directly from the sources are strongly suppressed between  $10^{15}$  eV to  $10^{17}$  eV due to pair production with the CMB. Typical interaction lengths are  $\sim 20$  kpc at  $10^{16.5}$  eV and  $\sim 300$  kpc at  $10^{18}$  eV [107]. Therefore, a direct photon flux from extragalactic sources is strongly suppressed, while favouring a Galactic origin. An observation of photons at these energies may be a smoking-gun signal of new physics, including superheavy dark matter [83, 21], axion-like particles [129], or an ultimate test of Lorentz-invariance [130]. On the other hand, a non-observation of a photon flux would allow for the estimation of an upper limit which imposes strong restrictions on the top-down models (see Sec. 1.2.2).

**Particle interactions** The identification of ultra-high energy photons with ground-based large-aperture detectors such as Pierre Auger can serve, through the interactions of the photons

with lower-energy photons or atmospheric nuclei, as probes of several aspects of Quantum Electrodynamics and Chromodynamics at energies currently unreachable by man-made collider experiments.

**Relation to IceCube neutrinos** The origin of the 28 neutrinos detected by IceCube is still uncertain. Although the preexisting photon limits at  $\sim 10^{15}$  eV support the extragalactic origin hypothesis, these photon surveys were conducted by Observatories located in the Northern Hemisphere whereas most of the 28 IceCube events are located in the Southern Hemisphere. This emphasizes the importance of a dedicated photon search around  $10^{15}$  eV in the Southern Hemisphere, particularly in the extended region around the Galactic center. Moreover, possible sub-dominant contributions from Galactic neutrino sources like SNRs are marginally consistent with present photon limits. Although there is no statistically significant neutrino event clustering at present [131], diffuse photon surveys around  $10^{15}$  eV are an important probe to break the degeneracy between Galactic and extragalactic neutrino contributions.

**Acceleration at Galactic center** The HESS Collaboration has recently reported strong evidence for  $10^{15} - 10^{17}$  eV protons coming from the Galactic center [31]. The supermassive black hole Sagittarius A\* was proposed as a feasible accelerator up to those energies. In this scenario, protons may interact with the surrounding matter and radiation fields, leading to the emission of photons of similar energies. Therefore, a photon signature at these energies may clarify the feasibility of Sagittarius A\* as a source of Galactic cosmic rays, alternatively to SNRs.

**The AMIGA Muon Detector** The Pierre Auger Observatory employs several independent detection techniques, which are thoroughly described in Chapter 3. The direct detection of the muon content of the air showers is provided by the AMIGA muon detector, which represents the base design for the upcoming *Underground Muon Detector* of the *AugerPrime* upgrade (see Sec 3.5). This muon detection system grants a unique approach to detect air showers with a sparse presence of muon content, which could be explained by a primary photon origin.

The unclear photon detection evidence reported by EAS-MSU and the lack of a complete picture around  $\sim 10^{17}$  eV are the main motivations for this thesis. The research described throughout the next chapters extends the Auger photon search to  $10^{16.5}$  eV, more than one order of magnitude below the current photon studies [114]. The information from one of the surface detector arrays of Auger and its dedicated muon detection system are the cornerstones to accomplish this goal.

## Chapter 2

# Physics of extensive air showers

We can judge our progress by the courage of our questions and the depth of our answers, our willingness to embrace what is true rather than what feels good.

---

Carl Sagan

Current balloon- and satellite-borne experiments are limited to an effective detection area of a few  $\text{m}^2$ . The extremely low cosmic rays flux in the energy region above the knee prevents to collect significant statistics with these types of experiments on reasonable time scales. Fortunately, cosmic rays entering the atmosphere interact with the atmospheric nuclei and produce cascades of secondary particles which emit Cherenkov and fluorescence light. These particle cascades are called *Extensive Air Showers*. Both its development in the air and the characteristics of the particles that reach the ground can be analyzed in order to indirectly infer the mass composition and arrival direction of the primary cosmic ray.

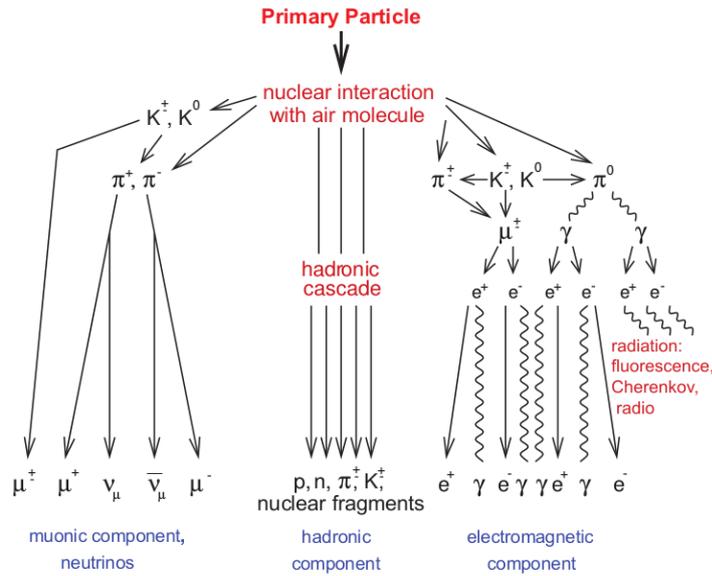
### 2.1 Components of the EAS

The *first interaction* with an air nucleus takes place at altitudes between 15 and 35 km, depending on the mass of the particle and angle of incidence [132]. The secondary particles of the first interaction are mainly pions, kaons, and nucleons<sup>1</sup>. In turn, each secondary interact again in the atmosphere, thus generating more particles. The particle multiplication continues either by interaction or decay processes.

The particles produced in an EAS can be classified in three components. The *hadronic component* is composed by secondary baryons, pions and kaons. The hadronic component feeds the other two components. High-energy photons from the decay of neutral pions are the dominant source for the *electromagnetic component*. On the other hand, the decay of charged pions and kaons gives rise to the *muonic component*. In addition, up to 10% of the low-energy muons are produced by the electromagnetic shower component through photoproduction or muon pair-production.

---

<sup>1</sup>In a typical interaction, more than 60% of the produced hadrons are pions (neutral and charged), followed by 10% kaons. The remaining fraction corresponds to neutrons and protons.



**Figure 2.1:** The schematic representation of the three components produced in an EAS [133].

Conversely, muon interaction and decay lead again to electromagnetic particles [132]. The relative weight of each channel depends intimately on the mass composition of the primary cosmic ray. A simplified scheme of the cascade of secondary particles is shown in Fig. 2.1.

**Hadronic component** The long-lived secondary hadrons form the hadronic shower core. The produced hadrons rapidly lose their energy by interacting with air nuclei, mainly by nuclear fragmentation, and thus relatively few of them arrive to ground. Secondary hadrons are produced at a typical, almost energy-independent transverse momentum of  $p_\perp \sim 350 - 400$  MeV, leading to a large angular dispersion of low-energy hadrons relative to the shower axis. Thus, hadrons exhibit a wider lateral distribution than electromagnetic particles do.

**Electromagnetic component** The neutral pions from the hadronic component decay almost immediately into two photons<sup>2</sup> since they have a very short decay length ( $c\tau = 25$  nm) [135]. The electromagnetic (EM) component represents 99.9% of the particles arriving to ground, out of which only  $\sim 10\%$  correspond to electrons and positrons.

For the electromagnetic components in the shower, there are mainly three processes for energy loss, namely pair-production of photons, and bremsstrahlung and ionization losses of electrons and positrons. The cross-section for the pair-production is  $\sim 60$  mb, which is  $7/9$  of the cross-section for bremsstrahlung. The electrons and positrons lose energy by ionization at the rate of  $\sim 2$  MeV  $g^{-1}$   $cm^2$  while at the same time radiating photons through bremsstrahlung. Thus, particle multiplication and ionization energy loss are competing processes in showers.

The EM component of the shower grows almost exponentially as it propagates through the atmosphere, so the primary energy is divided between an increasing number of particles until the mean energy per EM particle approaches the critical value of  $E_c \sim 86$  MeV. At this point, the ionization mechanism, which does not produce new particles, starts to dominate over the

<sup>2</sup>There are other possible decays for the neutral pion like  $\pi^0 \rightarrow \gamma + e^- + e^+$  and  $\pi^0 \rightarrow e^- + e^+ + e^- + e^+$  but they are negligible [134].

bremsstrahlung and pair-production mechanisms. In this way, the number of EM particles of the shower starts to decrease as it propagates to the ground.

The energy dissipated by the EM channel through air ionization is proportional to the number of low-energy particles. In consequence, the energy deposit in the atmosphere is proportional to the primary energy<sup>3</sup>. Therefore, the measurement of the energy deposit with fluorescence telescopes<sup>4</sup> is the most reliable way to estimate the primary energy (see Sec. 3.2).

**Muonic component** The charged pions have a longer decay length of  $c\tau \sim 8$  m so that they interact multiple times producing more secondaries before reaching a critical energy of  $E_c^\pi \sim 20$  GeV. At this point, the most probable process is the decay into muons and neutrinos. The muon inherits about 80% of the pion energy, thus its typical energy at the production point is of the same order as the critical energy  $E_c^\pi$ . Since the probability of interaction is inversely proportional to the medium density, most of the muons are generated in the upper layers of the atmosphere [136]. A similar scenario happens for charged kaons, albeit with a slightly shorter lifetime ( $c\tau = 3.7$  m). Consequently, they decay at higher energies.

The pions still have large Lorentz factors at the end of the hadronic cascade, which are inherited by the muons. Therefore, the muons form a collimated beam in the forward direction. Contrarily to hadrons, they do not lose a significant amount of energy by ionization during their propagation ( $dE/dX \simeq 3 \text{ MeV g}^{-1} \text{ cm}^2$ ). Furthermore, the energy losses due to bremsstrahlung and pair-production are subdominant<sup>5</sup> [137]. Having a small effective cross-section and a decay length of  $c\tau \sim 450$  m, they can easily reach the surface. The muonic component decays very slowly through bremsstrahlung and pair-production, feeding the EM component.

## 2.2 Topology of an EAS

The general structure of an EAS is shown in Fig. 2.2. The *shower axis* is defined as the extension of the initial momentum vector of the incident primary in the direction of cascade propagation. Experimentally its interception with the plane of observation is reconstructed from the measured lateral density distribution of the shower particles, as explained in Chapter 3. Its direction of incidence, i.e. its *zenith* and *azimuthal* angles, can be determined from measurements of the arrival time of the particles on the plane of observation. The *shower front* has the form of a disk that shows a slight curvature depending on the primary energy and direction of incidence. The bulk of the particles arrives in a narrow time interval, ranging from only a few nanoseconds in the vicinity of the shower axis to some  $\sim 10$  ns at larger distances from the *shower core*<sup>6</sup>. The particle disk broadens slightly with increasing radial distance from the shower axis because of larger path length fluctuations due to increased scattering at lower energies, and because of lower Lorentz

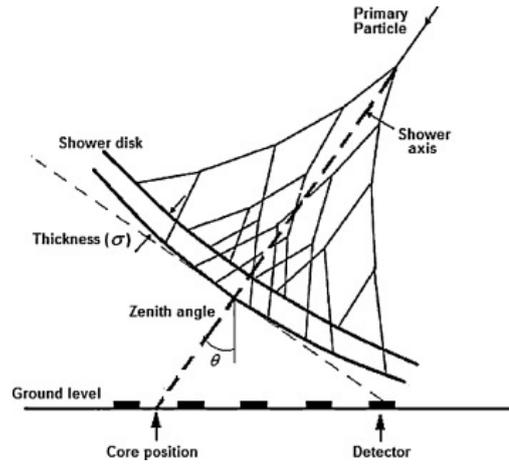


Figure 2.2: The schematic representation of an EAS [138].

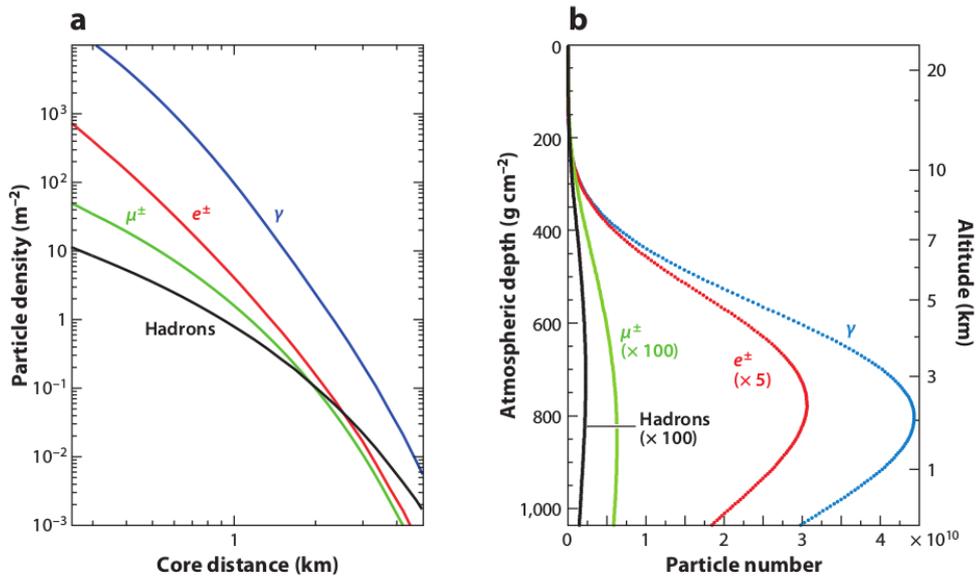
factors of the parent nucleons, responsible for the local sub-cascades.

The features of the different components of the EAS change with the altitude. As a characteristic unit of the amount of matter traversed by a particle, the atmospheric depth  $X$  from the top of the atmosphere downwards is preferred since it allows to measure distances independently of the density of the medium  $\rho$  [108]. Considering an impinging particle with a zenith angle  $\theta$ , the atmospheric depth,  $X$ , can be calculated as:

$$X(h, \theta) = \frac{\int_h^\infty \rho(h') dh'}{\cos\theta} \quad (2.1)$$

Fig. 2.3 shows the lateral (i.e., transverse to the shower axis) and longitudinal particle profiles of the different shower components, simulated with the CORSIKA<sup>7</sup> package [139], for proton-induced showers of  $10^{19}$  eV. The lateral distribution of muons is flatter than that of EM particles because at large lateral distances the latter component is more attenuated than the former one. Consequently, the muonic component dominates over the EM component at sufficiently large distances from the shower axis. Moreover, in showers with very large zenith angles ( $\theta > 65^\circ$ ), the muonic shower component and the EM particles produced in the decay of muons are the only particles that can be detected at ground. At these distances, muons are mainly produced in the decay of low-energy pions [140].

Since the primary particle may interact with an air nucleus at a non-fixed atmospheric depth, identical primary particles having the same energy and zenith angle have different lateral and longitudinal profiles<sup>8</sup>. These intrinsic variations of the shower development are jointly called *shower-to-shower fluctuations*. In Fig. 2.4, the fluctuations on the longitudinal profile are illustrated for different primary particles.



**Figure 2.3:** The average (left) lateral and (right) longitudinal shower profiles for vertical, proton-induced showers at  $10^{19}$  eV. The lateral distribution of the particles at ground is calculated for  $870 \text{ g cm}^{-2}$ , the depth of the Pierre Auger Observatory [135].

## 2.3 The Heitler model

In the early years of cosmic-ray physics, shower properties were calculated solving cascade equations. Nowadays, it is common to simulate air showers in much more detail with complete Monte Carlo (MC) simulation packages. But there is a simple model developed by Walter H. Heitler in 1950s that predicts the main characteristics of an EAS created by an electromagnetic primary [142].

The *Heitler model* assumes that photons create electron-positron which in turn produce photons through the bremsstrahlung process after traversing a fixed distance. In addition to radiative energy losses, electrons are subjected to ionization energy loss. The total energy loss  $dE/dX$  of electrons can be written as:

$$\frac{dE}{dX} = -\frac{E}{X_R} - \alpha(E) \quad (2.2)$$

where  $\alpha(E)$  is the ionization energy loss given by the Bethe-Bloch formula [75] which depends logarithmically on energy and the linear term accounts for the bremsstrahlung radiative energy loss. The radiation length is  $X_R \sim 37 \text{ g cm}^{-2}$  in air, which defines the exponential loss length of the energy by bremsstrahlung. Therefore the energy loss due to bremsstrahlung is dominant at

<sup>3</sup>Around 80% at  $10^{15}$  eV growing to 90% at  $10^{20}$  eV of the primary energy is transferred to the EM sector by the ionization process.

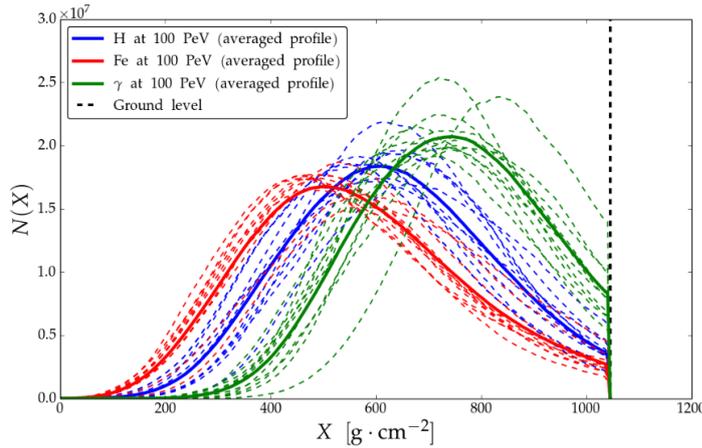
<sup>4</sup>The remaining fraction of the primary energy is transferred to the muonic component through charged pion production. Since this fraction is not measured by the fluorescence technique, it is called *invisible energy*. The primary energy estimation contains a correction accounting for this effect (refer to Sec. 3.2.4).

<sup>5</sup>In particular, these effects start to be dominant in the muon energy losses at energies of the order of  $10^3$  GeV.

<sup>6</sup>The tail of the arrival time distribution which contains almost exclusively low-energy particles, can extend beyond  $1 \mu\text{s}$ .

<sup>7</sup>The CORSIKA software is described in Sec. 4.2.1.

<sup>8</sup>These fluctuations are also due to the several stochastic processes that compete in the development of air showers, as a consequence of the low density nature of the atmosphere.



**Figure 2.4:** An illustrative example of shower-to-shower fluctuations in the longitudinal profiles of EAS generated by photons (green), protons (blue), and iron nuclei (red), with the same primary energy of  $10^{17}$  eV. Each dashed line corresponds to a simulated air-shower longitudinal profile, whereas the solid lines represent the average longitudinal profile for each primary [141].

the highest energies.

The cascade development is depicted in Fig. 2.5, left, where each segment can be thought of as a photon or an electron and each vertex as the interaction point after traversing an atmospheric slant depth  $\lambda_e$ . After the shower front has traversed an atmospheric depth  $X$ , the primary energy  $E_0$  is divided equally among all the particles  $N$ . Therefore, the energy per particle after  $n$  generations, with  $n = X/\lambda_e$ , is:

$$E(X) = \frac{E_0}{N(X)} = \frac{E_0}{2^{X/\lambda_e}} \quad (2.3)$$

The particle-multiplication process continues until ionization-energy losses dominate over radiative losses<sup>9</sup>. The number of particles in the shower reaches the maximum  $N_{\max}$  at  $E = E_c$ . Finally, two important relations are predicted: the number of particles at the shower maximum is proportional to  $E_0$ , as in Eq. 2.4, and the depth of the shower maximum depends logarithmically on the primary energy  $E_0$ , as in Eq. 2.5.

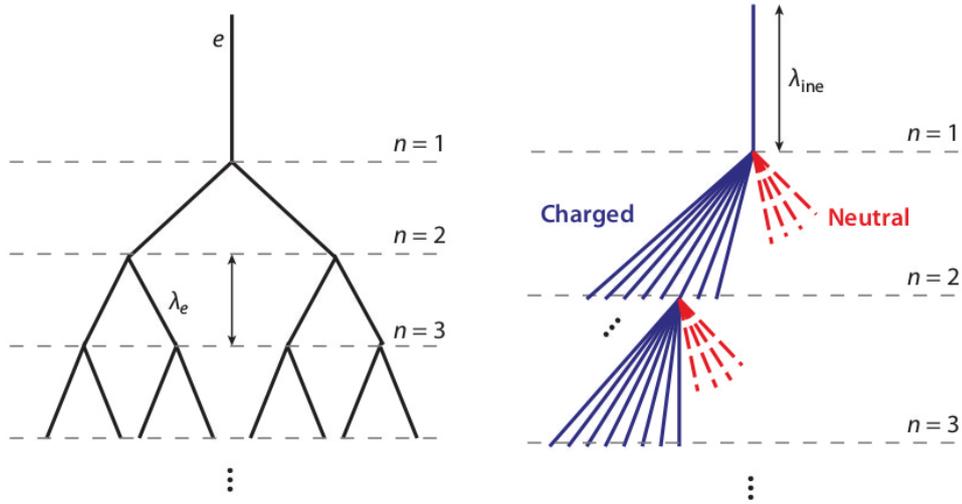
$$N_{\max} = N(X_{\max}) = \frac{E_0}{E_c} \quad (2.4)$$

$$X_{\max}^{(\text{EM})}(E_0) \sim \lambda_e \ln \left( \frac{E_0}{E_c} \right) \quad (2.5)$$

Although the model is very simple, these predictions are in qualitative agreement with much more detailed MC simulations of EAS [108]. The model overestimates the number of electrons  $N_e = \frac{2}{3}N_{\max}$  at the shower maximum since several photons can be produced by bremsstrahlung and not just one as the model assumes. Moreover, some electrons may leave the shower to integrate the medium. Numerically, it is observed that the number of photons exceeds by approximately an order of magnitude the number of electrons [143].

The longitudinal shower profile of an electromagnetic EAS can be calculated from cascade theory, and a related parametrization due to Thomas K. Gaisser and Anthony M. Hillas [144] is often used to fit measured shower profiles:

<sup>9</sup>After the maximum is reached, the particles can decay, or be absorbed, thus reducing the number of particles, but these processes are not in the scope of the simplified Heitler model.



**Figure 2.5:** The schematic representation of the Heitler model describing the electromagnetic component (left) and the hadronic component (right) of an EAS. Each segment represents a particle and each vertex represents a process by which several particles are generated. On the right panel, only one charged hadron interaction is shown for each generation [135].

$$N(x) = N_{\max} \left( \frac{X - X_0}{X_{\max} - X_0} \right)^{(X_{\max} - X)/v} \exp \left( \frac{X_{\max} - X}{v} \right) \quad (2.6)$$

where  $X_0$  and  $v$  are shape parameters. The dependence of the particle density  $\rho_e$  on the distance to the shower core  $r$ , i.e. the lateral distribution, is determined mainly by the multiple Coulomb scattering of electrons. Detailed calculations of the lateral shower profile by Jun Nishimura and Koichi Kamata [145] were parametrized by Kenneth Greisen [146] in the so-called *NKG function*:

$$\rho_e(r) = C(s) N_e(X) \left( \frac{r}{r_1} \right)^{s-2} \left( 1 + \frac{r}{r_1} \right)^{s-4.5} \quad (2.7)$$

where  $s$  is the shower parameter, often defined as  $s \simeq 3X/(X + 2X_{\max})$ ,  $C(s)$  a normalization constant and  $r_1$  the Molière radius which depends on  $E_c$  and  $\lambda_e$ .

### 2.3.1 Extension to hadronic showers

Although the simplified model describes only the EM component of the shower, some insight into the features of hadronic showers can be gained by generalizing the Heitler model, as done by James Matthews [147]. In this extended version,  $n_{\text{ch}}$  charged pions together with  $\frac{1}{2} \cdot n_{\text{ch}}$  neutral pions are produced in each interaction, as depicted in Fig. 2.5, right. Neutral particles decay immediately into EM particles ( $\pi^0 \rightarrow 2\gamma$ ). After having traveled a distance corresponding to the mean interaction length  $\lambda$ , charged particles interact again with air nuclei if their energy is greater than some typical decay energy  $E_c^\pi$ . Once the energy of the charged hadrons falls below this threshold, they decay and produce one muon each.

In each hadronic interaction, one-third of the energy is transferred via  $\pi^0$  decay to the EM shower component. After  $n$  generations, the energies in the hadronic and EM components are given by, respectively:

$$E_{\text{had}} = \left(\frac{2}{3}\right)^n E_0 \quad (2.8)$$

$$E_{\text{EM}} = \left(1 - \left(\frac{2}{3}\right)^n\right) E_0 \quad (2.9)$$

With  $n \sim 6$ , approximately 90% of the initial shower energy is finally carried by EM particles and deposited as ionization energy in the atmosphere, as mentioned before.

The number of muons in an EAS can be estimated considering that the number of charged pions after  $n$  generations is  $N_{\text{ch}} = n_{\text{ch}}^n$ , while the energy per charged pion is:

$$E^\pi = \frac{E_0}{\left(\frac{3}{2} \cdot n_{\text{ch}}\right)^n} \quad (2.10)$$

where also the number of neutral pions has to be taken into account. The particle cascade stops, when the critical energy  $E_c^\pi$  is reached after  $n_c$  generations:

$$n_c = \frac{\ln\left(\frac{E_0}{E_c^\pi}\right)}{\ln\left(\frac{3}{2} \cdot n_{\text{ch}}\right)} \quad (2.11)$$

Assuming that all charged pions at this stage of the cascade decay into muons, the total number of muons  $N_\mu$  can be calculated as:

$$N_\mu = n_{\text{ch}}^{n_c} = \left(\frac{E_0}{E_c^\pi}\right)^\beta \quad (2.12)$$

where  $\beta \simeq 0.82 \dots 0.94$  [132]. Therefore, the number of muons increases with energy almost linearly and depends on the air density, through  $E_c^\pi$ , and the charged pions multiplicity.

While the predictions of the extended Heitler model are in general in good agreement with detailed MC simulations, it should be noted that it neglects some fundamental aspects of the hadronic interactions relevant in the development of a hadronic cascade, for example the inelasticity  $\kappa$  of the interactions [147]. Furthermore, the model assumes that the energy of the primary particle is distributed equally between all secondary particles, although usually a fraction  $1 - \kappa$  of the energy is transferred to one leading particle.

### 2.3.2 Superposition model

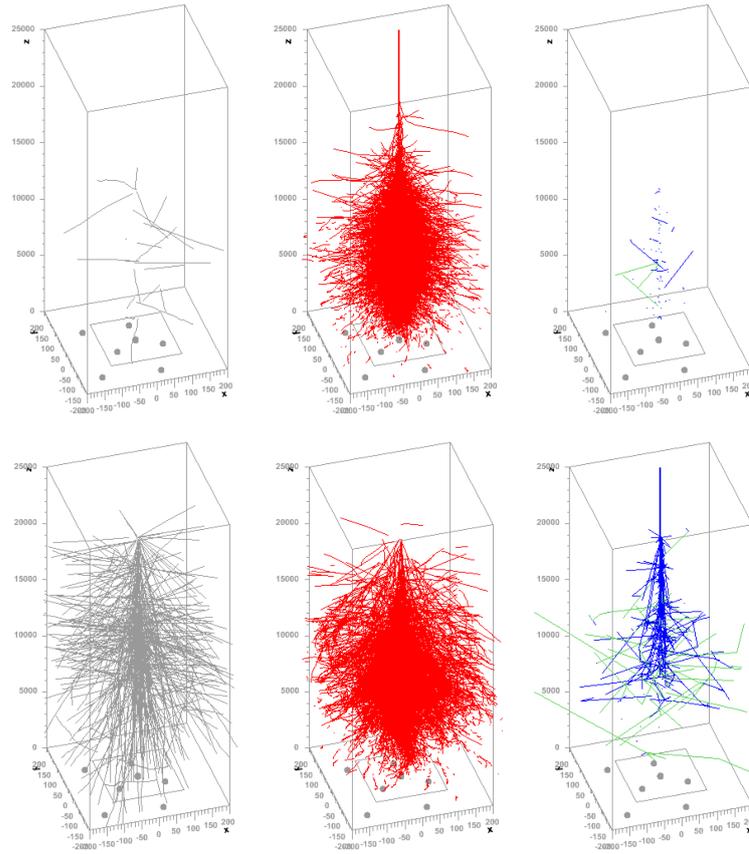
Because the binding energy of  $\sim 5$  MeV per nucleon is much smaller than the typical interaction energies, a nucleus of mass  $A$  can be considered as  $A$  independent nucleons. In this superposition model, a nucleus with mass  $A$  and energy  $E_0$  is considered as  $A$  independent nucleons with energy  $E_0/A$ . This leads to the predictions:

$$N_{\text{EM,max}}^{(A)}(E_0) = A \cdot N_{\text{EM,max}}^{(P)}\left(\frac{E_0}{A}\right) \approx N_{\text{EM,max}}^{(P)}(E_0) \quad (2.13)$$

$$N_\mu^{(A)} = A \cdot \left(\frac{E_0/A}{E_c^\pi}\right)^\beta = A^{1-\beta} \cdot N_\mu^{(P)}(E_0) \quad (2.14)$$

$$X_{\text{max}}^{(A)}(E_0) = X_{\text{max}}^{(P)}\left(\frac{E_0}{A}\right) \quad (2.15)$$

where the labels (p) and (A) denote proton- and nucleus-induced showers, respectively. The



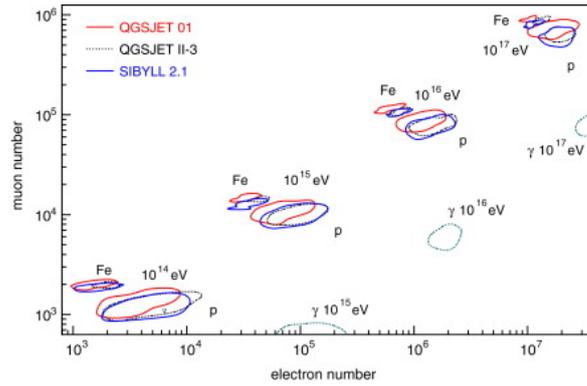
**Figure 2.6:** The tracks of secondary particles of EAS induced by a photon (top) and proton (bottom) primaries. From left to right, the longitudinal development of muons, EM particles and hadrons are shown. The distances are measured in meters. The primary energy in each case is  $10^{13}$  eV [132].

number of muons and the depth of maximum depend on the mass of the primary particle. The heavier the shower-initiating particle is, the more muons are expected for a given primary energy and the shallower the depth of maximum. Iron showers contain approximately 40% more muons than proton showers of the same energy, and they reach their maximum  $\sim 90 \text{ g cm}^{-2}$  higher in the atmosphere.

## 2.4 Properties of EAS produced by primary photons

The basic features of showers generated by primary photons are fairly well described by the simplified Heitler model, since their development is governed by the bremsstrahlung and pair production mechanisms. Therefore, the photon-initiated showers almost lack of hadronic and muonic components, as opposed to the proton- and nucleus-initiated EAS. The tracks of the secondary particles from the muonic, EM and hadronic shower components are shown for photon-initiated and a proton-initiated EAS in Fig. 2.6.

Although the muon pair-production is possible in a photon-initiated EAS, the process is suppressed by a factor  $m_e^2/m_\mu^2 = 2.3 \times 10^{-5}$  with respect to the electron pair-production. Similarly, the photonuclear interactions of high-energy photons with nuclei from the medium, responsible for the generation of secondary charged pions and kaons, are suppressed by almost  $\sim 10^{-2}$  with



**Figure 2.7:** The expected number of muons and electrons in vertical showers at sea level. The curves show the full-width half-maximum of the distributions for different primary particles and energies, as obtained with the quoted hadronic interaction models [63].

respect to pair-production in the Coulomb fields of the nuclei [148]. In this way, the scarce muonic content is one of the most notorious signatures of the EAS initiated by photons.

As muons are mainly produced in hadronic interactions, their number and lateral distribution can be used as composition-sensitive observables. The predicted muon distributions depend on the assumptions on hadron production in air showers. On the contrary, since the air showers induced by photons are almost purely electromagnetic, the choice of a specific hadronic-interaction model has a negligible impact<sup>10</sup>. This is demonstrated in Fig. 2.7 where the expected number of muons and electrons is shown for showers initiated by proton, iron, and photon primaries, as calculated with different interaction models.

The pure electromagnetic behaviour of photon showers also alters their longitudinal development. It is expected that showers initiated by primary photons develop, on average, deeper in the atmosphere. This is due to the small multiplicity of the EM interactions, in contrast to the large number of secondaries produced in inelastic interactions of high-energy hadrons [149]. The average atmospheric depth of the shower maximum for EAS induced by photon, proton and iron primaries as a function of the primary energy is shown in Fig. 2.8. Different hadronic interaction models, used in CORSIKA, are shown as well<sup>11</sup>. The difference in the average  $X_{\max}$  between photon and proton primaries is about  $80 \text{ g cm}^{-2}$  at  $10^{17} \text{ eV}$  and increases with the primary energy.

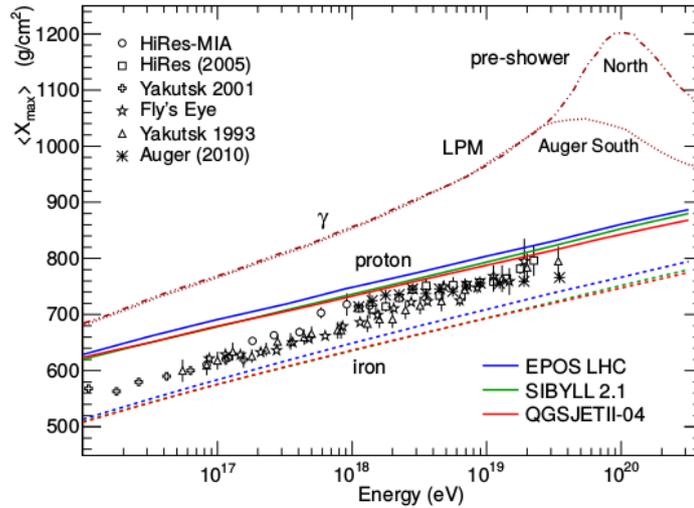
The delayed longitudinal development leaves also a clear signature on ground. Since less atmospheric mass is traversed by a photon-initiated compared to hadronic-initiated shower, the particles arriving at ground are more focussed near the shower core. In addition, the hadronic EAS have a prominent muonic component arriving at ground, which dominates over the EM component at large distances from the shower core, as mentioned in Sec. 2.2. Since the muonic component is almost absent in photon-initiated EAS, the lateral profile of particles on ground is expected to be steeper in the case of photon primaries than in their hadronic counterparts.

### 2.4.1 LPM and Preshower effects

As it is already seen in Fig. 2.8, there are certain phenomena affecting the longitudinal development that are unique to EAS initiated by photons.

<sup>10</sup>For example, the differences in the average  $X_{\max}$  between different hadronic-interaction models are usually less than  $5 \text{ g cm}^{-2}$  for such showers [127].

<sup>11</sup>The CORSIKA software is described in Sec. 4.2.1.



**Figure 2.8:** The average atmospheric depth of the shower maximum in terms of the primary energy for photon (brown dotted lines), proton (solid lines) and iron (dotted lines) primaries. Three hadronic interaction models used by the simulation software are shown. Measurements are represented with different marker styles [149].

Above  $\sim 10^{18}$  eV, subsequent interactions of photons or electrons with air can no longer be considered as independent and the scattering amplitudes have to be added coherently. This is the *Landau-Pomeranchuk-Migdal* (LPM) effect, which farther delays the shower development, leading to an even larger average  $X_{\max}$ . This effect reduces the cross-section for pair-production and bremsstrahlung at high energies or high matter densities [150].

Above  $\sim 3 \times 10^{19}$  eV, an ultra-high energy photon may convert into an electron-positron pair in the geomagnetic field above the atmosphere, initiating a *preshower* cascade [151]. When the preshower enters the atmosphere, a multitude of electromagnetic air showers is initiated, where the individual primary particles have a lower energy than the initial photon, thus reducing the average  $X_{\max}$ . Due to the superposition of many showers with lower energy, shower-to-shower fluctuations of *converted* primary photons are significantly reduced.

In summary, the high-energy muons that reach the surface carry valuable information regarding the shower development. On one hand, the study of the muon production profile could serve as a test bench for the hadronic interaction models which are employed in the EAS simulation. Indeed, since muons do not suffer greatly from energy losses, they provide a suitable imaging of the hadronic cascade that led to their production. On the other hand, the muon content, characterized by the muon density directly measured with the AMIGA muon detector, is sensitive to the mass composition of the primary particle, as deduced from Eq. 2.14, and particularly, offers a unique tool to discriminate between photon and hadronic primaries.



## Chapter 3

# The Pierre Auger Observatory

It is difficult to say what is impossible, for the dream of yesterday is the hope of today and the reality of tomorrow.

---

Robert H. Goddard

The discovery of the CMB in 1965 led to the prediction of the suppression of the energy spectrum of the cosmic rays above  $5 \times 10^{19}$  eV, provided that the mass composition at the sources is essentially light. This feature is known as the GZK cut-off (see Sec. 1.4). In 1990, there were already a few dozens of events with estimated energies above the GZK cut-off. Their interpretation was not clear at all: on one hand, the spectrum measured by the AGASA Collaboration did not present a suppression [44], but at the same time the HiRes Collaboration claimed the opposite [152]. The rather small sample of events was not enough to unveil the source of this discrepancy. Moreover, the different detection techniques (hence, different systematic uncertainties acting in the measurements) employed in each Observatory may had produced a wide variety of biases in an event-by-event basis.

A series of workshops, starting in Paris in 1992 and culminating in a 6-month study at Fermi National Accelerator Laboratory in 1995, gave birth to the design of the Pierre Auger Observatory (Auger) that would give a definite answer to the matter of the high energy end of the cosmic rays spectrum, make an effort to individualize sources, study other primaries rather than nuclei (photons and neutrinos), delucidate chemical composition of primaries, and study hadronic models at the highest energies. The design was sustained over the *hybrid detection* idea: a large-scale Observatory that could combine the surface detection technique employed by Haverah Park and the fluorescence technique employed by HiRes.

### 3.1 Site location and base design

Several aspects were taken into account in the choice of the Auger location, both related to the scientific goals and the construction feasibility [153]. These included the need for a location at  $\sim 1400$  m above sea level to optimize the detection of cosmic rays at  $10^{19}$  eV. In addition, the communications and deployment requirements made a relatively flat site with scarce vegetation

desirable. The need to detect the faint fluorescence signals produced by the EAS required a location with optical characteristics close to those sought by astronomical telescopes. The pre-existing infrastructure and human resources had to guarantee a smooth construction, operation and maintenance of the Observatory for at least 20 years but without creating anthropogenic pollution.

A broad sky coverage at the chosen location was one of the important considerations in the selection of water Cherenkov detectors (WCDs) for the surface detector array. WCDs were relatively deep (e.g. 1.2 m at Haverah Park), while scintillators commonly used in arrays were much thinner (9 cm at Volcano Ranch). As such, a WCD array would have approximately twice as much sky coverage as would have a scintillator array. According to simulations, a nearly uniform exposure of the sky could be accomplished at a latitude close to  $35^\circ$  [154]. After evaluating several sites in Argentina, Australia and South Africa, the Pampa Amarilla site ( $35.1^\circ - 35.5^\circ$  S,  $69.0^\circ - 69.6^\circ$  W and 1300 – 1400 m above sea level) in the south of the Province of Mendoza, Argentina, close to the town of Malargüe, was selected [155].

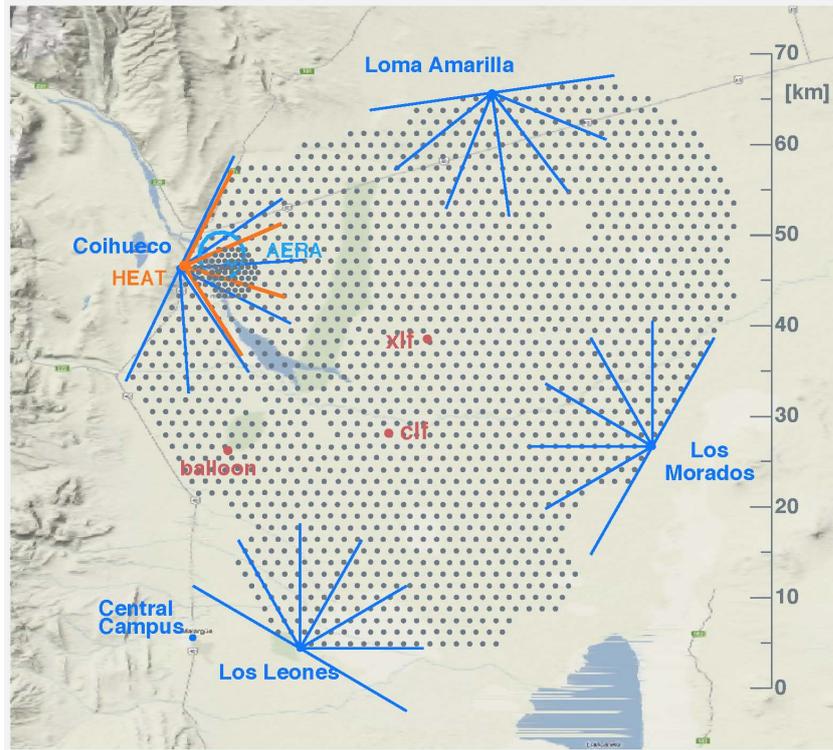
Above  $10^{20}$  eV, the rate of events is about  $1 \text{ km}^{-2} \text{ sr}^{-1} \text{ century}^{-1}$  so that vast areas must be monitored to collect a large statistical sample. The Auger Observatory has been planned as a pair of arrays, each of  $3000 \text{ km}^2$ , in both hemispheres though only the Southern hemisphere Observatory was finally built. The base design comprises more than 1600 WCDs arranged on a triangular grid, with the sides of the triangles being 1.5 km, overlooked from four sites by optical stations, each containing six telescopes. The WCDs respond to the particles of the EAS (mainly muons and electromagnetic particles) and the fluorescence telescopes measure the emission from atmospheric nitrogen, which is excited by the charged particles of the shower as they traverse the atmosphere. Both techniques were brought together, for the first time, in a *hybrid* detector to observe showers simultaneously with complementary detection techniques. The array of WCDs is known as the surface detector (SD) and its base design is often called “standard array” or SD-1500. The optical stations form the fluorescence detector (FD).

The combination of information from the two detection systems enhances the reconstruction capability with respect to the individual detector reconstruction [156]. A *hybrid event* is an EAS that is simultaneously detected by the FD and the SD. If an EAS independently triggers both detectors the event is tagged as a *golden hybrid* and these events can be fully reconstructed in both detection modes.

The construction of Auger began in 2001 with an Engineering Array (EA) consisting of two fluorescence telescopes and 40 WCDs deployed within the fields of view of the two telescopes. The construction and operation of the EA allowed to optimize the techniques related to the production and deployment of the detector components, the trigger algorithms, the data acquisition, monitoring software and telecommunications for both detection systems [157]. In 2004, the project entered into the production phase and in 2008 the construction of the base design was concluded.

Subsequently, two significant enhancements have been incorporated into the baseline detectors that significantly extended the Observatory science capability.

The *Auger Muons and Infill for the Ground Array* (AMIGA) enhancement, which is one of the cornerstones of this thesis, constitutes a hybrid upgrade to the Auger Observatory. On one hand, it included the deployment of an additional SD array of 61 WCDs separated by 750 m (SD-750) to lower the energy threshold of the SD-1500 to  $\sim 10^{17.5}$  eV. The original proposal also encompassed the installation of a denser triangular grid of WCDs separated by 433 m (SD-433) among several positions of the SD-750. On the other hand, the AMIGA upgrade includes the deployment of a Muon Detector (MD) associated with the SD-750. Currently, an EA of MD stations is coupled to seven WCDs from the SD-750. Both aspects of the AMIGA enhancement are extensively described



**Figure 3.1:** The layout of the Pierre Auger Observatory, where more than 1600 WCDs were deployed within the dotted area. The four FD sites are located at Los Leones, Coihueco, Los Morados and Loma Amarilla. The blue lines encompass  $30^\circ$  angles and define the azimuth acceptance of each fluorescence telescope [158].

in Sec. 3.3.5, 3.3.6 and 3.4.

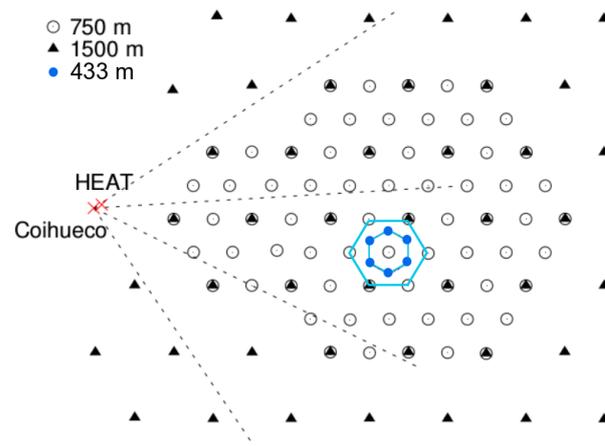
The *High Elevation Auger Telescopes* (HEAT) consisted in the installation of three additional telescopes at the Coihueco FD site, near the SD-750. The telescopes are tilted upwards compared to the FD base design so that the faint fluorescence light from low-energy EAS can be measured. In analogy to the SD-750, HEAT allows to lower the energy threshold of the FD base design. This enhancement is described in Sec. 3.5.1. The base design and the aforementioned enhancements are shown in Fig. 3.1. A more detailed view of the SD-750 is shown in Fig. 3.2.

## 3.2 The Fluorescence Detector

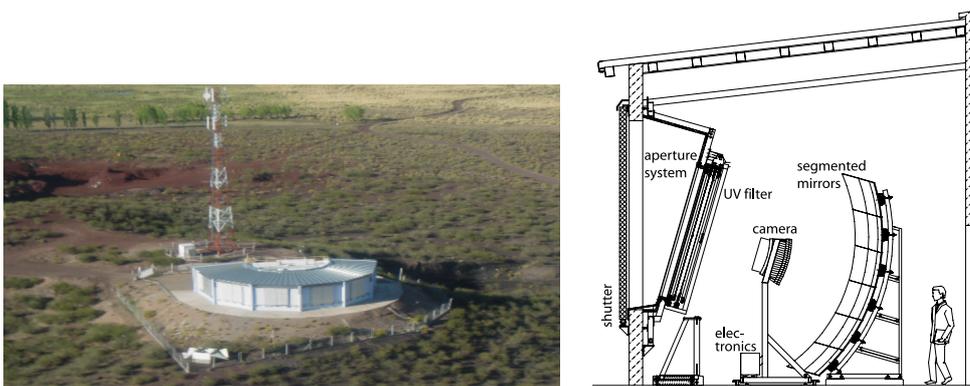
The secondary electromagnetic particles in the EAS excite nitrogen molecules in the atmosphere. A fraction of the energy deposited in the air is then re-emitted during the de-excitation of the nitrogen molecules. Typically, for each MeV of energy deposited, five photons are emitted in the wavelength range between 300 and 400 nm in dry air [160]. These fluorescence photons can be measured with the FD telescopes to estimate various parameters of the EAS, especially the energy of the primary particle.

### 3.2.1 Telescope hardware

The base design of the FD consists of 24 telescopes at four sites surrounding the SD: Los Leones, Los Morados, Loma Amarilla and Coihueco. At each site, there is a FD building housing



**Figure 3.2:** The schematic overview of the SD-750 array. The WCDs that are part of the base design (SD-1500) are symbolized with full triangles, the 49 additional WCDs of the SD-750 with empty circles and the six additional WCDs of the SD-433 with blue filled circles. The blue hexagons represent the current status of the SD-433 (modified from [159]).



**Figure 3.3:** (Left) Los Leones, one of the FD sites of the Pierre Auger Observatory. (Right) The inside structure of a FD telescope [162].

six telescopes, like the one shown in Fig. 3.3, left. Each telescope has a field of view (FoV) of  $30^\circ \times 30^\circ$  with a minimum elevation of  $1.5^\circ$  above horizon. The shutters are usually closed during day.

The FD operation is limited to clear, moonless nights, resulting in a duty cycle of around 15%. Whenever the weather conditions are sound, the shutters can be opened remotely and the data acquisition may be initialized. If a FD telescope is close enough to observe an EAS, the potentially scattered fluorescence photons will enter through a circular diaphragm with a diameter of 1.1 m. All telescopes are built with a *Schmidt* optics and equipped with an UV filter so that only photons with wavelength between 300 and 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<sup>1</sup>. The light is focused on the camera by a  $13 \text{ m}^2$  segmented mirror. The reflectivity of the mirror is above 90% in the UV range of interest. The camera is composed by 440 pixels arranged into 22 rows and 20 columns of *photomultiplier tubes* (PMTs). Each one gives a signal that is digitised with a 100 MHz *Flash Analog to Digital Converter* (FADC). All these elements are shown in Fig. 3.3, right.

### 3.2.2 Atmospheric monitoring

The FD operation and the post-processing of the measured signals require the knowledge of several atmospheric parameters [163]. The temperature, humidity and air pressure influence the longitudinal development of the EAS and the amount of emitted fluorescence light. Also clouds and aerosols may scatter and attenuate the fluorescence light during its propagation. By blocking the line of sight, cloud layers can bias the showers observation towards deeper penetrating events. Therefore, an extensive array of instruments were designed and are deployed to monitor these parameters [164].

Aerosol monitoring is performed using the *Central and eXtreme Laser Facilities* (CLF/XLF) and four elastic scattering Lidar stations, among other smaller facilities. During FD data taking, aerosol optical depth profiles are measured hourly by the CLF and the XLF [165], located at sites towards the centre of the SD-1500 array (see Fig. 3.1, top). The laser wavelength is fixed at 355 nm and the mean energy per pulse is around 7 mJ, similar to the amount of fluorescence light produced by an EAS with an energy of  $10^{20}$  eV. The laser beam is observed by all the FD sites<sup>2</sup>. The number of photons reaching the telescopes depends on the atmospheric conditions between the laser and the detector. The aerosol transmission and the presence of clouds within the laser site and FD can then be estimated.

Four elastic back-scatter Lidar stations located near each FD building are equipped with a UV laser and 3 PMTs for the detection of the elastic back-scattered light. During the data taking, a set of scans of the sky is performed hourly to record local aerosol and clouds coverage<sup>3</sup>.

Although these atmospheric monitors are sensitive to the presence of clouds, they do not provide a detailed all-sky map of cloud distributions. Thus an infrared camera is mounted on the roof of each FD site (7  $\mu\text{m}$  to 14  $\mu\text{m}$  wavelength range). It is sensitive to the temperature differences between clouds and clear sky (see Fig. 3.4). The cameras scan the FoV of the FD telescopes every 5 minutes, and also generate a full sky scan every 15 minutes. Information from the CLF, Lidars and clouds cameras are store in a database for crossing all the available information.

In addition to measurements with local weather stations, the density profiles of the atmosphere are estimated using the *Global Data Assimilation System* (GDAS) [167]. The day-to-day fluctuations of the pressure, temperature and humidity are estimated using meteorological radiosondes launched locally.

### 3.2.3 FD trigger

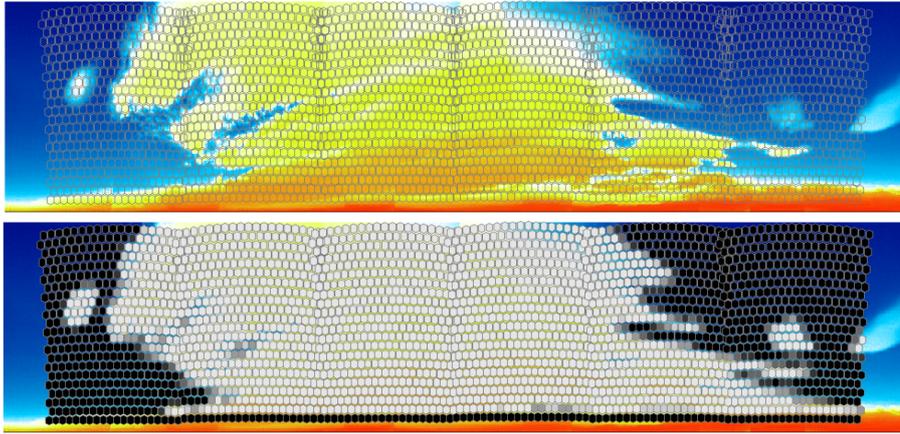
The camera of each FD telescope may register an EAS as a straight track in the  $20 \times 22$  pixel matrix above the sky light background. These tracks are identified by the hardware trigger system using a pattern recognition algorithm. The trigger logic has two hardware levels and one software level and it is optimized to have maximum efficiency at energies above  $10^{19}$  eV.

The first level trigger (FLT) decides the pixel status. The pixel response is a current pulse of 100 ns to about 2.5  $\mu\text{s}$  width [168], which is digitized at 40 MHz. The logic of the FLT consists of a

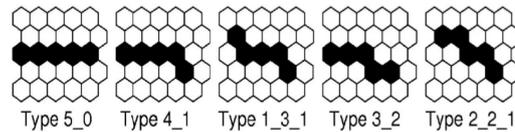
<sup>1</sup>Ongoing studies show that dust on the various parts of the system has a major impact on the level of 10% on the final energy scale of Auger [161].

<sup>2</sup>The light is collected with separate mirrors and PMTs at each FD site. Thus, this monitoring does not usually interfere with FD data acquisition.

<sup>3</sup>When a very high energy event is detected by FD and SD, the Lidar provides an additional monitoring of the aerosol and clouds content ("Shoot the Shower") by inspecting the region of the expected arrival direction of the shower [164].



**Figure 3.4:** Clouds observed by the infrared cloud camera: (top) the raw image and (bottom) the pixel mask. Lighter value of the gray scale denotes the presence of clouds [166].



**Figure 3.5:** The basic patterns used by FD trigger algorithm [169].

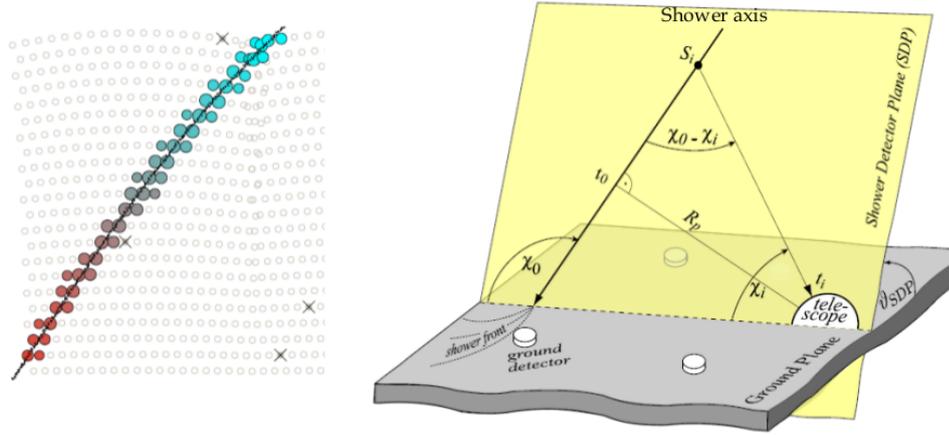
running sum over the last  $n$  ADC bins (each one of 25 ns), with  $5 \leq n \leq 16$ , which is compared to an adjustable threshold. A pixel trigger is generated if this sum exceeds the threshold, which is dynamically adjusted to keep the trigger rate close to 100 MHz<sup>4</sup>. The low statistics of the sky background (2.7 photoelectrons) is what dominates the threshold definition. The second level trigger (SLT) identifies track segments. A trigger is generated if there is a pattern of at least four triggered pixels that looks like a straight track produced by an EAS. In Fig. 3.5, five topological pattern types are shown. Finally, the third level trigger (TLT) rejects tracks conformed by pixels with a wrong time ordering and merges track segments by software. Events passing the TLT are sent to the *Central Data Acquisition System* (CDAS), where they are checked for coincidences with other telescopes triggers and the SD triggers to search for hybrid events.

The TLT may then act as an external trigger for the SD, allowing to record events below the SD full efficiency threshold. At these low energies, only one or two SD stations survive the trigger conditions, but this information together with the FD measurement is enough to ensure a high-quality hybrid reconstruction. The hybrid approach supplements the traditional FD direction fitting method with the arrival time of the shower at the ground measured by a single SD station. Accurate knowledge of the shower arrival time at ground level removes a degeneracy in the traditional FD monocular approach that uses pixel timing to reconstruct the shower axis. In hybrid mode, the resolution of the direction and of the position of the impact point at the ground are better than  $0.6^\circ$  and 50 m respectively [170, 171].

### 3.2.4 FD reconstruction

The start time of each pixel is used to determine the arrival direction of the EAS by treating the shower development as a series of point sources travelling along the shower axis (see Fig. 3.6,

<sup>4</sup>When the running sum is smaller than the threshold, the pixel trigger is extended for 5 to 30  $\mu$ s, which increases the probability of coincident pixel triggers.



**Figure 3.6:** (Left) The light track on the camera with the color scale representing the time ordering of the pixel triggers. The crossed pixels are rejected during the reconstructing process. (Right) The illustration of the shower geometry with the SDP in yellow [172].

left). The geometrical features considered in the FD reconstruction are shown in Fig. 3.6, right. The plane that contains both the shower axis and the telescope is named the *shower detector plane* (SDP).  $R_p$  represents the distance between the telescope and the shower axis. The angle between the shower axis and the ground is  $\chi_0$ . For each photon travelling from position  $S_i$  on the shower axis to the telescope, the angle between the shower axis and the telescope on the SDP is  $\chi_0 - \chi_i$ . The arrival time of the light at the pixel  $i$ ,  $t_i$ , is then:

$$t_i = t_0 + \frac{R_p}{c} \cdot \tan\left(\frac{\chi_0 - \chi_i}{2}\right) \quad (3.1)$$

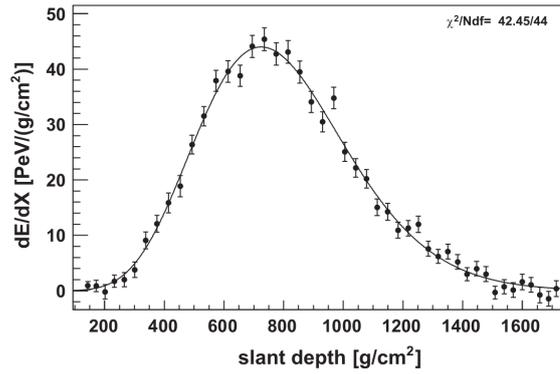
where  $\chi_i$  is the angle of the pointing direction of the pixel  $i$  with respect to the ground, and  $t_0$  the time at which the shower axis crosses the point at distance  $R_p$  from the telescope. The parameters  $R_p$  and  $\chi_0$  are obtained by a  $\chi^2$  minimization, in which the predicted values of  $t_i$  are compared to the data timing information. This procedure, that is solely based on FD signals, is named *mono reconstruction*. In addition, it is possible to combine SD and FD for the reconstruction, which is called the *hybrid reconstruction*.

Knowing the shower geometry, the FD absolute calibration<sup>5</sup> and the attenuation of the light flux in the atmosphere, the light collected by the FD telescopes as a function of the atmospheric depth  $X$  can be converted to the longitudinal profile of the energy deposit ( $dE/dX$ ) of the EAS, such as the example shown in Fig. 3.7.

The shower profile is fitted with the Gaisser-Hillas function (see Sec. 2.3, eq. 2.6). The calorimetric energy  $E_{\text{cal}}$  representing the electromagnetic component of the shower is obtained by integrating the curve. However, the energy carried by neutrinos, neutrons and high-energy muons cannot be measured by the FD. This fraction of energy is known as the *invisible energy* and has to be added to the calorimetric energy measured by the FD [173].

The invisible energy ( $E_{\text{inv}}$ ) can be calculated for each shower using the FD measurement of the longitudinal profile and the SD signal at 1000 m from the axis,  $S(1000)$ , which is introduced in Sec. 3.3.4.  $E_{\text{inv}}$  can be reliably estimated only above  $10^{18.5}$  eV (the energy above which the SD-1500 is fully efficient) as below this energy  $S(1000)$  is biased by upward fluctuations of the shower signals.  $E_{\text{inv}}$  is parameterised with an analytical function above that threshold, with the function

<sup>5</sup>An end-to-end calibration of all elements of each telescope is periodically performed with a drum-shaped light source placed in front of the diaphragm [169].



**Figure 3.7:** The energy deposit of the shower with respect to the slant depth of the atmosphere. The calorimetric energy of the shower is calculated by integrating the area under the profile [169].

being extrapolated to  $10^{17}$  eV. A set of hybrid showers was used to find the relation between  $E_{\text{inv}}$  and the calorimetric energy  $E_{\text{cal}}$ :

$$E_{\text{inv}} = a_0 \cdot \left( \frac{E_{\text{cal}}}{\text{EeV}} \right)^{a_1} \quad (3.2)$$

The fit was performed by minimising a  $\chi^2$  function which takes into account the fluctuations of both FD and SD measurements, yielding the parameters [174]:

$$\begin{aligned} a_0 &= (0.174 \pm 0.001) \times 10^{18} \text{ eV} \\ a_1 &= (0.914 \pm 0.008) \end{aligned} \quad (3.3)$$

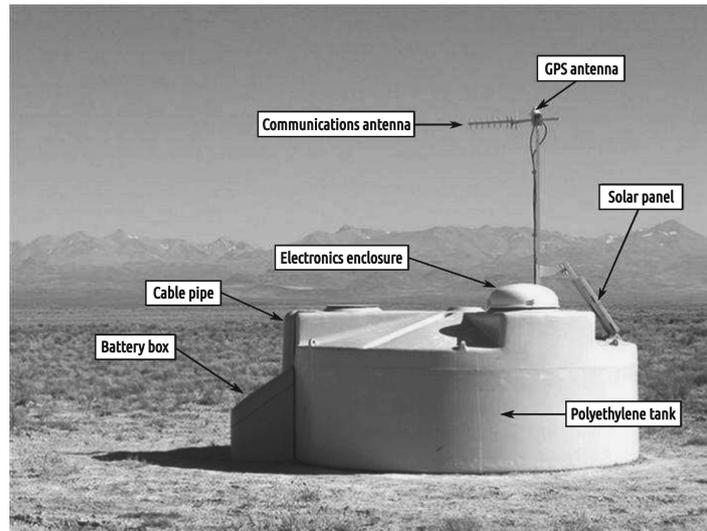
This contribution to the measured energy  $E_{\text{cal}}$  ranges between 15% at  $10^{18}$  eV and 11% at the highest energies. Due to the stochastic nature of air showers, the invisible energy is also affected by shower-to-shower fluctuations. These are parameterised and account for an uncorrelated uncertainty of about 1.5% [175].

### 3.3 The Surface Detector

The surface detector of the Auger Observatory consists of an array of more than 1600 stations separated by a distance of 1.5 km and arranged in a triangular grid, as mentioned before. Each WCD is a rotomolded polyethylene cylindrical tank of 1.6 m height and 3.6 m diameter filled with 12000 lt of ultra pure water (8 to 10 M $\Omega$  cm [176]). A sealed laminated liner with a reflective inner surface provides an interface between the water volume and the light sensors. The purpose of this liner is double: as a water-tight container to seal out external light, and to isotropically reflect Cherenkov light produced within the detector by traversing particles<sup>6</sup> independent of their arrival direction, given its high reflectivity (about 98% depending on wavelength). The Cherenkov light is collected by three 230 mm PMTs that look through windows of clear polyethylene into the purified water.

Access to the liner and the PMTs is through three hatches located on the top of the tank. An electronics box containing front-end charge amplifiers, shapers, trigger logic, signal buffers, power control, radio transmitter and receivers is located on the top roof of the tank on one of the

<sup>6</sup>Although only charged particles produce Cherenkov light, the WCDs are sensitive to photons since they cascade down through pair production in the water.



**Figure 3.8:** A picture of a water Cherenkov detector in the field, showing its main components [162].

hatch-covers and is protected by a dome. All the cables connecting the electronics and the light sensors run inside the tank and connect to the electronics via feed-throughs in the hatch covers. The different elements that compose an operating WCD are depicted in Fig. 3.8.

Each PMT provides two signals, which are digitized by 40 MHz 10-bit FADCs (bins of 25 ns). One signal is directly taken from the anode of the PMT and the other signal is provided by the last dynode, amplified 32 times and inverted within the PMT base electronics. The two kinds of signals are used to provide sufficient dynamic range to cover with good precision detection scenarios near the shower core ( $\sim 1000$  particle/ $\mu\text{s}$ ) and those produced far from the shower core ( $\sim 1$  particle/ $\mu\text{s}$ ) [157].

The WCDs are calibrated locally as discussed in Sec. 3.3.1. A hierarchical trigger system at the station level (see Sec. 3.3.2) is used to discriminate physical from spurious and accidental events. Once a physical event is detected, the traces are sent through a radio transmitter to the closest FD site, where a communication tower sends the data to the CDAS, where higher trigger conditions are applied.

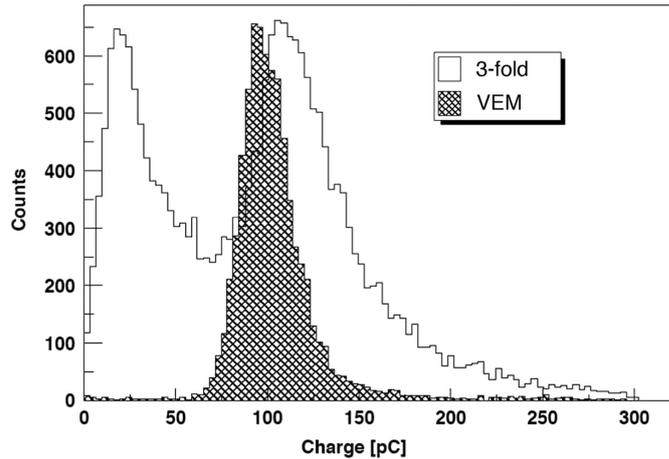
Each set of traces is tagged with a timestamp provided by a *Global Positioning System* (GPS) receptor attached to the WCD with a resolution of 7.2 ns. The GPS receptor is also used to monitor the position and altitude of each WCD with a  $\sim 1$  m precision. Each station has its power supply running autonomously with two 53 Wp solar panels and two 12 V batteries connected in series<sup>7</sup>. The hardware of the surface detector is extensively described in [162, 157].

### 3.3.1 Calibration of the WCD

With a large number of stations on the field, scattered over a very large area and often difficult to reach, it is important to be able to routinely calibrate and monitor each station remotely, with stable and robust procedures, and ensure a uniform response of the array in terms of trigger rates and performance.

The goal of the calibration procedure is to determine with good accuracy the reference signal

<sup>7</sup>A station with fully charged batteries can operate 7-10 days without further charging during a cloudy period. During all the operation of the observatory there has not been any general loss of operation due to extended cloudiness [177].



**Figure 3.9:** The charge spectrum obtained when a WCD is triggered by a 3-fold coincidence among its PMTs (open histogram). The hatched histogram shows the spectrum when triggered on central vertically aligned plastic scintillators. The bin containing the peak of the scintillator triggered spectrum is defined as a *vertical equivalent muon*. The leftmost peak in the open histogram is due to low energy and corner-clipping muons convolved with the 3-fold low-threshold coincidence [162].

unit for each PMT in terms of electronic units [178]. Because the total detected signal at each WCD depends on several parameters like the water quality, the liner reflectivity, the coupling between PMT and water and the PMT amplification factor, it is necessary to perform the calibration for each PMT of each station separately. Moreover, the calibration procedure has to be systematically repeated every 60 seconds to correct for temporal changes of the mentioned factors. The calibration is performed locally because of the limited bandwidth for data transmission to the CDAS (1200 bits/s).

The WCD calibration is inferred from atmospheric muons. The charge deposit of muons is proportional to their track length, as muons are not stopped in the tank. The reference signal is induced by a vertical and central through-going muon. Therefore, signals are measured in units of *vertical equivalent muons* (VEM). Since inclined muons cannot be distinguished from vertical ones in a WCD, the charge distribution of vertical and central muons ( $Q_{\text{VEM}}$ ) and the photocurrent they produce in the PMTs ( $I_{\text{VEM}}$ ) have been measured at a test WCD equipped with additional scintillators [179]. The corresponding charge histogram is shown in Fig. 3.9.

In order to determine the value of 1 VEM in integrated FADC channels, the following three steps are required:

1. The end-to-end gain of each of the three PMTs has to be chosen so that the current distribution has its peak  $I_{\text{VEM}}$  at ADC channel 50 above the baseline. Therefore, the sum of the three PMTs has a peak at channel 150. The choice of  $50 \text{ ch}/I_{\text{VEM}}$  results in a mean gain of  $\sim 3.4 \times 10^5$ .
2. After setting the gain of the three PMTs, the peak  $I_{\text{VEM}}$  is constantly adjusted, in order to compensate for drifts.
3. The corresponding charge  $Q_{\text{VEM}}$  is determined using charge histograms, in order to convert the integrated signals of the PMTs to VEM units.

Due to atmospheric muons, the charge distributions of the PMTs always exhibit a clear muon

peak. The charge distribution for each individual PMT has a peak at  $(1.03 \pm 0.02)$  VEM. For the sum of the three PMTs of a WCD, the charge distribution peaks at  $(1.09 \pm 0.02)$  VEM. This shift is the result of the asymmetry in the distribution of the track length of muons, as well as of defects in collecting the Cherenkov light.

The calibration constants  $Q_{\text{VEM}}$  and  $I_{\text{VEM}}$  are obtained with a 2% resolution and sent to the CDAS together with every triggering event.

### 3.3.2 WCD local triggers

The SD data acquisition trigger hierarchy must fulfill both physical and technical requirements. The main limitation to the rate of recordable events comes from the wireless communication system which connects the SD stations to the Central Campus. The latter must serve continuously 1600 stations spread over  $3000 \text{ km}^2$ , each using an emitter consuming  $< 1 \text{ W}$  power to transmit to collectors as far as 40 km away. The maximum sustainable rate of events per detector is  $< 1$  per hour, to be compared to the 3 kHz counting rate per station, due to atmospheric muons. Thus, the SD trigger system must reduce the single station rate, without inducing loss of physics events. At the same time, it must allow the data acquisition down to the lowest possible energy and should distinguish between real events and background. To satisfy these requirements, a hierarchical trigger system has been developed, where at each level the single station rate becomes smaller by means of stricter discrimination against background.

The first two trigger levels, namely T1 and T2, operate at the WCD electronics level. Their aim is to separate possible events from background noise. The T1 trigger contains two trigger modes conceived to detect, in a complementary way, the electromagnetic and muonic components of an EAS: the *Threshold* (Thr) trigger and the *Time over Threshold* (ToT) trigger.

In the first case, a coincidence in the three PMTs crossing the threshold of  $1.75 I_{\text{VEM}}$  in a single time bin is required<sup>8</sup>. This trigger is used to select large signals that are not necessarily spread in time. Thus, it is particularly effective for the detection of very inclined showers that have crossed a large amount of atmosphere and are consequently dominantly muonic. The rate of this trigger is about 100 Hz, in comparison to the rate due to atmospheric muons of  $\sim 3 \text{ kHz}$ .

The ToT condition asks for a coincidence of two PMTs with more than 12 ADC bins ( $> 325 \text{ ns}$ ) above  $0.2 I_{\text{VEM}}$  within a window of 120 ADC bins of a sliding window of  $3 \mu\text{s}$ <sup>9</sup>. This trigger criteria is very efficient in choosing small but spread signals, caused by low energy showers near the core or high energy showers at large distances from the shower core. The rate of the ToT trigger is  $< 2 \text{ Hz}$ . Single muon signals are ignored by this trigger since they are too short ( $\sim 150 \text{ ns}$ ).

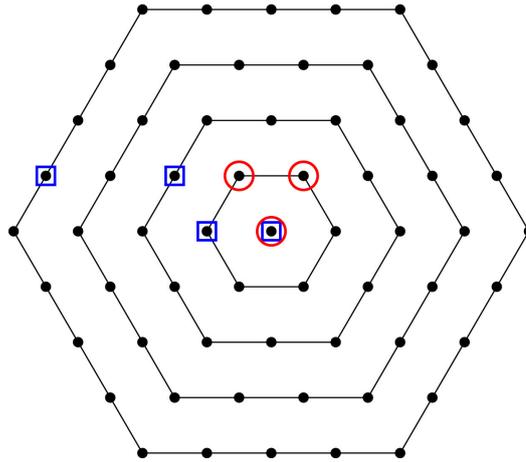
Signals passing the ToT trigger are automatically promoted to the second trigger level T2. In contrast, a software selection of the T1-Thr trigger with a higher threshold at  $3.2 I_{\text{VEM}}$  in coincidence of the three PMTs is performed<sup>10</sup> (referred as to T2-Thr). The T2-Thr trigger rate is about 20 Hz. The T2 triggered events are used to calculate the exposure of the array and to monitor the SD array performance. The station controller attaches GPS timestamps to the T2 signals and sends them to the CDAS for global (T3) trigger determination (see Sec. 3.3.3).

An improved version of the ToT trigger, the ToTd, was proposed as an optimisation especially suited for lower energy events [180, 181] and to improve the sensitivity to photon and neutrino

<sup>8</sup>For detectors with only two (one) operating PMTs the threshold is 2 (2.8)  $I_{\text{VEM}}$ .

<sup>9</sup>For detectors with only two (one) operating PMTs, the algorithm is applied to two (one) PMTs.

<sup>10</sup>For WCD with only two (one) operating PMTs the threshold is set to 3.8 (4.5)  $I_{\text{VEM}}$ .



**Figure 3.10:** Schematic representation of four hexagons of an ideal array containing SD stations. For a 3-fold coincidence, a T3 is issued if the 3 T2s are ToT, and if one of them is found in the first hexagon of the central station, and the other one no further than the second hexagon. A 4-fold coincidence applies to any kind of T2 and the additional station may be as distant as in the fourth hexagon. Two examples of the topology of triggers are shown: a 4-fold coincidence in which the triggered stations are identified by open blue squares, and a 3-fold coincidence identified by open red circles [162].

initiated showers. The underlying idea of this trigger is the deconvolution of the traces using the average response to single peaks. Due to multiple reflections, diffuse light and the decay time of the electronics, single peaks appear in the traces with an exponential tail. As a consequence, a deconvolved background trace would be reduced to a single peak, which would not survive the ToT condition, while a real signal would present several peaks, easily distinguishable from background. To account for some instabilities in the electronics, a revised version of the ToTd, the MoPS, was also introduced [182]. These new triggers are officially working since June 2013.

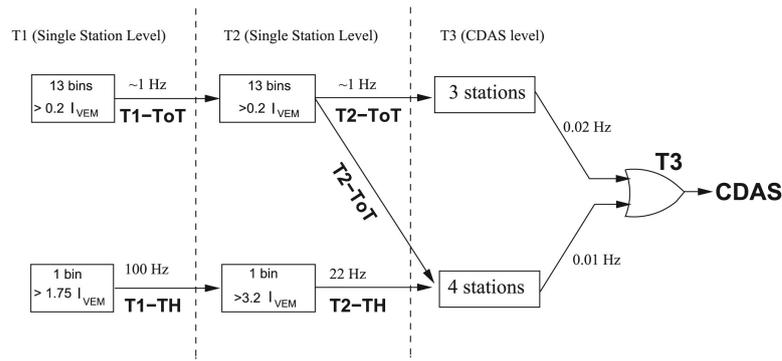
### 3.3.3 SD array triggers

**T3 trigger** The aim of the T3 trigger level is to select the possible physical events from the transferred data by looking for spatial and time coincidences between the T2 signals, and associating them to an EAS. Two kind of patterns are taken into account: 3-fold and 4-fold.

In the 3-fold scenario, T3 requires a coincidence of three neighbour stations with ToT trigger, such that two stations have to be in the first two crowns around the first one considered (see Fig. 3.10, empty circles). If this criteria is fulfilled, the pattern is tagged as T3-3ToT. As the ToT triggered events have a very low background, in comparison to Thr, the 3ToT trigger selects 90% of the physical events that are mainly vertical showers. The rate of T3-ToT with the full array in operation is around 1600 events per day, meaning that each detector participates in an event about three times per day.

For the 4-fold case, four stations with T2 (T2-Thr or ToT) have to be in coincidence. In this case, the distance requirement is looser, the fourth station being accepted if it is within four crowns around the reference station (see Fig. 3.10, empty squares). This condition is only relevant for showers of large zenith angle with triggered stations spread over larger areas, as nearly all vertical showers fulfil the T3-3ToT condition. This trigger selects about 1200 events per day.

However, an event fulfilling any of the T3 conditions may not necessarily be a real EAS. The T3 trigger was designed as a compromise between selection efficiency and purity. The trigger



**Figure 3.11:** The block diagram of the SD trigger hierarchy before the T3 algorithm [183].

hierarchy up to the T3 level is summarised in Fig. 3.11. The T3 triggered events are stored and then processed off-line at the CDAS.

**T4 physics trigger** To select only real physical events from the stored T3 data, the T4 trigger algorithm is applied. This physical trigger also consists of two different trigger conditions, the T4-3ToT and the T4-4C1 trigger. The T4-ToT is a stricter version of the T3-ToT that asks for a non-aligned compact 3ToT trigger (see Fig. 3.12b). The number of accidental coincidences passing the T4-3ToT condition over the full array is  $< 1$  per day, thanks to the very low rate of the T2-ToT. Due to their compactness, less than 2% of real EAS with zenith angles  $\theta < 60^\circ$  are discarded by this trigger<sup>11</sup>.

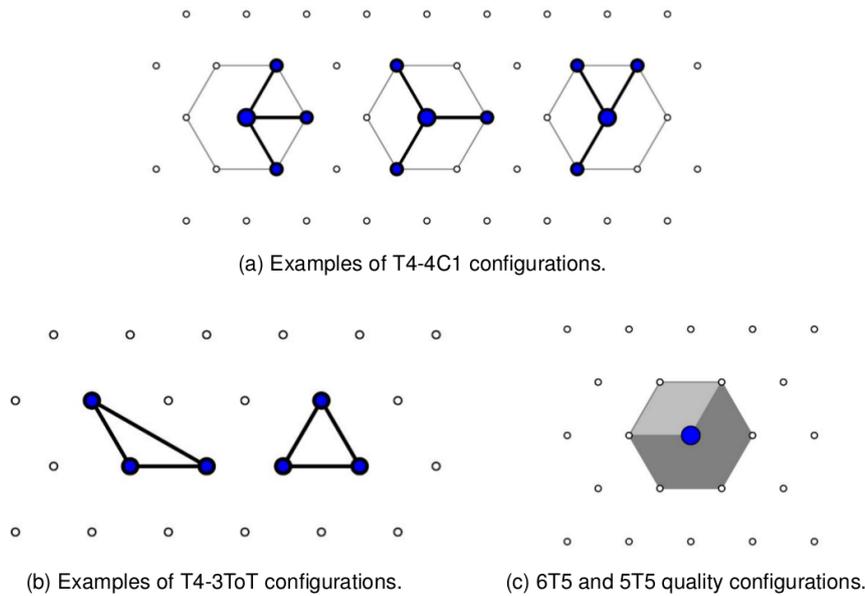
The T4-4C1 trigger chooses events with four stations with any T2 trigger aligned in a such a way that there are four stations in the first four crowns (see Fig. 3.12a). This trigger ensures the selection of nearly-horizontal shower events, discarded by the T4-3ToT trigger, bringing to  $\sim 100\%$  the efficiency for showers below  $60^\circ$ . For both trigger methods, the difference in their trace start time has to be smaller than  $d/c$ , being  $d$  the distance between two stations and  $c$  the speed of light in vacuum. In Fig. 3.13, the zenith angle and energy distributions of both trigger conditions are illustrated.

Stations that are triggered within the time window of an EAS, but originate from atmospheric muons or background radiation, are called accidental. In order to remove those WCDs from the event, a triangle conformed by the three WCDs with the highest signals, called the seed, is used to define the shower front of the event. Subsequently, all other stations are examined, and are defined as accidental if their time delay with respect to the front plane is outside a time window of  $[-2 \mu s; +1 \mu s]$ <sup>12</sup>.

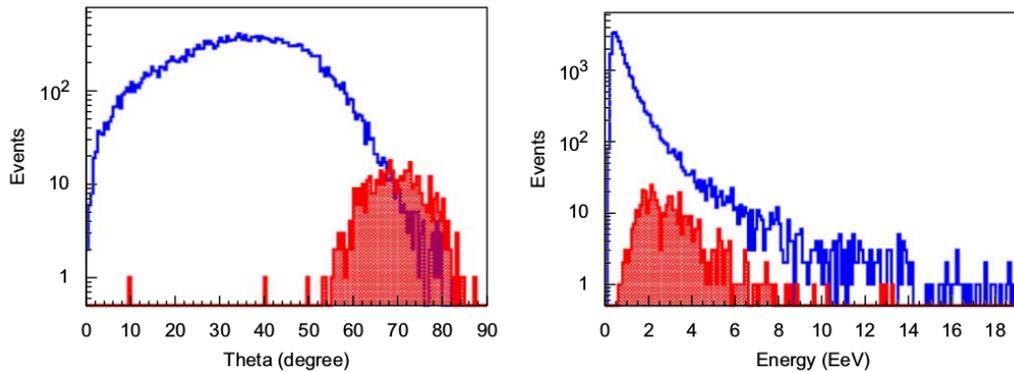
**T5 fiducial trigger** The main purpose of the last trigger level is to select only events for which an accurate energy and geometry reconstruction can be achieved. Events located close to an array border are often wrongly reconstructed because they span an incomplete footprint. These events

<sup>11</sup>The air showers with an inclination above  $60^\circ$  have different properties and require specific selection criteria described elsewhere [184].

<sup>12</sup>In addition, a triggered station with no other triggered WCD in 1800 m, or only one in 5000 m, around it is discarded. Stations with lightning-like signals (oscillations in the ADC traces of all three PMTs), those belonging to doublets (i.e. pairs of stations located very close that are used to study signal and timing accuracy) or those used for testing purposes are also discarded.



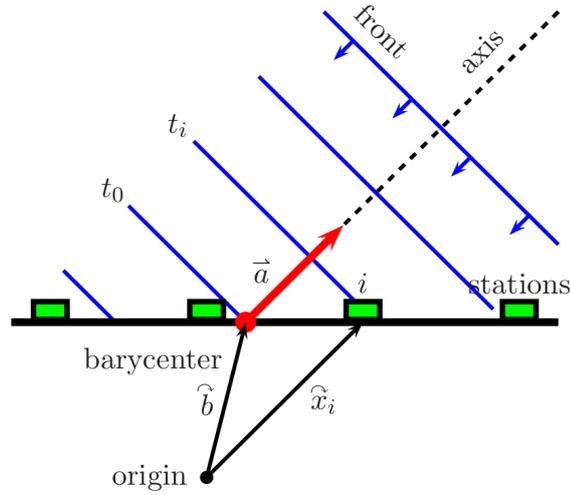
**Figure 3.12:** T4 and T5 configurations. (a) The three minimal compact configurations for the T4-4C1 trigger. (b) The two minimal compact configurations for the T4-3ToT configuration. (c) Example of the 6T5 hexagon (shadow) and the 5T5 hexagon (dark shadow) [185].



**Figure 3.13:** The zenith angle (left) and the energy (right) distribution of T4 triggered events. The events fulfilling the T4-3ToT (T4-4C1) condition triggered events are represented by the red shaded (blue) histogram. The events passing both T4 triggers are counted as T4-3ToT events [186].

should be rejected by the quality trigger T5. The most common T5 triggers are the 6T5 and 5T5 triggers. The 6T5 trigger requires an active unitary cell, i.e. the station with the highest signal (hottest) has to be surrounded by six active stations (not necessarily triggered). On the other hand, the 5T5 is a loose version of 6T5 requiring that the hottest station has to be surrounded by five active stations and the core has to be located within an active triangle of stations. Both trigger methods are shown in Fig. 3.12c. It has to be noted that this criterion also discards events that, though contained in an area of active WCDs, fall close to a non-working station<sup>13</sup>.

<sup>13</sup>Even with constant maintenance over the whole SD array, about 1% of the stations are expected to be not functioning at any given moment.



**Figure 3.14:** The schematic illustration of the planar-front approximation used for a first estimate of the incoming direction of the primary particle [188].

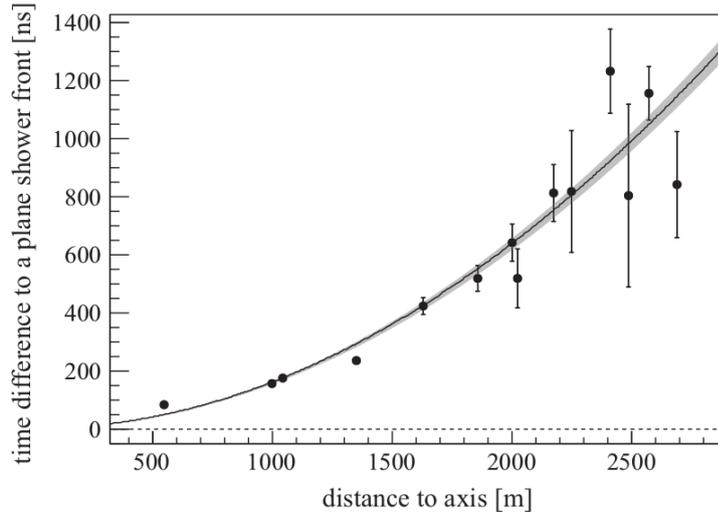
### 3.3.4 SD reconstruction

After the trigger hierarchy selection and the accidental stations removal, more than 99% of the events can be successfully reconstructed, i.e. the arrival direction described by the zenith angle  $\theta$  and the azimuth angle  $\varphi$ , core position  $\vec{x}_0$  and energy  $E_{\text{rec}}$  can be determined. General details on the SD-only reconstruction procedure, without specifying a particular SD array, are given in this section. In Sec. 3.3.6, details about the geometrical and energy reconstruction in the case of SD-433 are discussed.

The reconstruction procedure uses the signals and the traces start time of the triggered WCDs as the main ingredients [187]. The objective is to determine the lateral distribution function (LDF), the incoming direction of the primary and the impact point on ground of the EAS. Afterwards, the primary energy estimator  $S_{\text{opt}}$  is found by evaluating the obtained LDF at a pre-defined distance from the shower axis  $r_{\text{opt}}$ , which depends on the particular spacing of the detector grid. Although the reconstruction is performed by using an iterative maximum likelihood process, it can be broken down into three steps, which are discussed in the following.

**Plane fit to the shower front** The seed for the geometry reconstruction is produced during the T4 trigger stage. In a simplified description, the geometry of the shower can be reconstructed by fitting the time and space information of the triggered WCDs. A first estimate of the arrival direction of the cosmic ray consists in approximating the shower front to a plane front moving at the speed of light  $c$  with a normal direction  $\hat{a} = (u, v, w)$  pointing towards the primary particle. The first estimate of the core position  $\vec{x}_0$  (i.e. the shower impact point on the ground) is a signal-weighted barycenter ( $\vec{b}$ ) of the stations involved in the event (see Fig. 3.14) while the shower arrival time to ground  $t_0$  is the weighted bary-time.

To infer the time  $t(\vec{x})$  when the shower plane passes the point  $\vec{x} = (x, y, z)$ , the point has to be projected to the shower coordinate system (Eq. 3.4). Specifying the generic point  $(t, \vec{x})$  as the signal start time  $t_i$  and the WCD position  $\vec{x}_i$  (which is supposed to be given with absolute precision), the unknown variables are the three coordinates of the shower axis  $\hat{a}$ .



**Figure 3.15:** The dependence of signal start times (relative to the timing of a plane shower front) on perpendicular distance to the shower axis. The shaded line is the resulting fit of the evolution model and its uncertainty [162].

$$ct(\vec{x}) = ct_0 - (\vec{x} - \vec{b}) \cdot \hat{a} \quad (3.4)$$

Therefore, the function to minimize is the square of the time differences between the measured start of each signal  $t_i$  and the predicted start time  $t(\vec{x}_i)$ :

$$\chi^2 = \sum_i \frac{|(t_i - t(\vec{x}_i))|^2}{\sigma_{t_i}^2} = \sum_i \frac{|c(t_i - t_0) + \vec{x}_i \cdot \hat{a}|^2}{c^2 \sigma_{t_i}^2} \quad (3.5)$$

with  $\sigma_{t_i}$  as the time uncertainty of the signal start of the station  $i$  at position  $\vec{x}_i \equiv \vec{x}_i - \vec{b}$  [189].  $\hat{a}$  is a unitary vector by construction so this problem is non-linear. An approximate solution could be obtained if it is supposed that all stations lie close to a plane, so  $z_i \ll x_i, y_i$  and thus the  $z$  component is neglected (more details in [187]).

This approximation serves as a starting point to a more elaborate model which takes into account the varying altitude of the stations and is based on a curved front fit. It extends the plane fit method with a parabolic term that describes the curvature of the shower front near the impact point (see Fig. 3.15). In this model, the time propagation of the shower front is described as a speed-of-light inflating sphere:

$$c(t_i - t_0) = |\vec{x}_{\text{sh}} - \vec{x}_i| \quad (3.6)$$

where  $\vec{x}_{\text{sh}}$  and  $t_0$  are a virtual origin and start time of the shower development. From this 4-parameter fit the radius of curvature of the inflating sphere is determined from the time at which the core of the shower is inferred to hit the ground. The difference between this 3D fit and the first approximation ( $z_i \ll x_i, y_i$ ) is of few tens of meters in the core position and  $0.1^\circ$  in the shower axis direction.

**The lateral distribution function** The lateral dependence of the signal measured in the WCDs is modeled as:

$$S(r) = S_{\text{opt}} \times f_{\text{LDF}}(r) \quad (3.7)$$

where  $f_{\text{LDF}}(r)$  is a particular shape parametrization normalized such that  $f_{\text{LDF}}(r_{\text{opt}}) = 1$  holds. Several functional forms of the LDF have been investigated [190, 191], and it was found that the best description of data is given by a modified Nishimura–Kamata–Greisen function [145]:

$$f_{\text{LDF}}(r) = \left(\frac{r}{r_{\text{opt}}}\right)^{\beta} \times \left(\frac{r + r_{\text{scale}}}{r_{\text{opt}} + r_{\text{scale}}}\right)^{\beta} \quad (3.8)$$

Similarly to the fit of the shower front, the initial guess for  $S_{\text{opt}}$  is solely determined on the signal of the station closest to the distance  $r_{\text{opt}}$  from the shower axis. The optimal distance  $r_{\text{opt}}$  is defined as the distance on the shower plane where the LDF slope fluctuations minimally affects the signal model. In other words, it is the point where changes in the slope (induced by shower-to-shower fluctuations and statistical fluctuations in the number of particles) have the minimum influence on the LDF. This parameter depends on the array geometry, in particular on the mutual distance between the stations. This parameter has been fixed to 250 m for the SD-433 array [141], 450 m for the SD-750 array [159] and 1000 m for the SD-1500 array [192]. The scale parameter  $r_{\text{scale}}$  plays a role only at larger distances from the shower axis, and in addition it is strongly correlated to  $\beta$ , hence it has been kept fixed to its hard-coded value of 700 m.

The shape parameter  $\beta$  defines the steepness of the LDF. It can be included in the reconstruction procedure as a free parameter estimated by the global fit. However, a good reconstruction is achieved only if specific conditions regarding the spatial distribution of triggered stations (i.e. information about the footprint of the EAS) are met. The official choice is to fix the slope according to the following model:

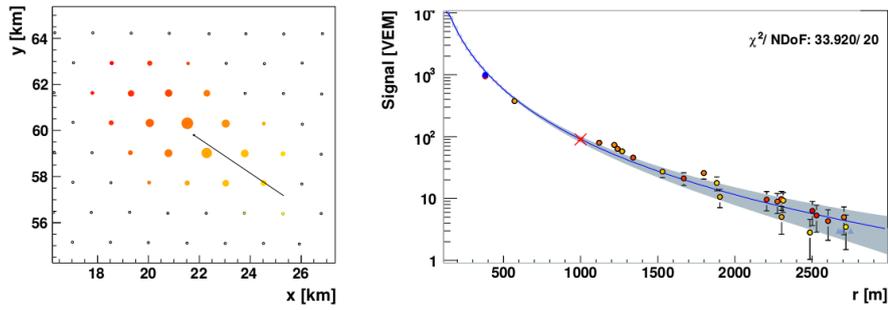
$$\beta(\log_{10} S_{\text{opt}}, \theta) = a + b \log_{10} S_{\text{opt}} + (c + d \log_{10} S_{\text{opt}}) \cdot \sec\theta + (e + f \log_{10} S_{\text{opt}}) \cdot \sec^2\theta \quad (3.9)$$

where the six parameters are obtained through simulations of the particular array geometry.

Lastly, a maximum likelihood method is performed to find  $S_{\text{opt}}$ ,  $\vec{x}_0$ ,  $\theta$  and  $\varphi$ . An example of a fitted LDF and the event footprint is displayed in Fig. 3.16. This procedure allows the inclusion of zero-signal stations, stations of small signals (i.e. small particle densities) by means of Poissonian statistics and the signal of saturated stations as (so far) lower signal limit. In all cases, the uncertainty on the measured signal  $S_i$  in the WCDs depends linearly on  $\sec\theta$  and  $\sqrt{S_i}$  [193].

**Treatment of the saturated stations** In order to have a linear PMT response in a large dynamic range, each PMT provides two outputs, as mentioned before. One signal is directly taken from the anode of the PMT (low gain channel), while the other is provided by the last dynode (high gain channel). Thus, when the dynode signal saturates, the anode signals are used to generate the signal trace.

But eventually, the anode signal may saturate when the detector measures a large number of particles. Different methods which take into account the ADC channel overflow as well as the PMT non-linearity have been tested in order to estimate a lower limit to the saturated signal. The recovered signal is used only when the second derivative of the normalized LDF is smaller than 1 in order to avoid the rapid increase of the LDF approaching the shower core. In the standard reconstruction, the saturation recovery is not activated by default, while the user could switch it



**Figure 3.16:** (Left) The footprint of an EAS at ground. The color code represents the arrival times, the red WCDs triggered later than the yellow ones. (Right) The reconstructed LDF. The saturated WCD close to the core (red) is recovered (blue) and used in the fit of the measured signals. The reconstructed energy is  $E_{\text{rec}} = 4 \times 10^{18}$  eV and the zenith angle is  $\theta_{\text{rec}} = 57^\circ$  [194].

on [195, 196].

The study on how to recover the saturated signals is relevant since even saturated station introduces a systematic uncertainty in the LDF determination that gets propagated into the estimation of the primary energy, as it is described in Sec. 4.4.

**Primary energy determination** The primary energy is estimated mainly through the LDF normalization  $S_{\text{opt}}$  (see Eqs. 3.7 and 3.8). However, this parameter decreases with increasing zenith angle because of the atmospheric attenuation of the shower particles and geometrical effects [197]. This dependency can be disentangled by means of the *Constant Intensity Cut* (CIC) method [198]. Assuming that the high-energy cosmic-ray flux is isotropic,  $S_{\text{opt}}$  is converted into a reference signal size  $S_{\text{ref}}$  by means of the attenuation curve  $\text{CIC}(\theta)$  as:

$$S_{\text{ref}} = \frac{S_{\text{opt}}}{\text{CIC}(x)} \quad (3.10)$$

with  $x := \cos^2\theta - \cos^2\theta_{\text{ref}}$  and  $\text{CIC}(x)$  a second-order polynomial determined by the method. An energy estimator  $S_{\text{ref}}$  for each event, independent of  $\theta$ , is the  $S_{\text{opt}}$  that a shower would have had if it had come with a zenith angle of  $\theta = \theta_{\text{ref}}$ . The reference angle  $\theta_{\text{ref}}$  represents the median zenith angle, so the impact of the correction factor is minimized. In the case of the SD-1500,  $\theta_{\text{ref}} = 38^\circ$ .

The energy density of shower particles at ground level is used to determine the cosmic-ray energy if the corresponding EAS is detected solely by the SD. On the other hand, the observation of the longitudinal profile of the shower with the FD allows the measurement of the calorimetric energy  $E_{\text{cal}}$  of the event. The sample of events reconstructed by both detectors, though small with respect to the whole SD sample, is very important since it constitutes the base data set for the energy calibration of the SD events [197]. Using the information provided by the FD, it is possible to obtain the energy corresponding to each  $S_{38}$  almost entirely from data, except for assumptions about the missing energy. Only data with a successful hybrid geometry reconstruction were selected to parametrize the CIC method [199]. Data is fitted with a power law [200]:

$$E_{\text{FD}} = a \cdot S_{38}^b \quad (3.11)$$

For example, the constants  $a = (1.68 \pm 0.05) \times 10^{17}$  eV and  $b = (1.035 \pm 0.009)$  have been obtained through a comparison of the measured  $S_{38}$  and the energy reconstructed from FD data

( $E_{FD}$ ) for a set of well-measured hybrid events above  $10^{18.5}$  eV with the SD-1500 [197].

### 3.3.5 The SD-750 array

The SD-1500 has a full efficiency threshold at  $10^{18.5}$  eV [183]. In order to study the energy spectrum window between  $10^{17.5}$  eV and  $10^{18}$  eV, where it was supposed that the transition from Galactic to extragalactic cosmic rays should lie (see Sec. 1.4), the SD-750 was built. Completed in September 2012, the SD-750 comprises an area of  $23.5 \text{ km}^2$ . This denser array benefits from the preexisting infrastructure, communication system and data acquisition tools. The extra 42 WCDs needed to complete the grid have the same design as the WCDs employed in the SD-1500.

Such an extension permitted a thorough study of the cosmic-ray energy spectrum down to  $\sim 10^{17.5}$  eV. Four other experiments have shown evidences for a *second knee* feature (i.e. change in the slope) in spectrum in the  $10^{17}$  eV decade that could be related to a change in the sources or the composition of the primary particles: Akeno [201], Haverah Park [193], Fly's Eye [202] and HiRes [203]. However, analyses performed with the SD-750 showed no clear indication of such break in the energy spectrum [159]. This pointed out the need to pursue the second knee at lower energies, motivating further spectrum studies with an even denser SD array.

### 3.3.6 The SD-433 array

The KASCADE-Grande Collaboration has observed a second knee structure in the cosmic-ray spectrum at  $\sim 10^{16.9}$  eV [204]. Since SD-750 is not fully efficient at such low energies, a denser array would confirm the existence of this feature.

The original proposal for the AMIGA enhancement included, not only a graded infill of 750 m separation, but also a denser infill with a separation of 433 m. The SD-433 array has been operational since January 2013. Although the base design was composed of 43 WCDs, it currently consists on 13 WCDs<sup>14</sup>. Six WCDs were arranged in an hexagonal shape around the *Kathy Turner* (station Id 1764) position, which is already part of the SD-750. Six more WCDs, also originally from the SD-750, surround this hexagon in an incomplete second crown, as illustrated in Fig. 3.17.

To get a better idea of the capabilities of the radio technique in determining the energy, or any other shower parameter, the SD-433 was also meant to be used as an independent way to establish the primary energy of events detected in coincidence with the Auger radio-detection extension (see Sec. 3.5.1). Analyses on the geometrical reconstruction of the SD-433, such as the uncertainty of the reconstructed core position and angular resolution, were carried out using data reconstructed simultaneously by the SD-750 and SD-433 [205]. A dedicated energy calibration and LDF parameterization were obtained and incorporated in the official reconstruction procedure [206]. A brief summary of the main features regarding the reconstruction procedure is given in this section.

The LDF parameters are essential to estimate the primary energy of a detected event, as explained in Sec. 3.3.4. An optimization of the values of  $r_{\text{opt}}$ ,  $S_{\text{opt}}$  and  $\beta$  was performed. It was found that for SD-433 events the distance where the slope fluctuations are minimal is 250 m [207]. Therefore, it is considered that  $r_{\text{opt}} = 250 \text{ m}$  and thus  $S_{\text{opt}} \equiv S_{250}$ . The  $\beta$  parameterization was

<sup>14</sup>In July 2017, three more WCDs were added to the second crown of the SD-433. However, as it will be discussed in Sec. 7.1.1, the considered data for the present analysis was taken without this extension.



**Figure 3.17:** The schematic map of the SD-433 array. The green (gray) icons represent the location of WCDs with (without) an associated MD station. The Id 1764 WCD and its associated MD station are located in the center of the SD-433 hexagon.

Parameter	a	b	c	d	e	f
Value	-4.5	1.2	3.4	-1.9	-0.9	0.7
Errors	1.2	0.8	1.9	1.3	0.8	0.5

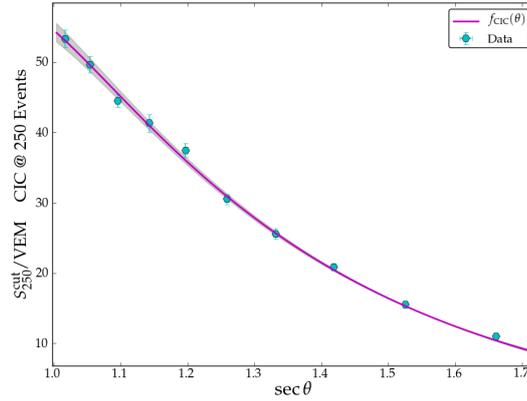
**Table 3.1:** The parameters of the  $\beta$  model specified in Eq. 3.9 for the SD-433 reconstruction with their associated uncertainties [205].

done adopting a multi-event fit approach, rather than an event-by-event approach, in order to increase the number of data points for the selected sample of events [205]. By minimizing the difference between the measured signal in the WCDs and the model given by Eq. 3.7, the set of parameters that determine the  $\beta$  dependence on the  $S_{\text{opt}}$  and  $\theta$  were determined (see Eq. 3.9). The results are shown in Tab. 3.1.

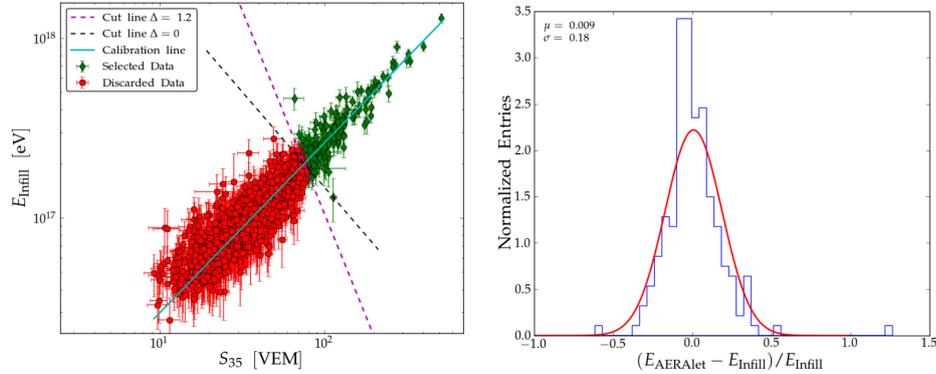
The CIC method was used to obtain a zenith-independent energy estimator. In this case, a reference zenith angle of  $\theta_{\text{ref}} = 35^\circ$  was chosen. Contrarily to the CIC method applied to the standard SD-1500 array, a third-order polynomial was adopted for the function  $\text{CIC}(x)$ , which is shown in Fig. 3.18 for a constant intensity cut at 250 events [206]. The energy calibration of the SD-433 was performed exploiting a sample of events that had been detected both by the SD-750 and the SD-433. This unusual approach (“SD-SD”) calibration was preferred due to the very scarce number of events detected by the SD-433 in coincidence with the FD. Thus, the reference energy used in Eq. 3.11 was the energy reconstructed by the SD-750, namely  $E_{\text{infill}}$ . Fig. 3.19, left, shows the relation between  $S_{35}$  and the reference energy  $E_{\text{infill}}$  with a least-square fit of Eq. 3.11. The resolution of the energy conversion is shown in Fig. 3.19, right. By means of a Gaussian fit to this histogram, the energy resolution for the SD-433 array was found to be 18% [206]. The final conversion formula is given by:

$$E_{\text{AERAlet}} = a \cdot S_{35}^b = (3.39 \pm 0.83) \cdot S_{35}^{0.94 \pm 0.05} \times 10^{15} \text{ eV} \quad (3.12)$$

This conversion formula was obtained by using events above  $10^{17.2}$  eV and optimized for estimated energies above  $10^{16.9}$  eV below which trigger efficiency effects of the SD-433 have to be taken into account [206]. However, as it will be shown in Chapter 4, it has a suitable performance at even lower primary energies depending on the primary zenith angle.



**Figure 3.18:** The fitted  $CIC(x)$  correction function as a function of  $\sec\theta$  compared to the data extracted by performing a CIC at 250 events. The shaded area represents the  $1\sigma$  region obtained from the parameters uncertainties [206].



**Figure 3.19:** (Left) The conversion between the SD-433 energy estimator  $S_{35}$  and the reference energy given by the SD-750. The cyan-solid line represents the energy-calibration function obtained by a least-squares fit of 193 events detected both by the SD-750 and SD-433 [206]. (Right) The relative differences between the converted energy  $E_{AERAlet}$  by Eq. 3.12 and the reference energy  $E_{infill}$  [206].

### 3.4 The AMIGA Muon Detector

As mentioned before, the AMIGA enhancement comprised the deployment of denser SD arrays and the deployment of a dedicated *muon detector*. The engineering array of the MD (called *Unitary Cell* or UC) was finished at the end of 2014. It consists on a hexagonal cell of seven MD stations, each one paired with a WCD from SD-750 from which it receives the trigger condition.

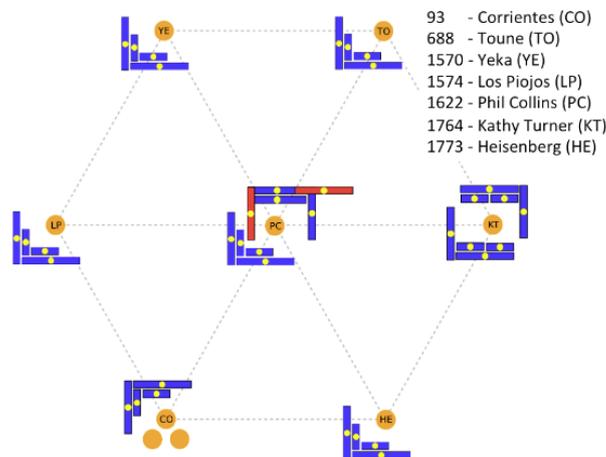
In this section, the elements composing a MD station are described. The order reflects the steps involved in the detection of a muon, starting with the spatial distribution of modules in the site and finishing with the off-line muon counting strategy and calculation of the *muonic LDF* (MLDF).

#### 3.4.1 Layout of the MD Unitary Cell

The MD was designed to measure the muon content of the EAS by employing an already



**Figure 3.20:** (Left) The deployment of the first MD module of the UC. The electronics kit, which is contained in a PVC casing, is shown in the inset [209]. (Right) Installation of an access tube during module deployment. The access tube is sealed to the module for water tightness [210].

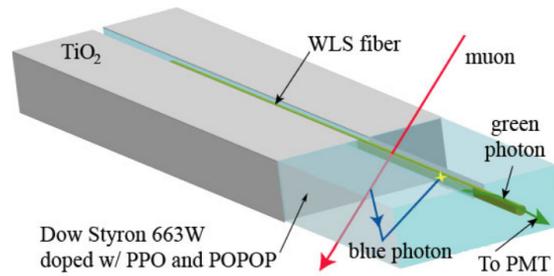


**Figure 3.21:** The layout of the UC, the engineering array of the MD, showing the locations of the MD stations and corresponding WCDs. The size of the modules represents their nominal area ( $10 \text{ m}^2$  or  $5 \text{ m}^2$ ) [210].

well-proven technique: a collection of buried scintillation plastic bars [208]. 64 scintillation bars with optical fibers are lodged within a *Polyvinyl Chloride* (PVC) casing and together with the optic system and electronics, both contained in a dome, form a MD *module*. Each module is deployed separately as shown in Fig. 3.20, left. Both the optic detection system and the electronics are located in the center of the MD module and can be reached from the ground through an access pipe, which can be seen in Fig. 3.20, right.

With a height of  $\sim 1 \text{ cm}$  and a width of  $4 \text{ cm}$ , the scintillation bars can have a length of  $4 \text{ m}$  or  $2 \text{ m}$ . In the first case, the resulting module would have a nominal area of  $\sim 10 \text{ m}^2$  while in the second case  $\sim 5 \text{ m}^2$ . Each MD station can be composed of three or four modules, depending on its position in the UC. In Fig. 3.21, the layout of the UC is shown with the corresponding distribution of modules. Twin MD stations of  $\sim 60 \text{ m}^2$  ( $4 \times 10 \text{ m}^2$  and  $4 \times 5 \text{ m}^2$ ) at two locations in the UC permitted thorough analyses regarding the systematic uncertainties of the detector [211].

The MD stations lie a few meters next to the SD station to prevent any possible shadowing of particles by the WCD [212]. The scintillators are buried at a depth of  $\sim 2.3 \text{ m}$  corresponding to a vertical overburden of  $\sim 540 \text{ g cm}^{-2}$  (considering an average local soil density of  $(2.38 \pm 0.05) \text{ g cm}^{-2}$  [210]). The soil shields nearly all the electromagnetic particles from the EAS, and thus, only en-



**Figure 3.22:** A schematic view of the scintillation process when a muon impinges a plastic bar. The WLS fiber collects the blue photons and re-emits them with a larger wavelength [209].

ergetic and penetrating particles (mostly muons) are capable of reaching the scintillation bars<sup>15</sup>. At the same time, the distance between the MD and its associated WCD is small enough to represent the same physical point in the shower front and to be able to share GPS time signals and telecommunications.

The modules are deployed in each position in an “L” shape. This allowed to study the *clipping corners* effect, which is produced by a muon transvering two scintillation bars and leaving a detectable signal at each one. In this scenario, a systematic overcounting may be produced depending on the temporal structure of the signal produced at each bar. A correction based on a parameterized model depending on the primary energy and zenith angle has been developed and is already included in the official MD event reconstruction chain [214]. Nevertheless, it is customary to impose a constrain on the maximum zenith angle of  $\theta = 45^\circ$  that an event can possess in order to reduce the impact of the clipping corners effect.

### 3.4.2 Scintillation bar - fiber tandem of a MD module

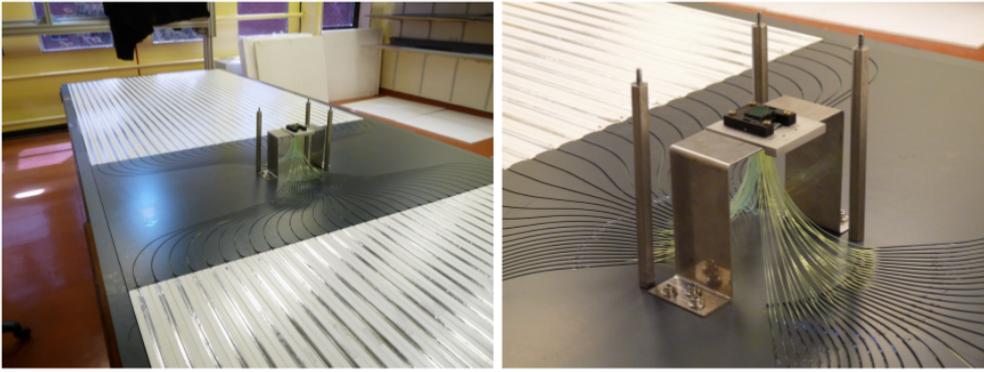
The scintillation bars have a middle groove which lodges a *wavelength shifter* (WLS) fiber glued into it and covered with reflective foil. The strips are co-extruded with a  $\text{TiO}_2$  reflective coating which prevents the light from leaving the scintillator. The fibers absorb the fluorescence photons emitted by the scintillation bar and re-emit them with a larger wavelength. Since the wavelength is shifted, most of these photons are not absorbed. Thus, they are transmitted towards the optic detection system. The scintillation process is illustrated in Fig. 3.22.

The 64 scintillation bars are arranged in two halves of 32 each. The corresponding fibers are grouped in a manifold in the middle of the MD module, as shown in Fig. 3.23, left. The endpoint of the fibers are glued into a PVC adapter or *cookie* which acts as an interface between them and the optic system, as seen in Fig. 3.23, right.

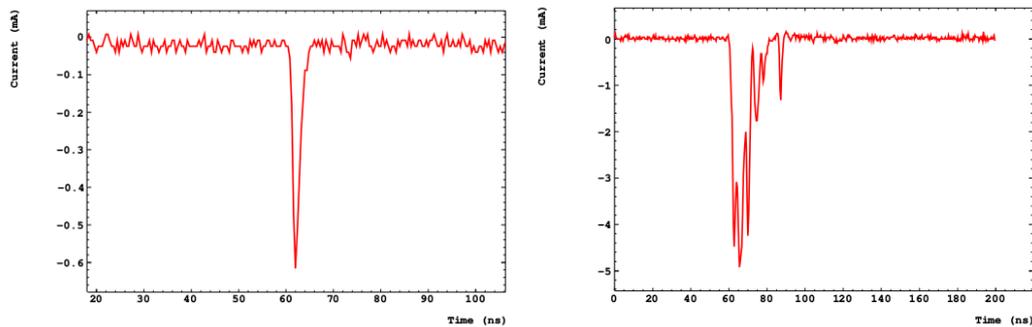
### 3.4.3 Optic system of the MD

The optic system employed in the UC is a *multi-anode* PMT. Although the PMT has a single photocathode, it is divided in 64 pixels. Each pixel is aligned to an optic fiber through the cookie. A special optic grease is employed in the interface to avoid possible air bubbles. Photons coming through an optic fiber imping on the photocathode and may strip an electron from

<sup>15</sup> The energy cut-off in the energy spectrum of vertical muons is close to 1 GeV [213].



**Figure 3.23:** (Left) A 5 m<sup>2</sup> MD module after the mechanical assembly of scintillation bars and fibers. (Right) The PVC adapter or *cookie* that mediates between the optic fibers and the optic detection system [215].



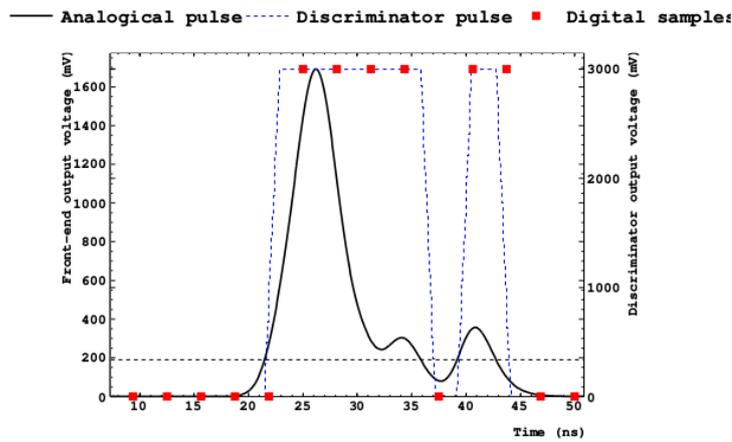
**Figure 3.24:** The typical shape of a pulse at the PMT anode generated by a single photoelectron (left) and a muon impinging on a scintillation bar (right) [209].

the photocathode atoms. Since these electrons emerge by the photoelectric effect, they are called *photoelectrons* (PE). In this sense, another advantage of using a specific WLS fiber is that its emission spectrum can be coupled to the absorption spectrum of the photocathode of the PMT so that its quantum efficiency is maximized<sup>16</sup>.

One PE may start an avalanche process through the internal multiplication stages of the corresponding pixel producing a detectable current at the end of the chain, i.e. at a PMT anode. This multiplication stage is fed by a voltage difference between successive dynodes. The tandem conformed by a single pixel, its multiplication chain and its corresponding readout anode is called *channel*. In Fig. 3.24, left, an output signal produced by a single PE is shown. The typical signal produced by a muon is a superposition of many PE distributed in time and amplitude, so that it has a more complex shape, as it can be seen in Fig. 3.24, right. Therefore, a counting strategy taking into account the possible structure of the signal was needed (see Sec. 3.4.5).

The main sources of systematic effects at this stage are the *dark current* and the *cross-talk*. The dark current is spurious noise generated in the absence of impinging photons on the photocathode. In this case, the avalanche is initiated at some point of the multiplication stage by thermal electrons. The cross-talk consists on a signal generated in a channel at some point of the multiplication stage of a neighbour channel. It can be classified into *optical* and *electronic* cross-talk. The optical cross-talk is produced at the very beginning of the avalanche process, i.e. an impinging photon on a neighbour pixel which strips a PE with high transverse momentum. The electronic cross-talk

<sup>16</sup>The quantum efficiency is the number of photoelectrons emitted from the photocathode divided by the number of incident photons. It depends on the wavelength of the incident radiation.



**Figure 3.25:** The discrimination and digital sampling of the pulse after the amplification phase by the front-end electronics. The input (output) signal to (from) the discriminator is represented in solid black (dotted blue). The threshold level is represented by the dotted black line. The FPGA digital samples are represented by red squares [209].

is produced at an intermediate point of the avalanche by a single electron that trespasses to a neighbour channel. In both cases, a channel which did not receive any photon on its pixel would be activated. These systematic effects have been quantified and introduced as correction factors in the MD reconstruction chain [209, 216]. A cross-talk signal does not necessarily leads to a muon counting in the corresponding channel. Indeed, as it is discussed in Sec. 3.4.5, the muon counting is performed by a suitable strategy that takes into account the structure of the digital signal. In this way, the impact of the cross-talk signals in the measurement of the muon density is then negligible.

### 3.4.4 Electronics and data acquisition of the MD

The digitalization of the analog signal from each PMT channel is performed by the front-end electronics in three steps. In the first step, the signal coming from the PMT anode is amplified and inverted. Also, the high-frequency Fourier components ( $\gtrsim 300$  MHz) are attenuated by means of a low-pass filter. Afterwards, the signals are broader and devoid of spurious noise. These signals are compared to a tunable threshold that would produce a positive (negative) digital sample if the signal is above (below) it. The threshold is chosen as the 30% of the typical signal amplitude produced by a single PE and it can be monitored and set manually or through a calibration algorithm for each PMT channel [217].

After the discrimination step, the output signal is a rectangular pulse with only two possible values, depending on the boolean result of the mentioned comparison. This signal is the input of a *Field Programmable Gate Array* (FPGA) which converts it to a collection of digital samples at 320 MHz ( $\sim 3.125$  ns). Thus, the output of the front-end electronics is a digital trace composed by 0s and 1s [218]. The three stages of the digitalization of the analog signal are summarized in Fig. 3.25.

This *1-bit electronics* method is very robust since it does not rely on deconvoluting the number of muons from an integrated signal<sup>17</sup>. Thus, it does not strongly depend on the characteristics of the PMT, such as its gain, stability or quantum efficiency. Nor the muon impact position on the

<sup>17</sup>Additionally to the digital procedure provided by the current electronics, an upgrade that allows for the integration of the total detected signal in a module is being developed. This independent electronic channel will be particularly useful for measurements close to the EAS core.

scintillation bar and neither the corresponding light attenuation along the fiber length neither have a large systematic impact if appropriate adjustments of the PMT gain factor and discrimination thresholds are performed. However, the 1-bit electronics does rely on a fine module segmentation to prevent under counting due to simultaneous impinging muons.

Each PMT channel uses a 2048 bits ( $6.4 \mu\text{s}$ ) buffer to store digital traces. The first 512 bits are used in a loop until a T1 trigger is received from the associated WCD. At that moment, the circular buffer is frozen and the subsequent 1536 bits are used for data acquisition of the event. Once the complete buffer is filled, it is copied to an external memory<sup>18</sup>. The external memory receives the buffers from the 64 PMT channels of an MD module. It can store up to 6 ms of data, or equivalently 1024 events. If the associated WCD participates in a T3, the memory transmits the traces to the CDAS. If no T3 is found, the data is overwritten by new traces.

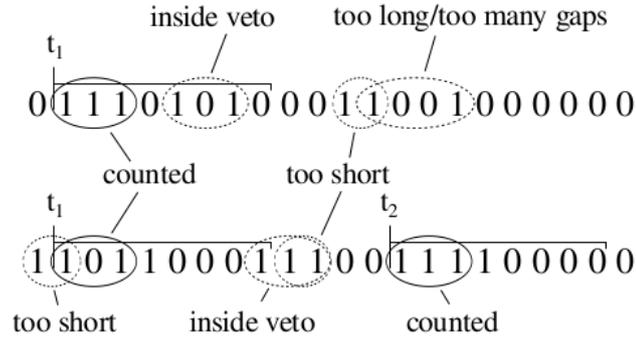
### 3.4.5 Muon counting strategy

Although the electronics of the MD facilitate the identification of pulses and thus allow for the counting of muons without detailed knowledge of signal structure and peak intensity, a certain *counting strategy* had to be implemented in order to translate the digital trace to the number of muons in a given channel.

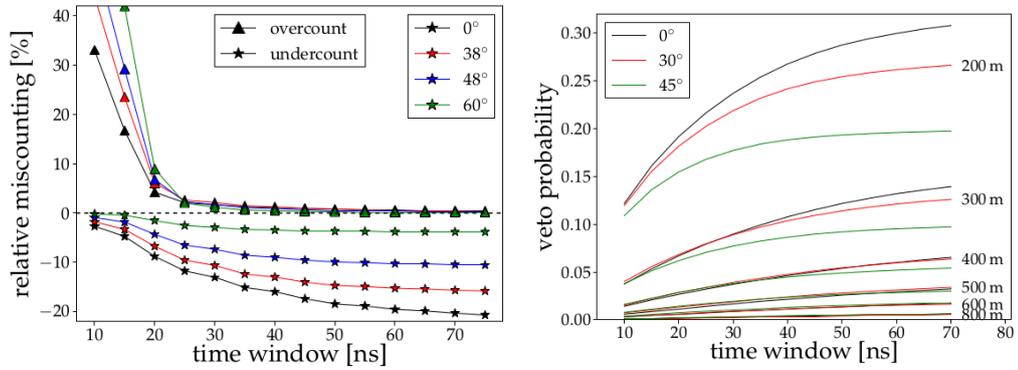
Since an analog signal of a muon pulse is composed by several single PE, the digital trace may have a succession of positive digital samples interluded by negative samples (see Fig. 3.25). In order to prevent the digital trace produced by one muon to be counted as two or more, an *inhibition time window*, over which the searching process for a muon identification pattern is stopped, has to be applied starting from the first identified pattern on the binary string. As a consequence of the amplitude discrimination, once this inhibition time window is triggered by a pattern identification, it is not possible to identify whether another muon arrives at the same scintillation bar over the time interval spanned by the window. This effect is called muon *pile-up* and may produce a systematical undercounting.

The muon counting strategy and the inhibition window must be jointly chosen as to minimize systematic effects such as over- and undercounting. Since the muon counting strategy is applied during off-line analysis, several approaches have been investigated [216]. The *gap counting strategy* is the most commonly used. It is based on the recognition of the pattern 1X1, being X a logical 0 or 1, in the digital trace of the channel. The composite concept of the gap strategy with an inhibition window is depicted in Fig. 3.26.

The gap strategy is particularly effective in suppressing false positives originating from cross-talk, since it generates “1” or “11” patterns<sup>19</sup> (see Sec. 4.2.3). The over-counting due to the electronic cross-talk has been measured to be  $\sim 4\%$  in the case of the gap strategy [209]. In general, the more stringent a strategy is with respect to overcounting, the more real muons from the far end of the scintillators are not registered correctly. An optimality criterion is not obvious because under- and overcounting have different impacts. While the latter mainly introduces systematic effects, the first discards part of the information irretrievably. The effects depending on the veto window



**Figure 3.26:** The illustration of pattern identification considering an inhibition window (veto) and the *gap* counting strategy. The solid (dashed) ovals indicate a valid (invalid) pattern. The times  $t_1$  and  $t_2$  are associated to positive matches. The horizontal lines span the inhibition window [219].



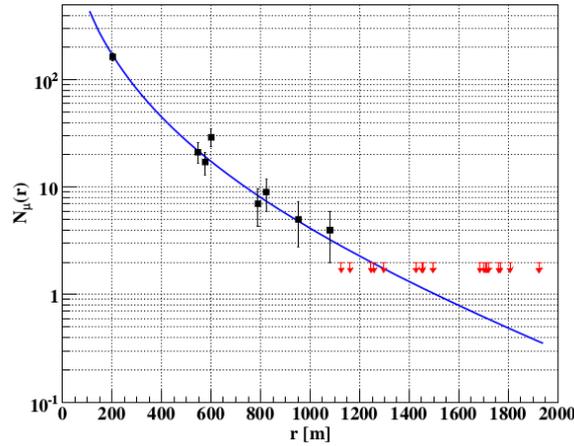
**Figure 3.27:** The influence of the inhibition window on the miscounting of several muons per channel using the gap strategy. (Left) Systematic muon miscounting classified between under- and overcounting in terms of the inhibition time window for the quoted zenith angles. (Right) Undercounting probability due to pile-up in terms of the inhibition time window for MD stations at the quoted distances to the shower axis. In both cases, the simulated showers were initiated by proton primaries at  $E_{MC} = 10^{18}$  eV [219].

are summarised in Fig. 3.27, left, for the gap strategy.

A large time window limits the dynamic range of an MD station because saturation occurs for high muon densities. The undercounting in Fig. 3.27, left, do not affect all MD stations in the same way. Indeed, the MD stations that are mostly affected by the systematic undercounting are those closer to the shower axis, as seen in Fig. 3.27, right. Motivated by the fact that a muon digital trace spans at most 30 ns (only  $\sim 8\%$  of the measured traces are longer [216]), an inhibition window of 25 ns has been found to be a good trade-off between the overcounting probability due to broad muon pulses and undercounting probability due to the arrival time distribution of high-density muons in the EAS [210]. In this case, the systematic undercounting is expected to be between 10% and 15%, as shown in Fig. 3.27, left.

<sup>18</sup>While the trace is being copied to the external memory, an auxiliary buffer is used in order to continue with the acquisition of the event.

<sup>19</sup>Most cross-talk pulses originate from one SPE, which extends on average 3.5 ns in time. Given the FPGA sampling of 3.125 ns, the expected pattern would be “1” or, at most, “11”.



**Figure 3.28:** The fitted MLDF for a simulated proton event with  $E_{MC} = 10^{18}$  eV and  $\theta = 30^\circ$ . The MD stations considered in this event have a nominal area of  $30 \text{ m}^2$ . The red arrows correspond to the MD station that registered less than two muons [221].

### 3.4.6 Muon LDF reconstruction

The MD relies on the SD for the event reconstruction, i.e., for triggering and for both geometry and energy reconstruction. Therefore, the information from the MD stations may be summarized in a single observable reflecting the composition of the primary particle. If an event has enough MD information (i.e. triggered MD stations), a LDF ansatz according to Eq. 3.13 may be fitted to the data points, with  $r_{\text{scale}} = 120 \text{ m}$ ,  $\alpha = 1$ ,  $\beta = 2.8 - \sec\theta$  and  $\gamma = 1.85$ . It was first used to analyze muon data in the KASCADE-Grande experiment [220] and then applied to the MD [221].

$$S(r) = N_\mu \left( \frac{r}{r_{\text{scale}}} \right)^{-\alpha} \left( 1 + \frac{r}{r_{\text{scale}}} \right)^{-\beta} \left( 1 + \left( \frac{r}{10 \cdot r_{\text{scale}}} \right)^2 \right)^{-\gamma} \quad (3.13)$$

The parameters were found empirically and reflect the typical shape of a MLDF above  $10^{17.5}$  eV. The shower size  $N_\mu$  is always fitted to data using a profile likelihood approach [222]. An example of a fitted MLDF for a simulated event is shown in Fig. 3.28.

In analogy to the SD, an expected signal at a reference distance can be used as a composition-dependent observable. However, as it is explained in Chapter 5, a different approach is developed throughout this thesis in order to discriminate photon and hadronic primaries.

## 3.5 Upgrades of the Pierre Auger Observatory

### 3.5.1 Current enhancements

**AERA** The FD detects ultraviolet light emitted by nitrogen molecules that are excited by EAS in the Earth's atmosphere. Another channel that can be exploited for the cosmic-ray detection is based on geomagnetic separation of electrons and positrons of an EAS. As a result, an EAS emits coherent radio pulses which can be measured by a radio-antenna array. It was proven some years ago that it is experimentally feasible and promising to study the emission of air showers in the MHz-regime [223]. Recently, it was shown by the LOFAR Collaboration that the measurement of the radio footprint allows the reconstruction of the energy of the primary particle as well as quantities that are related to its mass, like the depth of shower maximum or the shower curvature



**Figure 3.29:** A radio station prototype. The log-periodic dipole antenna and solar panels are visible. The electronics and digitizer are placed in the aluminum box beneath it [225].

[224].

The *Auger Engineering Radio Array* (AERA) is built within the SD-750 array and covers an area of 20 km<sup>2</sup>. 150 radio stations arranged in grids with different spacings enable a study of the radio emission of EAS together with measurements from SD, FD and MD in a multi-hybrid approach. One prototype radio station is shown in Fig. 3.29. Some first multi-hybrid events exist and are currently studied. With a duty cycle of nearly 100%, low costs and definite possibilities to reconstruct primary properties, the radio emission is a very promising candidate for the future composition measurements at ultra-high energies.

**HEAT** The field of view of the base design of the FD is limited to 30° above the horizon (see Sec. 3.2). At close distances only the lowest few kilometres of the atmosphere are within the field of view. However, low energy showers reach their maximum of development at higher altitudes. Thus, the crucial region around the shower maximum is generally not observed by the original FD telescopes.

In order to observe these faint low-energy showers, the FD was upgraded by the installation of three high-elevation telescopes. These telescopes are located 180 m north-east of the Coihueco FD site and work independently of the other FD sites. The HEAT telescopes were designed similarly as the original FD system, except for the possibility to tilt the telescopes upwards by 29°, covering the elevation range from 30° to 58°. The building housing the HEAT telescopes is shown in Fig. 3.30.

In combination with the information from the SD-750 close to the HEAT site, the energy range of high-quality hybrid EAS measurements has been extended down to 10<sup>17</sup> eV since May 2010 [227].

### 3.5.2 AugerPrime

The Pierre Auger Observatory is planned to be operated until the end of 2025. *AugerPrime* is a major upgrade to many aspects of the current design of the Observatory. Its main motivation is to provide additional measurements in order to address the following important questions [228]:



**Figure 3.30:** A photo of the HEAT telescopes tilted upwards. In the background, the telecommunication tower of Coihueco is visible [226].

- Elucidate the mass composition and the origin of the flux suppression at the highest energies, i.e. the differentiation between the energy loss effects due to propagation of the cosmic rays through the intergalactic medium, and the maximum energy of particles injected by astrophysical sources.
- Search for a flux contribution of protons up to the highest energies with a sensitivity to a contribution as small as 10% in the flux suppression region. The measurement of the fraction of protons is the decisive ingredient for estimating the physics potential of existing and future cosmic ray, neutrino, and gamma-ray detectors; thus prospects for proton Astronomy with future detectors will be clarified. Moreover, the flux of secondary gamma-rays and neutrinos due to proton energy loss processes will be predicted.
- Study of EAS and hadronic multiparticle production. This will include the exploration of fundamental particle physics at energies beyond those accessible at man-made accelerators, and the derivation of constraints on new physics phenomena, such as Lorentz invariance violation or extra dimensions.

The addition of new detectors and the enhancement of existing ones will provide additional composition sensitive information that will help to better reconstruct the properties of the primary particles at the highest energies and improve the measurements in the important energy range just above the ankle. The improved knowledge will likely then also allow a re-analysis of existing data, for better energy assignments, mass composition studies, and for photon and neutrino searches.

The AugerPrime upgrade consists of four main improvements shortly described in the following.

**Scintillator Surface Detector** A complementary measurement of the shower particles will be provided by a plastic scintillator plane above the existing WCDs. This allows the shower particles with two detectors having different responses to muons and electromagnetic particles to be sampled. The design of the *Scintillator Surface Detector* (SSD) is simple, reliable and they can be easily deployed over the full 3000 km<sup>2</sup> area of the SD.

The SSD stations consist in one module of  $\sim 4 \text{ m}^2$  extruded plastic scintillators which are read out by WLS fibers coupled to a single photo detector. The active part of the scintillator is a plane made by 12 extruded polystyrene scintillation bars. Each bar is 1.5 m long, 1 cm thick and



**Figure 3.31:** One SSD module already installed in the Pierre Auger Observatory [228].

10 cm wide. The fibers are positioned following the grooves of the routers at both ends, in a 'U' configuration that maximizes light yield and allows the use of a single photomultiplier. A deployed SSD station is shown in Fig. 3.31. The external detector enclosure is made from aluminum to guarantee light tightness, robustness for 10 years of operation in the field and enough rigidity for transportation.

**Surface Detector Electronics Upgrade** The current SD electronics was designed 15 years ago using the technology available at that time. Evolution in processors, power consumption of electronics components, and timing systems make it possible to design and implement a higher performance electronics system for the SD arrays. Use of the new electronics also aims to increase the data quality (with faster sampling of ADC traces, better timing accuracy, increased dynamic range), to enhance the local trigger and processing capabilities (with a more powerful local station processor and FPGA) and to improve calibration and monitoring capabilities of the SD stations. Furthermore, the proposed electronics provides an interface to allow the SSD stations co-located with the WCDs to make use of the data processing and communications infrastructure of the stations. The *Surface Detector Electronics Upgrade* (SDEU) can be easily deployed, and will have only minimal impact on the continuous data taking of the SD.

**Underground Muon Detector** In light of the Auger upgrade, the plan is to finish the deployment of AMIGA MD stations in the whole SD-750. The *Underground Muon Detector* (UMD) will provide important direct measurements of the shower muon content and its time structure, while serving as verification and fine-tuning of the methods used to extract muon information with the SSD and WCD measurements. The performance and characteristics of the EA of AMIGA MD match these requirements, and thus the completed AMIGA array will serve as the UMD.

The UC served as test bench to analyze the performance of the whole detection system in several aspects: mechanics, casing and deployment; electronics and communication; front-end and monitoring software; and data acquisition. The EA data over the past years have motivated several improvements and upgrades to the original MD design:

- The central dome where the electronics is housed has been re-designed. The PVC casing of the scintillation modules will be folded at the edge of the modules, instead of glueing extra pieces of PVC as in the base design. The original 1.3 m diameter access tube will be replaced by a 30 cm diameter PVC tube, so that the electronics would be replaced with a specialized

probe.

- One MD station from the UC is currently equipped with *silicon photo-multipliers* (SiPMs). These solid state devices show both a much cheaper price compared to PMTs and better photon detection efficiency around 450 nm, the maximum light emission wavelength of the fibres used in the UC. Other advantages are the lower power consumption and the avoidance of cross-talk. Therefore, it is foreseen that the PMTs will be replaced with SiPMs in the UMD design.
- The MD stations will have a dedicated integrator channel, in addition to the standard 64 electronic channels from the base design. This integrator was designed to measure the muon content of the EAS near the shower axis, where the high muon density does not permit to obtain an unbiased estimation of the muon number according to the standard counting procedure from the base design.
- The monitoring system of AMIGA has been fully integrated within the Observatory system since May 2014. It delivers graphical as well as exportable data of several parameters of the scintillator modules both from the real-time monitoring in the field and from the laboratory tests.

**Low-gain mode for the FD** In parallel with the SDEU, the FD operation will be extended to times at which a larger fraction of the moon is present in the sky. In such conditions, the PMTs gain must be reduced by lowering the supplied high voltage to avoid high anode current and, therefore, an irreversible deterioration of the PMTs sensitivity. This will allow an increase of about 50% in the current duty cycle of the FD.

The scientific goals of AugerPrime imply the thorough study of the mass composition of the cosmic rays and the hadronic interactions intervening in the EAS development, which are intimately interwoven. One of the most promising ways to solve these open problems is the measurement and characterization of the muonic component of the air showers.

## Chapter 4

# Simulation of the detector responses

I had rather be Mercury, the smallest among seven [planets], revolving round the Sun, than the first among five [moons] revolving round Saturn.

---

Johann W. von Goethe

The longitudinal development and the footprint on ground of the EAS can be registered with the fluorescence and surface detectors of the Pierre Auger Observatory. The features of the primary cosmic ray, such as its energy and arrival direction, are reconstructed from the raw data gathered in CDAS (see Sec. 3.3.3). Additionally, the muon content of an EAS may be characterized with the MD, given that its impact point is sufficiently close to a MD station (see Sec. 3.4).

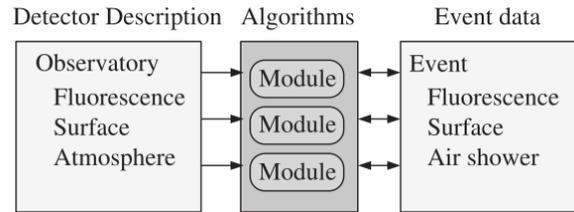
The procedures developed to reconstruct raw data were described in Sec. 3.2.4, Sec. 3.3.4 and Sec. 3.4.5. They are implemented within an evolving C++ collaborative framework called *AugerOffline* (Offline). In a similar way, procedures to simulate the responses of the different elements of the detectors have been programmed and are continuously upgraded. Therefore, the response of the detectors to an EAS with known characteristics can be also studied from a full Monte Carlo approach.

Several calibrations and parametrizations intervening in the reconstruction procedure are already included in Offline, they are particularly tuned to accurately describe data, which are known to be hadronic cosmic rays. The same code machinery can be used to reconstruct photon-initiated EAS, but the low muon content in these events produce several side effects and artifacts in the reconstructed observables.

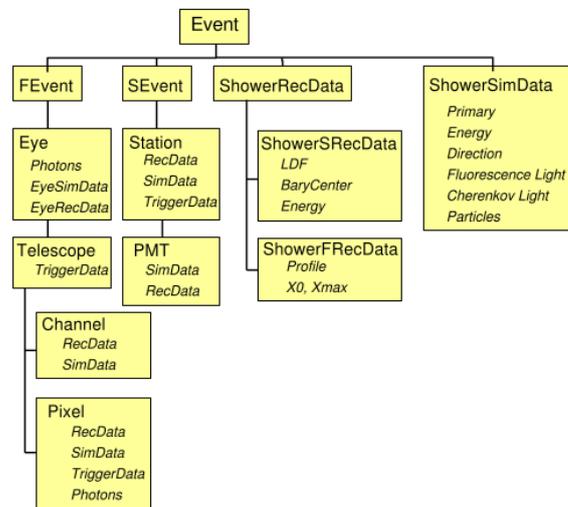
In this chapter, the performance of the reconstruction procedure to photon primaries, compared to the case of proton primaries, is extensively described, both from the point of view of the SD and the MD.

### 4.1 AugerOffline framework

The Offline Framework comprises three principal parts: a collection of processing *modules* which can be assembled and sequenced through instructions provided in an XML file, an *event data model* through which modules can relay data to one another and which accumulates all simulation



**Figure 4.1:** The general structure of the Offline framework. Simulation and reconstruction tasks are broken down into *modules*. Each module is able to read information from the *detector description* and/or the *event*, process the information, and write the results back into the event. Image taken from [229].



**Figure 4.2:** The hierarchy of the event interface. The top level *Event* encapsulates objects representing each detection system (here, only FD and SD are represented by *FEvent* and *SEvent*, respectively), as well as reconstructed and simulated shower data (*ShowerRecData* and *ShowerSimData*, respectively). These components are further subdivided into objects representing simulated, reconstructed and triggering data at the level of individual detector stations, and further into electronics elements. Image taken from [229].

and reconstruction information, and a *detector description* that contains static or relatively slowly varying information such as detector geometry, calibration constants and atmospheric conditions [229]. These ingredients are depicted in Fig. 4.1. These components are complemented by a set of foundation classes and utilities for error logging, physics and mathematical manipulation.

The event data model contains all raw, calibrated, reconstructed and MC data and acts as the principal backbone for communication between modules. The overall structure comprises a collection of classes organized following the hierarchy normally associated with the detection systems, with further subdivisions for accessing such information as reconstructed quantities, calibration information and raw data. A simplified illustration of this hierarchy is given in Fig. 4.2.

Primary	Energies (eV)	Energy distribution	Zenith angles (°)	Zenith distribution
1000x Photon	$10^{16.5} - 10^{17}$	$\sim E^{-1}$	$0^\circ, 30^\circ, 40^\circ$	Fixed angles
1000x Photon	$10^{17} - 10^{18}$	$\sim E^{-1}$	$0^\circ, 30^\circ, 40^\circ$	Fixed angles
2000x Photon	$10^{16} - 10^{18}$	Uniform in E	$0^\circ-45^\circ$	$\sin\theta \cos\theta$
1000x Proton	$10^{16.5} - 10^{17.5}$	$\sim E^{-1}$	$0^\circ, 30^\circ, 40^\circ$	Fixed angles
3000x Proton	$10^{16} - 10^{18}$	Uniform in E	$0^\circ-45^\circ$	$\sin\theta \cos\theta$

**Table 4.1:** The simulated shower libraries based on QGSJetII-04 that are used for this thesis.

## 4.2 Event simulation within AugerOffline

### 4.2.1 Production of EAS

An event simulation starts with an EAS with known characteristics, such as the primary particle type, its energy and zenith angle. These are referred throughout this thesis as *simulated* or *MC* characteristics. In order to have a wide variety of primary energies and zenith angles at disposal<sup>1</sup>, libraries of EAS initiated by photons and protons were produced. No other primary was simulated, since the EAS initiated by protons are the most similar to the electromagnetic-dominated EAS produced by ultra-high energy photons. Thus, the discrimination procedures are developed in this thesis in the most disfavoured scenario, at least from the point of view of the type of the primary cosmic ray.

Also, depending on the specific analysis, the primary energy and/or zenith angle may be fixed or following a continuous distribution. The energies of interest for this thesis, as mentioned before, range between  $10^{16.5}$  eV to  $10^{18}$  eV. The lower limit is chosen in agreement with the estimated trigger capabilities of the SD-433 in previous studies [212], whereas the upper limit is the lowest primary energy that current Auger photon searches can reach. The zenith angles range between  $0^\circ$  and  $45^\circ$ , the upper limit being imposed by the technical design of the MD stations (see Sec. 3.4.1). The shower libraries available for this study are specified in Tab. 4.1.

As already uttered before, the most popular software package to simulate EAS is the *COsmic Ray Simulations for KAscade* (CORSIKA) [230]. In particular, the version 7.4950 was employed to produce all the showers used in this thesis. The particle interactions are simulated with a high- and low-energy hadronic-interaction models<sup>2</sup>, depending if the energy in the center-of-mass frame is below or above 100 GeV. Several alternatives for hadronic models are available at compilation time. Each model makes different assumptions about the particle interactions or employs different techniques to perform it. As already seen in Fig. 2.7 and Fig. 2.8, the choice of the hadronic-interaction model (specially in the high-energy regime) introduces an intrinsic systematic uncertainties [232]. In Sec. 6.5, the impact of the choice of hadronic-interaction models on the muon content of simulated EAS, specially in the framework of this photon search, is discussed. Throughout this thesis, QGSJetII-04 [233] and FLUKA [234] are chosen as the hadronic interactions models for the high- and low-energy regimes, respectively.

The number of particles of an EAS increases with the energy of the primary cosmic ray. Therefore, at energies above  $\sim 10^{16}$  eV, a *thinning* method is employed to reduce both computation time and disk space. After each hadronic interaction, only one of the secondary particles that fall below the energy threshold of  $\varepsilon_{th} \times E_0$  continues in the shower development<sup>3</sup> and the rest of

<sup>1</sup>The azimuth angle is randomly selected between  $0^\circ$  and  $360^\circ$ . Thus, processes affecting the particles trajectories, such as the geomagnetic field, are averaged through several realizations of showers with the same primary features.

<sup>2</sup>The electromagnetic interactions are described by the EGS4 package [231].

<sup>3</sup>A particle has a probability of surviving the thinning method proportional to its energy.

them are discarded.  $\varepsilon_{th}$  is the *thinning level* and  $E_0$  is the primary energy, both included as input parameters of a CORSIKA run. The surviving particle is given a statistical weight such that the total energy is conserved. The showers used in this thesis share a thinning level of  $10^{-6}$ , which is the most common choice.

A large variety of information is available from CORSIKA, ranging from the longitudinal development of the shower (i.e. number of particles, energy per EAS component, energy deposit in the atmosphere, etc) to the list of particles arriving at a given observation level. By setting it to 1452 m ( $878 \text{ g cm}^{-2}$  vertical depth), corresponding to the mean altitude of the Auger location, a reliable simulation of the SD and MD responses can be performed<sup>4</sup>.

## 4.2.2 Monte Carlo procedure

The detector response is simulated in Offline v3r3p3 revision 31269. The WCDs are arranged in a hexagonal grid with a spacing of 433 m as the SD-433 array. 61 WCDs are considered for the event simulation. The events are reconstructed using only the 13 WCDs reflecting the real positioning of the SD-433 on the field, as shown in Fig. 3.17.

Similar to the SD array, an array of 61 MD stations, each one with the same module layout, is simulated. The  $60 \text{ m}^2$  MD station with Id 1764 (named *Kathy Turner* or KT), which is located at the center of the SD-433 hexagon, is taken as the reference layout since it is one of the cornerstones of the photon search of this thesis, as it is explained in Chapter 5. However, only the four MD stations located in the positions corresponding to the ones deployed in the field, i.e. one at the central position of the SD-433 hexagon and three in the second-crown (see Fig. 3.17), are considered for the reconstruction<sup>5</sup>. In Fig. 4.3, the ideal SD+MD layout used for the simulation and the realistic layout for the reconstruction are depicted.

The flexible detector layout used at the reconstruction phase is offered by a dedicated Offline module called *PositionRejector* (see Appendix A) written specifically for this thesis<sup>6</sup>. It is suited to take into account certain WCD positions, MD modules or even channels chosen by the user to perform the event reconstruction and the muon counting. The added flexibility is desired in order to study specific systematic effects due to low number of available stations, to estimate observables at a single station level with high statistics or to design parameterizations in ideal scenarios without introducing artifacts from the array incompleteness.

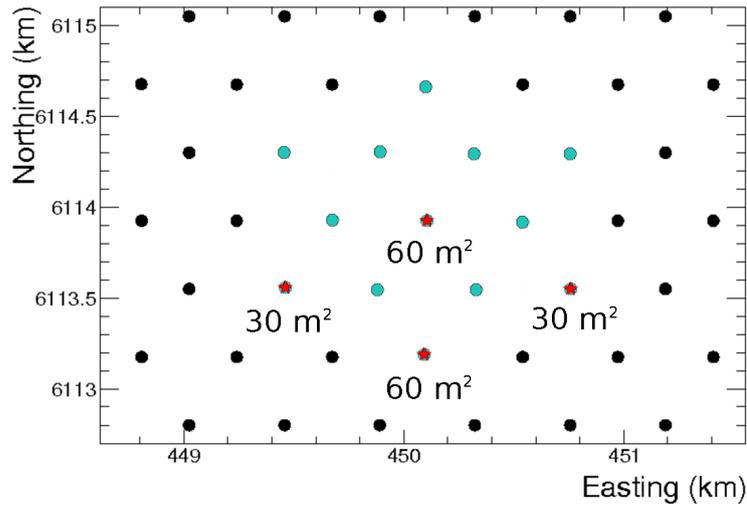
At the time of writing, this thesis is one of the few MC studies related to the SD-433. Therefore, a functionality regarding the array spacing has been added to the *CentralTriggerSimulatorXb* Offline module, which is responsible for the arrangement of the T4 configuration in an event. The module itself has been minimally changed and the extra features have been included via optional XML configuration. The change has been committed to the official Offline repository in revision 30941.

The standard Offline simulation/reconstruction pipeline for SD+MD is employed for the event production [235]. In Appendix A the Offline module sequences for both phases are shown. The SD and MD stations are simulated outside a region of 50 m radius around the shower core to avoid

<sup>4</sup>Although the Pampa Amarilla location is not perfectly flat, the maximum altitude difference between any pair of WCD in Auger is in the order of a few tens of meters and its systematic impact is negligible.

<sup>5</sup>Two of the four MD stations have a detection area of  $30 \text{ m}^2$ . This is also considered in the reconstruction.

<sup>6</sup>This Offline module is foreseen to be made public soon.



**Figure 4.3:** Schematic view of the detector layout for the event simulation (black dots) and for the realistic event reconstruction (cyan dots). The central hexagon represents the SD-433 array and thus the MD stations (red stars) resemble the field configuration respective to their positions and quoted sensitive areas.

long usage of computational resources due to the large particle densities<sup>7</sup>.

Each shower is used ten times to generate an event well-contained in the central hexagon of the SD array. The shower cores are placed randomly in the elemental hexagon (i.e. the first Brillouin zone of the hexagonal lattice) around the central WCD<sup>8</sup>.

### 4.2.3 Tuning the MD simulation

The foundations of the MD simulation pipeline are well-established [216]. However, some important contributions have been done in order to carry out the analyses presented in this thesis. A thorough study on the characteristics of SPE pulses at the PMT anode and after the front-end electronics of the MD module have been performed [236]. This is particularly important since these parameters, measured under laboratory conditions, are then implemented in the MD simulation to model the SPE pulses as Gaussian functions. A brief summary of the main results of that study is presented in this section.

The width<sup>9</sup> and the amplitude of the SPE pulses at the output of six PMTs<sup>10</sup> have been investigated in a dedicated testing facility [236]. In Fig. 4.4, top, the width of SPE pulses for a single pixel are shown. The first bin corresponds to the baseline of the signals (measurements with low signal-to-noise ratio) and have not been considered in the analysis. Approximately  $10^4$  measurements for each pixel were obtained. In Fig. 4.4, bottom, peak measurements for the same pixel are displayed.

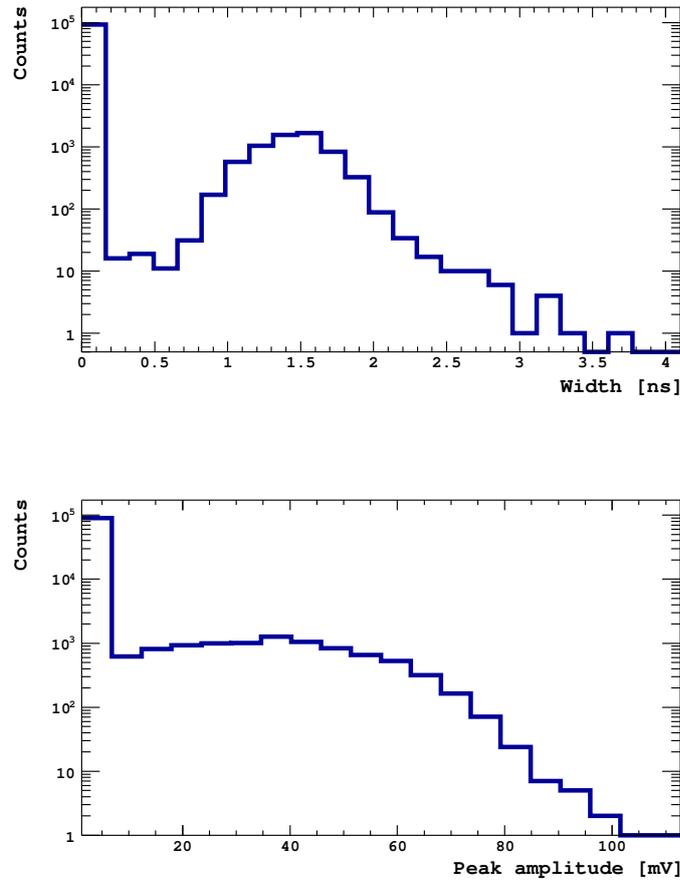
The distributions from each pixel are used to calculate the average and sample variance for

<sup>7</sup>In any case, hadronic models are currently not capable of describing the EAS phenomena in the shower core in an accurate manner and also the core reconstruction uncertainty is within 30 – 40 m.

<sup>8</sup>For some specific analyses, the shower cores are placed in a greater area, for example in a rectangle containing the whole central hexagon. Nevertheless, whenever the core distribution is changed, it is explicitly mentioned.

<sup>9</sup>The width of the SPE pulses is defined by the full-width half-maximum criterium.

<sup>10</sup>The models of the PMTs as well as the electronics board used in this analyses resemble the ones installed in the MD Engineering Array.



**Figure 4.4:** The width (top) and peak amplitude (bottom) of SPE pulses at the output of a single channel of a PMT.

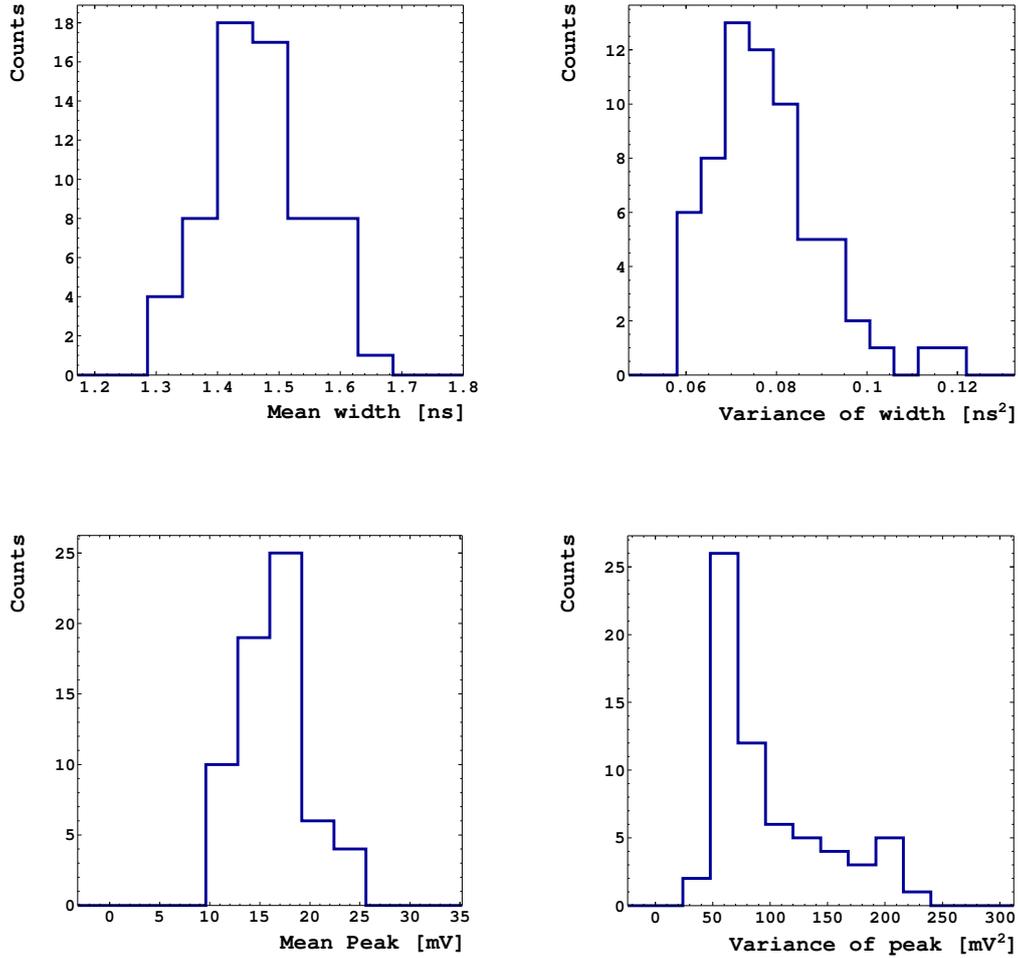
each observable (width and amplitude of the SPE pulses). The 64 mean values and their variances are presented in Fig. 4.5 for a single PMT are shown. There are intrinsic fluctuations in the width among different pixels mainly due to variations in the gain factor. Lastly, the mean SPE width and amplitude are calculated by averaging over the six PMTs. The results are summarized in Tab. 4.2.

The SPE pulses are modelled by Gaussian functions in the MD simulation chain, since the SPE pulses have a qualitative resemblance to it (see Fig. 3.24, left). Having found the characteristic values which describe the distributions of widths and amplitudes of SPE pulses at the output of the PMT, a set of  $\sim 10^4$  pulses have been simulated<sup>11</sup>. On the other hand, an independent set of  $10^4$  SPE signals have been measured in the laboratory. Both sets are then processed according to the MD electronics simulation sequence [235]. Therefore, the responses of the front-end electronics

<sup>11</sup>For each SPE pulse, a random width and amplitude is chosen according to the parameters of Tab. 4.2.

Mean width (ns)	Std. dev. of width (ns)	Mean peak amplitude (mV)	Std. dev. of peak (mV)
$(1.53 \pm 0.01)$	$(0.28 \pm 0.01)$	$(35.1 \pm 0.8)$	$(16.4 \pm 0.3)$

**Table 4.2:** The width and peak amplitude (mean and standard deviation) of SPE pulses at the output of a PMT. The values are obtained by averaging over 64 channels of six PMTs as used in the MD Engineering Array.



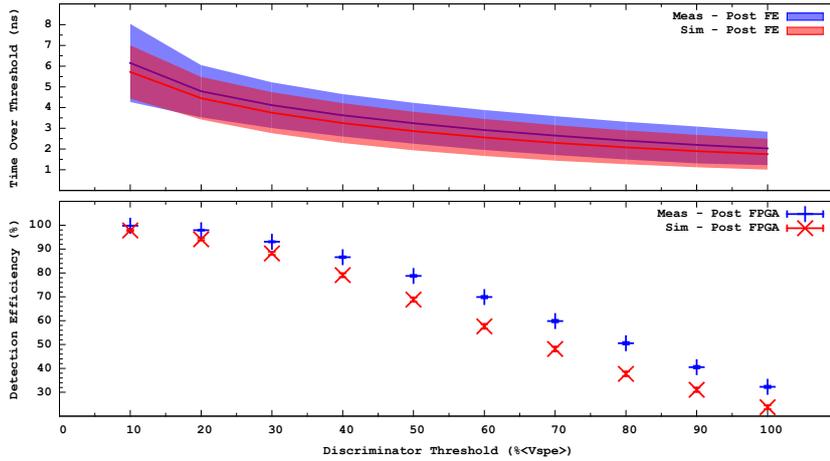
**Figure 4.5:** Top row: The average width (left) and its variance (right) for the 64 pixels of a single PMT. The means of the histograms are  $(1.47 \pm 0.01)$  ns and  $(0.078 \pm 0.002)$  ns<sup>2</sup> respectively. Bottom row: the mean peak amplitude (left) and its variance (right) for 64 pixels of a single PMT. The means of the histograms are  $(16.24 \pm 0.44)$  mV and  $(96.95 \pm 6.26)$  mV<sup>2</sup> respectively.

to the modelled and real SPE pulses can be compared.

The time-over-threshold (ToT) is defined, for the sake of this particular analysis, as the elapsed time while the pulse exceeds the discriminating threshold. In Fig. 4.6, top, the mean ToT values and their standard deviations are shown for each discrimination threshold, defined as a certain fraction of the mean SPE amplitude after the amplification stage of the front-end electronics  $\langle V_{\text{SPE}} \rangle$ <sup>12</sup>. Both sets are in good agreement through the entire range of thresholds. The ToT obtained from simulated pulses is lower than for the measured set. The difference ranges between 7% to 13% in the entire range which is compatible with the expected fluctuations between PMTs [236].

For the set of simulated SPEs, the mean ToT equals the FPGA sampling rate (3.125 ns) at  $\sim 50\% \langle V_{\text{SPE}} \rangle$ , while for the measured set this happens at  $\sim 65\% \langle V_{\text{SPE}} \rangle$ . The probability of obtaining at least a digital 1 from the FPGA is defined as *detection efficiency*. As it is shown in Fig. 4.6, bottom, the difference in ToT is translated in different detection efficiencies depending on the

<sup>12</sup>Note that this value is calculated separately for the simulated and real SPE pulses sample.



**Figure 4.6:** (Top) The Time-over-Threshold for different discrimination levels. The shaded areas correspond to the  $1\sigma$  region. (Bottom) The detection probability, defined as the ratio of SPEs producing at least a positive FPGA sample to the total number of SPEs considered. The blue (red) areas and markers correspond to measured (simulated) SPE pulses.

threshold. The lower ToT of the simulated SPEs produces a lower detection probability than for the measured set. This relative difference is  $\lesssim 10\%$  up to  $50\% \langle V_{\text{SPE}} \rangle$ . The current discrimination threshold for the MD stations is set at  $30\% \langle V_{\text{SPE}} \rangle$ . Therefore, the Gaussian model constitute a very good approximation to the real SPEs shape. The disagreement starts to be significant at higher threshold ( $\gtrsim 20\%$  at discrimination thresholds above  $70\% \langle V_{\text{SPE}} \rangle$ ) since the measured SPEs are wider than the simulated set.

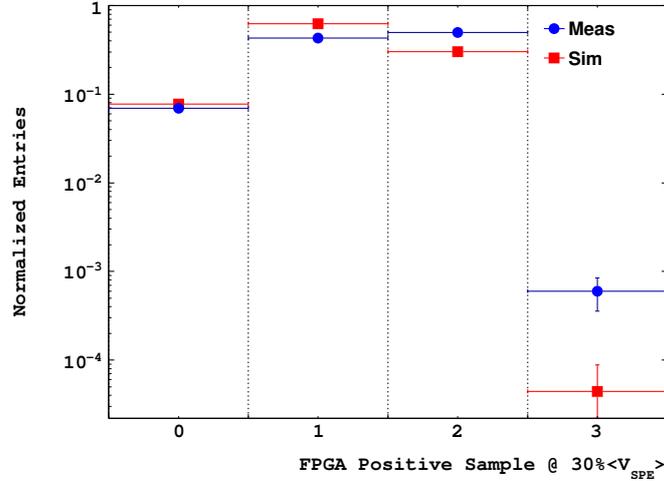
The detection probability rises the question of how many FPGA positive samples can a SPE pulse generate. When setting a threshold at  $30\% \langle V_{\text{SPE}} \rangle$  the number of positive FPGA samples is shown in Fig. 4.7 for the measured and simulated sets. The probability of getting one or two digital samples is  $\gtrsim 92\%$  for both sets. It is worth mentioning that all the cases with two digital samples correspond to the pattern 11, i.e. the samples are always consecutive both in real and simulated pulses. The probability of getting three or more digital samples, and thus counting an SPE pulse as a muon (see Sec. 3.4.5) is  $< 0.1\%$ .

In summary, this study permitted an accurate estimation of the parameters describing the SPE as Gaussian pulses and used to define the gap counting strategy, which have been incorporated in the MD simulation chain and used for all the MC studies of this thesis.

### 4.3 WCD and array triggers

Following the rationale behind the data acquisition, the first aspect to be analysed is the WCD triggers followed by the array trigger efficiency.

As mentioned before, the ToTd and MoPS trigger algorithms are included in this analysis. They provide a substantial improvement to the WCD trigger probability in the case of photon showers due to their high sensitivity to electromagnetic-dominated traces and they were specially designed for  $E > 10^{17.5}$  eV [237]. Nearly half of the triggered stations are recovered with these algorithms independent on the primary energy, as it can be seen in Fig. 4.8 for the case of vertical



**Figure 4.7:** The number of FPGA positive samples at a threshold of  $30\% \langle V_{SPE} \rangle$  for measured (blue circles) and simulated (red squares) SPEs.

incidence. The dependence does not change remarkably with the zenith angle. The WCDs triggered exclusively with the ToTd and MoPS triggers, jointly called *new triggers* for historical reasons, are effective at larger distances, as seen in Fig. 4.9, left. On the contrary, the stations close to the core usually trigger with the ToT and Th algorithms, as shown in Fig. 4.9, right. The same behaviour is seen for photon primaries.

As it is discussed in the next sections, the triggered WCDs at the energies of interest are located mainly in the central hexagon. Thus, events satisfying the 6T5 condition would have most of the triggered stations located within a radius of  $\sim 800$  m from the shower core. Therefore, the new triggers are not particularly useful for the energies of interest, since most of the stations close enough to the shower core are triggered with the Th or ToT algorithms.

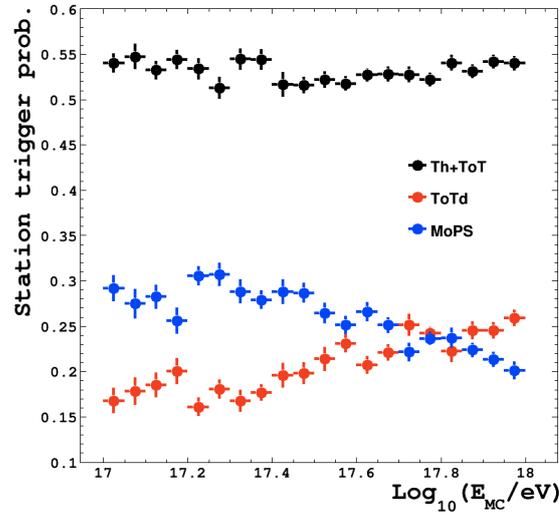
### 4.3.1 Array trigger efficiency

The size of the footprint of particles that an EAS imprints on the ground decreases with the primary energy. Therefore, if the stations are deployed sufficiently close, a suitable number of triggered WCDs may be obtained to perform the event reconstruction. However, there is a certain primary energy threshold below which the number of triggered WCDs falls below 3, i.e. the minimum required number to build the first array trigger T3 (see Sec. 3.3.2). This intrinsic limitation is referred, from now on, as *trigger efficiency* and it is described in this section.

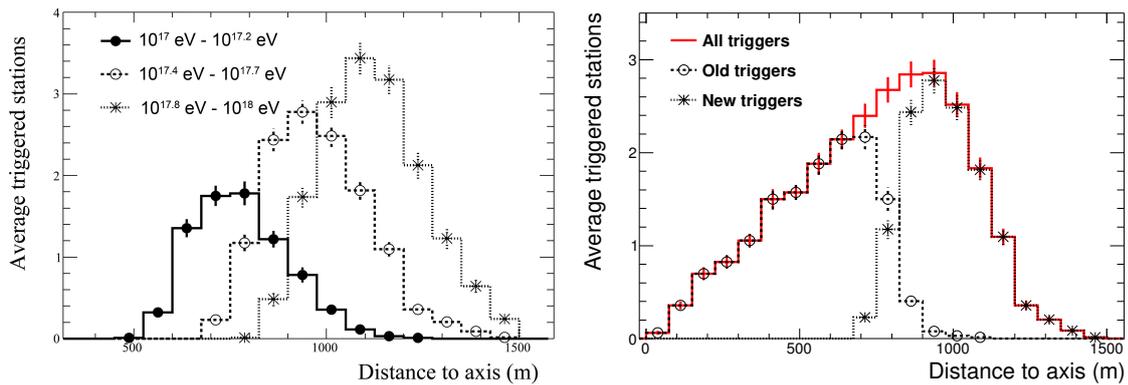
The trigger efficiency is defined as:

$$\varepsilon(E, \theta) = \frac{N^{\text{trig}}(E, \theta)}{N^{\text{tot}}(E, \theta)} \quad (4.1)$$

where  $N^{\text{trig}}$  is the number of events generating an array trigger in a sample of  $N^{\text{tot}}$  events at a fixed energy  $E$  and zenith angle  $\theta$ . By construction,  $\varepsilon$  takes values between 0 and 1. Customarily an array is considered as “fully efficient” if  $\varepsilon \geq 0.95$ .



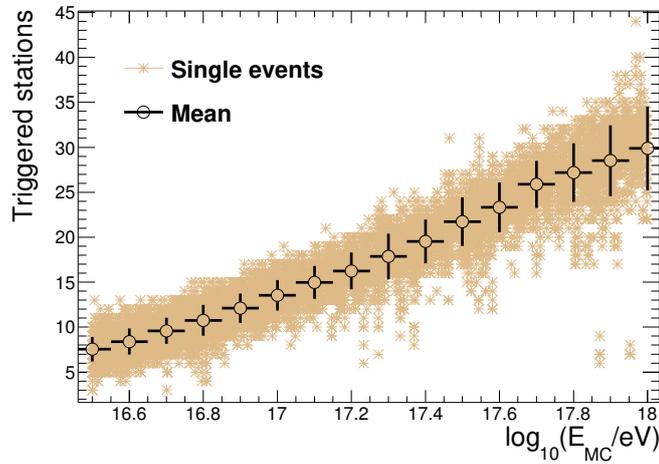
**Figure 4.8:** The ratio of WCDs triggered with the quoted algorithms to the total number of triggered WCDs in an event-by-event basis. The subsets corresponding to the ToTd (red dots) and MoPS (blue dots) algorithms include only WCD triggered with the corresponding algorithm and not with Th or ToT (black dots). The events are generated by primary protons impinging with  $\theta_{MC} = 0^\circ$ .



**Figure 4.9:** (Left) The average number of triggered stations per event with the ToTd and MoPS algorithms in terms of the distance to shower axis for three different energy ranges (marker styles). (Right) The average number of triggered stations per event with Th or ToT algorithms and exclusively with the ToTd or MoPS algorithms for  $E_{MC} = 10^{17.5}$  eV. Stations satisfying both trigger conditions are included only in the old triggers subset. In both panels, the events are generated by primary protons with  $\theta_{MC} = 30^\circ$ .

The trigger efficiency depends not only on the primary energy and zenithal incidence, but on its mass composition. Although there is not a large difference in the number of electromagnetic particles of EAS initiated by photons and protons, the lack of muons in the former case tends to increase the minimum required energy for an array to be fully efficient. Therefore, it is instructive to study the trigger efficiency in the case of photon and proton primaries separately.

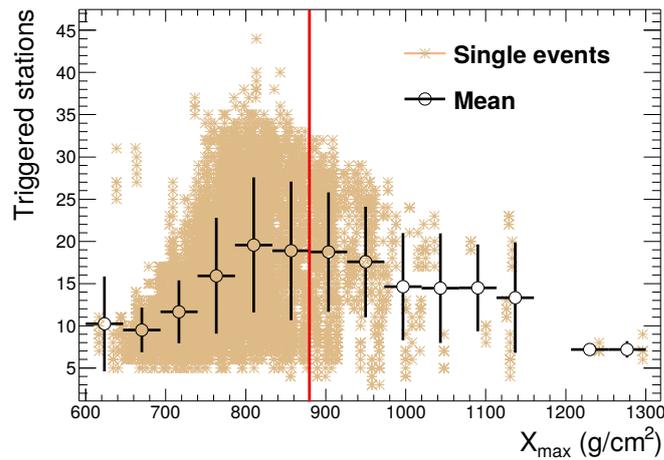
**Photon primaries** The vertical photon showers with energies above  $E_{MC} = 10^{16.5}$  eV trigger the SD array and verify the T5 selection criterion with a probability above 98.9%. The showers responsible for the events without a T5 trigger have energies between  $E_{MC} = 10^{16.5}$  eV and  $10^{16.6}$  eV. This can be partially inferred from the number of triggered stations in Fig. 4.10. The events with the lowest MC energies have an average of  $\sim 7$  triggered stations with a standard deviation of 1.36. The mean number of triggered stations increases with the primary energy, as expected.



**Figure 4.10:** The number of triggered SD stations in terms of the primary energy for vertical photon primaries. The beige crosses represent the single events while the empty black markers represent the mean values with the  $1\sigma$  deviation. The outlier events with  $E_{MC} \gtrsim 10^{17.8}$  eV and less than 10 triggered stations are different realisations of deeply penetrating showers with  $X_{\max} > 1200 \text{ g cm}^{-2}$  (see text for details).

However, there is a small population of events with an unusual low number of triggered stations at the highest energies. For example,  $\sim 2\%$  of the photon events with  $E_{MC} > 10^{17.8}$  eV have less than 10 triggered stations, while an average of  $\sim 28$  stations are expected to trigger at these energies. This is understood in terms of the shower development: since vertical showers transverse less matter, they reach the observation level ( $878 \text{ g cm}^{-2}$ ) at an earlier stage of development. Therefore, their particles are strongly boosted around the shower axis, which in turn may trigger a low number of stations. In particular, nearly a third of the showers arrive at ground without reaching the maximum development at  $E_{MC} \gtrsim 10^{17.8}$  eV, qualitatively in agreement with previous studies [238]. This behaviour is depicted in Fig. 4.11. In particular, the events with  $E_{MC} \gtrsim 10^{17.8}$  eV and less than 10 triggered stations are different realizations of two unique showers with  $X_{\max} = 1242 \text{ g cm}^{-2}$  and  $X_{\max} = 1296 \text{ g cm}^{-2}$ .

The presence of photon showers with a hadronic-like development is also seen in the top left of Fig 4.11. These shallow showers have a copious production of charged hadrons in the first interaction. In these particular cases, the interaction between the primary photon and the atmosphere have a multiplicity larger than 200, producing hundreds of pions and kaons, while the



**Figure 4.11:** The number of triggered SD stations in terms of the depth of the shower maximum. The observation level depth is represented as a red line. The beige crosses represent the single events while the empty black markers represent the mean values with the  $1\sigma$  deviation.

expected multiplicity is hardly larger than 20 with a rather poor production of pions and kaons. This extraordinary behaviour appears once and again in several studies throughout this thesis and it is a direct consequence of the photonuclear process mentioned in Sec. 2.4.

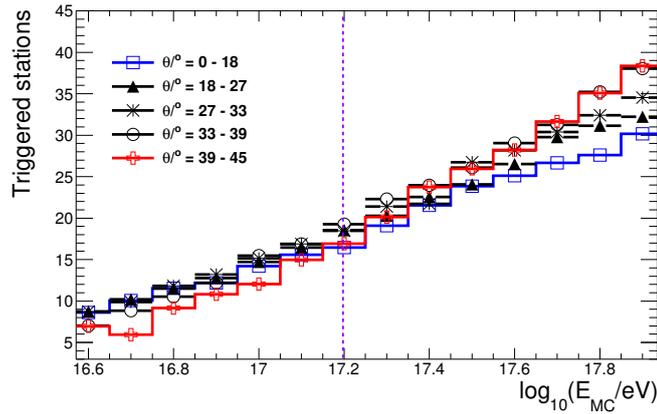
When a primary particle impinges the atmosphere with an inclined incidence, the subsequent EAS would arrive at the observation level in a more mature stage of development. This is translated in a flatter lateral distribution of particles, as discussed in Sec. 2.2. This does not necessarily indicate that more particles would arrive to the ground. As the zenith angle increases, the number of particles arriving to the ground decreases due to the strong atmospheric attenuation if the primary energy is not high enough, and thus the footprint gets smaller. On the other hand, the array spacing is reduced in the coordinate system co-moving with the shower front as the zenith angle increases, thus favoring the proliferation of triggered stations.

This two effects compete in the number of triggered stations, as shown in Fig. 4.12. Indeed, the non-vertical showers tend to fire a larger number of stations given a large enough primary energy. On the contrary, at lower energies they suffer a stronger atmospheric attenuation and thus the number of triggered stations is smaller than in the vertical case. The primary energy at which the transition occurs is  $\sim 10^{17.2}$  eV. Therefore, for the energies of interest of this photon search less triggered SD stations are expected as the zenith angle increases.

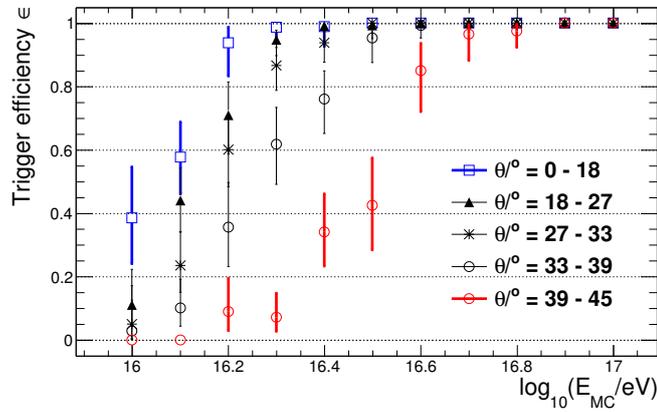
In consequence, the trigger efficiency strongly depends on the zenith angle, as it can be seen in Fig. 4.13. The vertical events populate the lower energies and the SD array is fully efficient at  $10^{16.3}$  eV. Air showers with  $\theta \sim 40^\circ$  trigger the SD array with a probability higher than 0.95 when the primary energy is  $E \gtrsim 10^{16.7}$  eV.

The energy threshold for full efficiency is not affected by the usage of the ideal or realistic array, since at sufficiently low energies the only triggered stations are located in the first hexagon, which is complete in both scenarios (see Fig. 4.3). The trigger efficiency integrated in zenith angles  $\theta < 45^\circ$ , together with the probability of generating a more stringent T4 and T5 array trigger, is shown in Fig. 4.14. It can be seen that all triggered events are at least T4, imposing a zenith-integrated energy threshold of  $10^{16.6}$  eV. The probability of generating a T5 condition is hampered due to the presence of silent stations<sup>13</sup>, leading to a higher energy threshold of  $\sim 10^{16.8}$  eV.

<sup>13</sup>The silent stations are those without any T1 or T2.



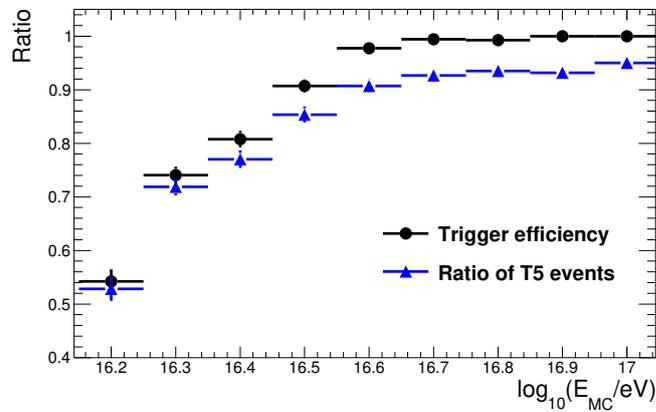
**Figure 4.12:** The average number of triggered SD stations in terms of the primary energy for photon-initiated showers for different zenith intervals. The most vertical (less vertical) showers are distinguished with blue (red) markers.



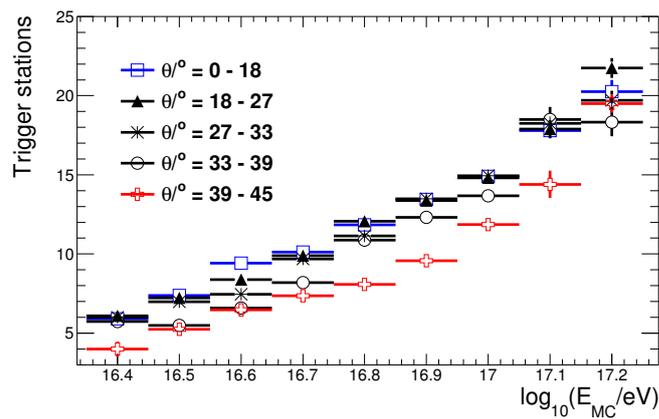
**Figure 4.13:** The trigger efficiency of the SD array in terms of the primary photon energy for different zenith intervals. The most (less) vertical events are represented with blue (red) markers, while intermediate zenith bins are represented with different black markers. The error bars represent the binomial confidence interval according to the Clopper-Pearson method [239].

**Proton primaries** As mentioned before, the trigger efficiency generally depends on the type of primary particle. However, the number of triggered stations does not show a strong dependence on the primary particle, since the averages for proton primaries in Fig. 4.15 are barely distinguishable from the expected means for photon-initiated showers shown before in Fig. 4.12. The average number of triggered stations for vertical events is 9 at  $10^{16.6}$  eV in both cases, increasing to 14 and 17 at  $10^{17}$  eV for photon and proton primaries. Thus, the increase in the number of triggered WCDs starts to be noticeable at the latter energies. Therefore, the systematic effect produced by the primary composition is negligible considering that the realistic array is composed by only 13 WCDs.

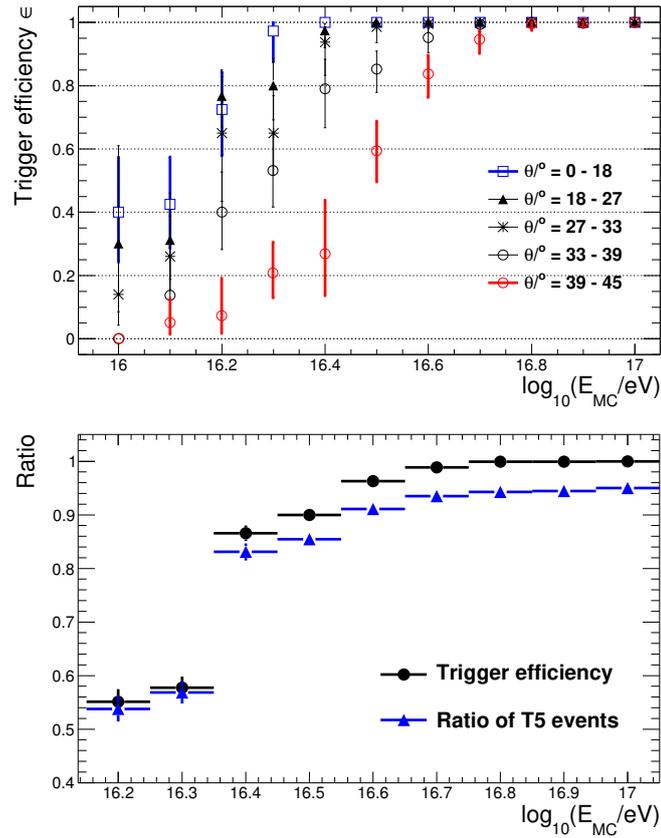
For completeness, the trigger efficiency in terms of primary energy for proton primaries is shown in Fig. 4.16, top. In the same fashion as before, the integrated trigger efficiency is shown in Fig. 4.16, bottom. As expected from the previous discussion, there is not any remarkable difference in the trigger efficiency for proton primaries with respect to the discussed results in the



**Figure 4.14:** The trigger efficiency  $\varepsilon$  (black markers) integrated in zenith angles between  $0^\circ$  and  $45^\circ$  following a  $\sin\theta\cos\theta$  distribution. The ratio between the events with a T5 trigger and all events is shown with blue markers.



**Figure 4.15:** The average number of triggered stations in terms of the primary energy for different zenith intervals. The showers are produced by primary protons. The more (less) vertical showers are distinguished with blue (red) markers.

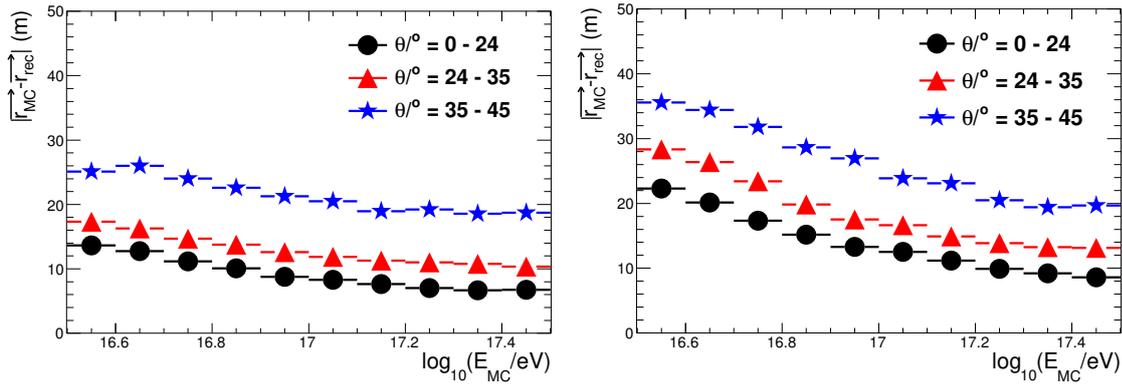


**Figure 4.16:** (Top) The trigger efficiency of the SD array for proton primaries in terms of the primary energy for different zenith intervals. The more (less) vertical events are represented with blue (red) markers, while intermediate zenith bins are represented with different black markers. The error bars represent the binomial confidence interval according to the Clopper-Pearson method [239]. (Bottom) The trigger efficiency (black markers) integrated in zenith angles between  $0^\circ$  and  $45^\circ$  following a  $\sin\theta\cos\theta$  distribution. The ratio between events with a T5 trigger and the total number of events is shown with blue markers.

case of photon primaries. A full efficiency threshold of  $10^{16.3}$  eV and  $10^{16.7}$  eV for  $\theta < 18^\circ$  and  $39^\circ < \theta < 45^\circ$  respectively is found for proton primaries, while the integrated trigger efficiency is at  $\sim 0.95$  for energies above  $10^{16.9}$  eV.

## 4.4 Features of the SD reconstructed events

The reconstruction of the simulated events is performed with the standard Offline sequence [187], but with the suitable configuration files to account for the 433 m spacing [205], such as the energy calibration discussed in Sec. 3.3.6. Since it is fine-tuned to describe events generated by hadron-initiated EAS, the set of reconstructed photon events exhibit some distinctive features that are addressed in this section.



**Figure 4.17:** The absolute difference between the reconstructed and simulated core position in terms of the simulated energy for photon (left) and proton (right) primaries. The markers represent the quoted zenith intervals.

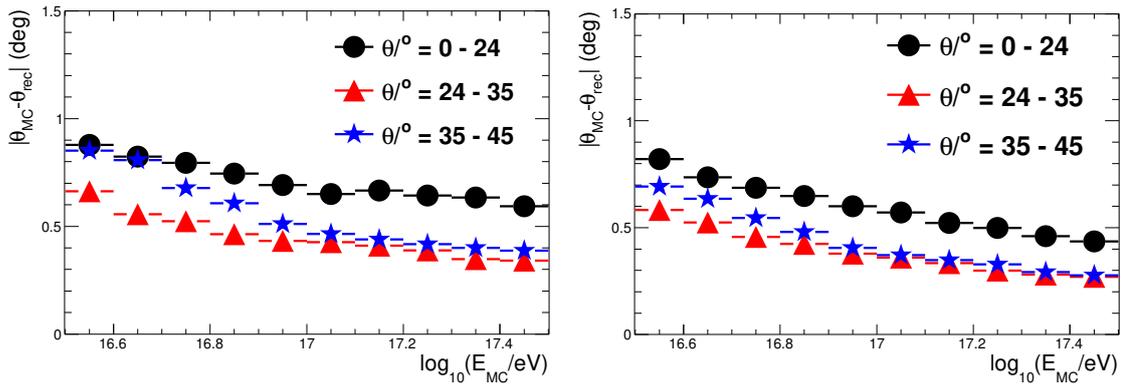
#### 4.4.1 Geometry reconstruction

The quality of the geometrical reconstruction of an event can be assessed through several aspects. It is customary to compare the reconstructed core position  $\vec{r}_{\text{rec}}$  with the simulated core position  $\vec{r}_{\text{MC}}$ , both measured from the origin of a local coordinate system<sup>14</sup>. The absolute difference is shown in Fig. 4.17 for photon and proton primaries. In average, the reconstructed core position deviates at most 30 m (40 m) from the simulated values for  $\theta < 35^\circ$  ( $35^\circ < \theta < 45^\circ$ ). The difference increases with decreasing primary energies which is due to the diminishing number of particles in the EAS.

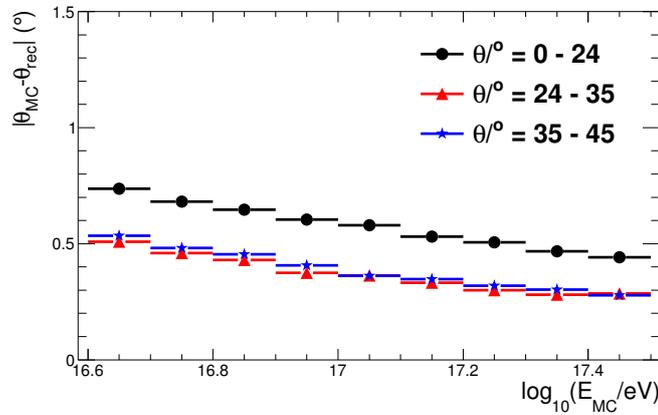
Another aspect to take into account is the precision of the angular reconstruction, in particular the zenith angle. The absolute difference between the reconstructed and simulated zenith angle,  $\theta_{\text{rec}}$  and  $\theta_{\text{MC}}$  respectively, is shown in Fig. 4.18. The average difference in both cases is less than  $1^\circ$ , independently of the features of the primary particle. The angular reconstruction of non-vertical events (blue stars in Fig. 4.18) show a particular trend at  $E_{\text{MC}} < 10^{17}$  eV, independently of the primary particle type. Indeed, the number of triggered WCDs in this energy domain is tightly connected with the zenith angle, as discussed on Fig. 4.15. In particular, events with a  $\theta > 35^\circ$  are composed, in average, by a smaller number of triggered WCDs than in the case of vertical events. Hence, the reconstructed zenith angle starts to depart from the simulated value as the number of available stations decreases. This behaviour is represented in Fig. 4.19, where the angular difference is shown in terms of the primary energy for events with, at least, 7 triggered stations (i.e. with the complete first crown). The trend for the non-vertical events is erased due to the multiplicity constraint.

The angular reconstruction of non-vertical events shows a slight dependence on the position of the shower core at energies below  $\sim 10^{16.9}$  eV. Indeed, if the shower core is close to the central WCD, the stations from the first crown are located at nearly the same distances from it. This degeneracy affects mostly the non-vertical events because of the smaller number of available triggered stations (see Sec. 4.3.1). In any case, this effect is notorious for energies far below the full efficiency threshold.

<sup>14</sup>The origin of the coordinate system is situated in the center of the SD-1500. This point belongs to the zone 19H of the Universal Mercator System (UTM) [240].



**Figure 4.18:** The absolute difference between the reconstructed and simulated zenith angle in terms of the simulated energy for photon (left) and proton (right) primaries. The markers represent the quoted zenith intervals.



**Figure 4.19:** The absolute difference between the reconstructed and simulated zenith angle in terms of the simulated energy for proton events with at least 7 triggered WCDs. The markers represent the quoted zenith intervals.

#### 4.4.2 LDF reconstruction

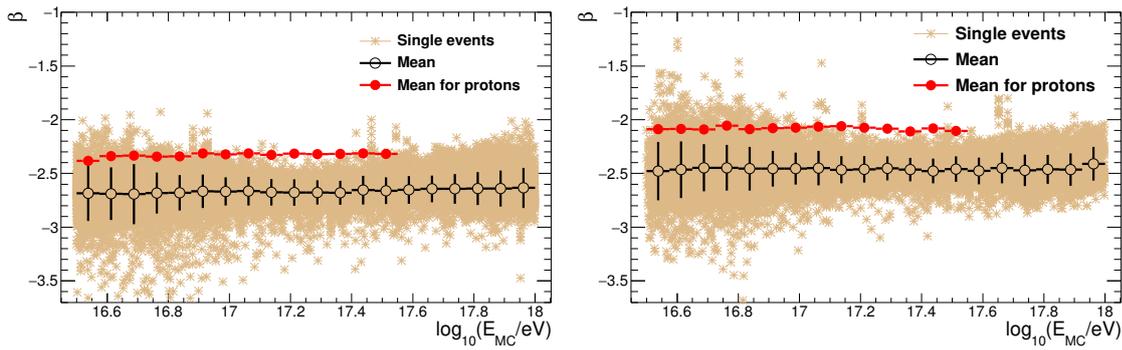
The LDF is fitted with an NKG-like function, as explained in Sec. 3.3.4. The mathematical expression is repeated here for clarity:

$$S(r) = S_{\text{opt}} \left( \frac{r}{r_{\text{opt}}} \right)^{\beta} \left( \frac{r + r_{\text{scale}}}{r_{\text{opt}} + r_{\text{scale}}} \right)^{\beta} \quad (4.2)$$

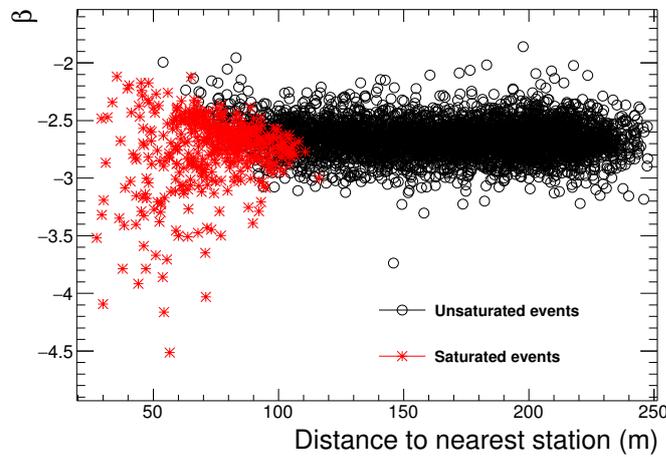
The parameter  $r_{\text{opt}}$  is fixed to 250 m, where slope fluctuations were found to be minimal [207]. The parameter  $r_{\text{scale}}$  is fixed to 700 m.

While a dedicated LDF slope parameterization for the SD-433 was investigated previously [206], it was obtained directly from data, i.e. hadronic events. Thus, that model does not describe photon events well enough because of the fundamental differences in the shower development. Therefore, a dedicated  $\beta$  parameterization has to be obtained a priori in the case of photon primaries. The features of the LDF produced by a photon-initiated EAS are described in this section by fitting the LDF slope in an event-by-event basis.

The fitted slope has a fairly weak dependence on the zenith angle and nearly no dependence on



**Figure 4.20:** The fitted LDF slope  $\beta$  in terms of the primary energy for photon events with  $\theta_{MC} = 0^\circ$  (left) and  $\theta_{MC} = 40^\circ$  (right). The beige crosses represent the single events while the empty black markers represent the mean values with the  $1\sigma$  deviation. The red profiles are the average LDF slope for proton primaries.

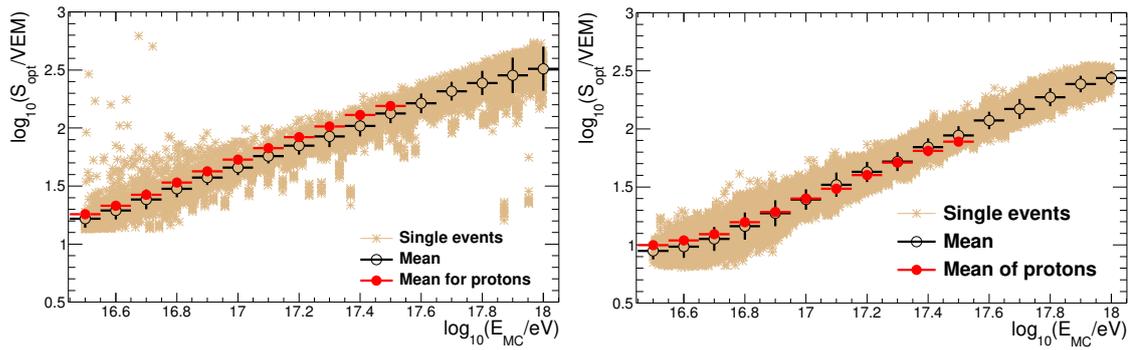


**Figure 4.21:** The fitted LDF slope  $\beta$  in terms of the distance to the nearest station for the subset of photon events with  $E_{MC} < 10^{17}$  eV and  $\theta_{MC} = 0^\circ$ . The black markers (red stars) represent the unsaturated (saturated) events.

the simulated energy, as can be seen in Fig. 4.20. As the primary energy decreases, the fluctuations of the fitted slope are enhanced. This is interpreted in terms of the smaller number of triggered stations, which is translated into less degrees of freedom for the LDF fit. On the other hand, there is a population of events which is visibly departed from the expected average at  $E_{MC} > 10^{17.6}$  eV. These events have a saturated SD station whose signal is underestimated and thus it acts as a lever arm to finally produce a flatter LDF than expected.

The impact of the presence of a saturated station is even greater at low energies, where the spatial distribution of the few triggered stations is not enough to fairly probe the shower front, as can be seen in Fig. 4.21 for the particular subset of vertical events with  $E_{MC} < 10^{17}$  eV. In this scenario, the fitted LDF is strongly dependant on the core position, which in turn causes an increase in the fluctuations of  $\beta$ . On the contrary, there is not any dependence on the distance to the nearest station in the subset of unsaturated events.

The second fitted parameter is the shower size  $S_{opt}$ , which is shown in terms of the primary energy in Fig. 4.22. The value of  $S_{opt}$  increases in average with the primary energy, as expected, but in the case of vertical events its fluctuations also increase. This is due to the presence of high-energy events that, being deeply penetrating, could generate lower signals in the SD stations



**Figure 4.22:** The energy estimator  $S_{\text{opt}}$  in terms of the primary energy for photon events with  $\theta_{\text{MC}} = 0^\circ$  (left) and  $\theta_{\text{MC}} = 40^\circ$  (right). The beige crosses represent the single events while the empty black markers represent the mean values with the  $1\sigma$  deviation. The red profiles are the average energy estimator for proton primaries.

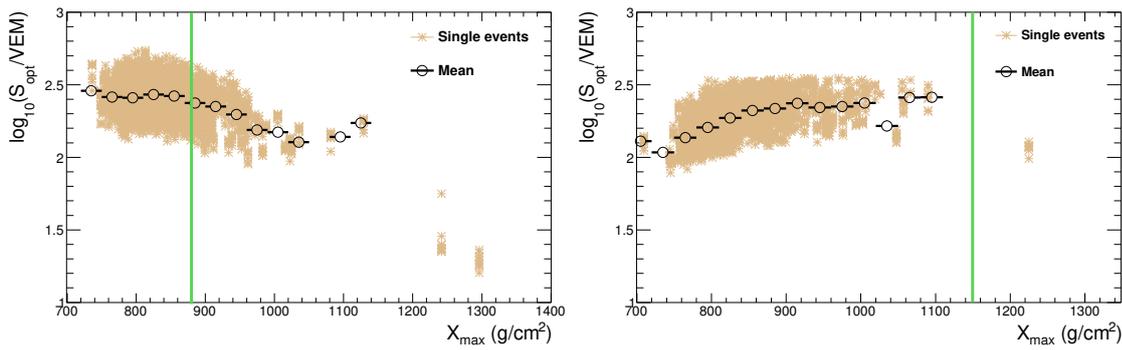
and also more concentrated around the shower axis. Note that these outliers are also seen in Fig. 4.11. The presence of this kind of outliers is suppressed in the case of non-vertical events because the EAS arrive to the observation level after traversing more matter. In fact, the observation level at  $\theta_{\text{MC}} = 40^\circ$  is located at a depth of  $\sim 1146 \text{ g cm}^{-2}$ . The relation between the shower size and the depth of the shower maximum between vertical and non-vertical events is depicted in Fig. 4.23.

On the other hand, upward fluctuations on  $S_{\text{opt}}$  are visible at lower energies. These are the same steeper saturated events explained before. They are reconstructed with an energy much larger than the Monte Carlo value based on the value of  $S_{\text{opt}}$ . They can be distinguished from valid high-energy events by their the footprint size (i.e. the number of triggered stations) which remains small. This kind of outliers are not present in non-vertical events due to the smaller saturation probability, as seen in Fig. 4.24. As the zenith angle increases, the EAS suffer a stronger atmospheric attenuation, as explained before, and thus the expected signal in a WCD near the shower axis is smaller. Therefore, the number of saturated events (i.e. events with at least one saturated WCD) decreases for non-vertical showers. However, as the primary energy also increases, the opposite behaviour is seen. Since the photon-initiated EAS have a deep development (deeper as the energy increases), the shower front arrives to the observation level less developed as the zenith angle decreases. Therefore, inclined EAS tend to produce higher signals in the WCDs at larger zenithal incidence. Similarly to the discussion around Fig. 4.12, there is a certain energy at which the atmospheric attenuation starts to give in to the delayed shower development, which in view of Fig. 4.24, is around  $E_{\text{MC}} = 10^{17.2} \text{ eV}$ .

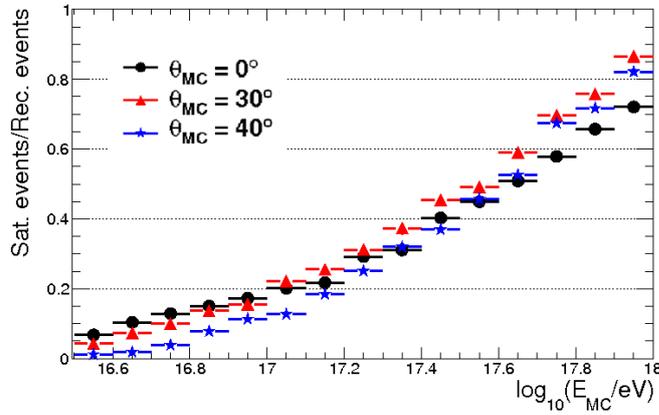
### 4.4.3 Energy reconstruction

The energy calibration implemented in `Offline` for the SD-433 was previously obtained through data analysis (see Sec. 3.3.6). In the general case, the photon energy measured by means of a WCD array is usually underestimated due to their smaller muon content. Several methods have been used to overcome this difficulty in the past years (see [148] for details). The fluorescence technique offers an unbiased estimation of the photon energy due to the near-calorimetric nature of the measurement and the small impact of the missing energy [241].

The late development and, mainly, the lack of muons produce a discrepancy between the reconstructed and MC primary energy named *energy bias*, which is displayed in Fig. 4.25 for



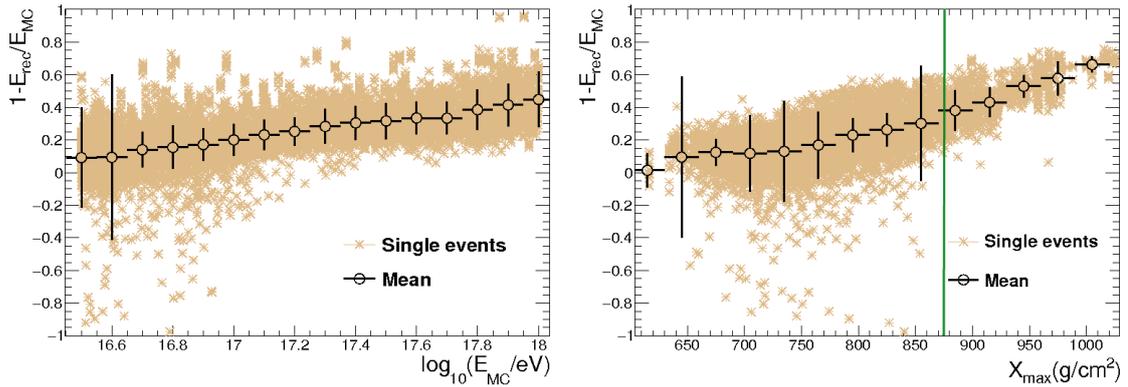
**Figure 4.23:** The fitted parameter  $S_{\text{opt}}$  in terms of the depth of the shower maximum  $X_{\text{max}}$  for photon-initiated showers at  $E_{\text{MC}} > 10^{17.6}$  eV with  $\theta_{\text{MC}} = 0^\circ$  (left) and  $\theta_{\text{MC}} = 40^\circ$  (right). The beige crosses represent the single events while the empty black markers represent the mean values with the  $1\sigma$  deviation. The green lines represent the depth of the observation level at each  $\theta_{\text{MC}}$ .



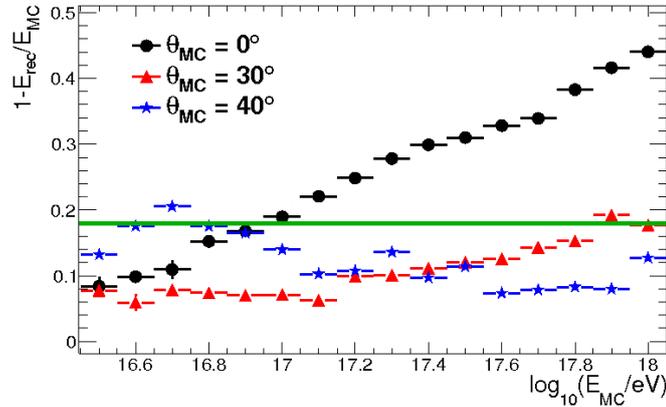
**Figure 4.24:** The ratio between the number of saturated events (i.e. events with at least one saturated WCD) and the number of reconstructed events in terms of the simulated primary energy for EAS produced by photons with the quoted zenith angles (represented by different markers).

vertical events. The energy offset from the MC value grows from 9% at  $E_{\text{MC}} = 10^{16.5}$  eV to 45% at  $E_{\text{MC}} = 10^{18}$  eV for vertical showers. Due to the delayed development of EAS generated by primary photons, the particles tend to be more collimated around the shower axis in the case of vertical incidence. In turn, this behaviour is responsible for steep LDFs with underestimated values for  $S_{\text{opt}}$  (with respect to the hadronic EAS expectation). This effect is accentuated as the primary energy increases because the first interaction occurs at larger atmospheric depth (i.e. closer to the ground). Therefore, the discrepancy between the reconstructed energy, by means of a calibration suited for hadronic events, and the MC primary energy increases with the energy. This can also be deduced from Fig. 4.25, right, where it can be seen that as the shower reaches the observation level at earlier stages of its development, this discrepancy increases on average, independently of the zenith angle.

The upward fluctuations of  $S_{\text{opt}}$  below  $E_{\text{MC}} = 10^{16.9}$  eV produce negative energy offsets as explained before, leading to an increase in the energy offset fluctuations. Since these events with large offsets ( $\lesssim -0.4$ ) are caused by the reconstruction procedure, they exhibit a flat distribution in  $X_{\text{max}}$ . For the same reason, the fluctuations of the energy offset do not follow any particular trend with  $X_{\text{max}}$ .



**Figure 4.25:** The relative energy bias in terms of the MC energy (left) and the shower maximum (right) for photon primaries with  $\theta_{MC} = 0^\circ$ . The beige crosses represent the single events while the empty black markers represent the mean values with the  $1\sigma$  deviation. The green line corresponds to the depth of the observation level.



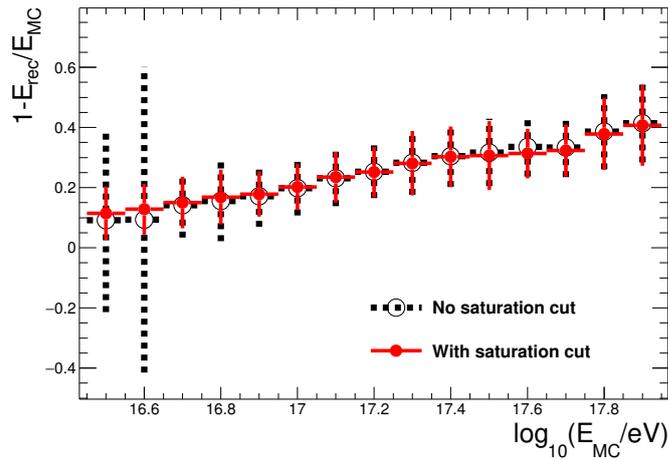
**Figure 4.26:** The relative energy bias in terms of the simulated energy for photon primaries with the quoted zenith angles. The green line represents the SD-433 energy resolution, as found in [206].

The impact of the late development is reduced if the shower traverses sufficient matter, and thus reaching the observation level in a more advanced stage. In Fig. 4.26, the relative energy bias for different zenith angles is shown. The showers with  $\theta_{MC} = 30^\circ$  are reconstructed with an energy offset growing from  $\sim 9\%$  at  $E_{MC} = 10^{16.5}$  eV to  $\sim 18\%$  at  $E_{MC} = 10^{18}$  eV, thus following the same trend as the vertical ones but with a much weaker energy dependence. For  $\theta_{MC} = 40^\circ$ , the atmospheric attenuation starts to be relevant, specially at low energies, and it is the dominant effect at  $E_{MC} < 10^{16.8}$  eV producing a worsening of the energy estimation with respect to the vertical case.

The energy resolution of the SD-433 was calculated as  $\sim 18\%$  between  $E = 10^{16.6}$  eV and  $10^{18}$  eV [242] and it is given by the green line in Fig. 4.26. Therefore, the energy offset is fairly well contained within this resolution for  $E_{MC} < 10^{17}$  eV. At higher energies, the SD array is already fully efficient for non-vertical events, which compose the vast majority of the events in data because of geometrical reasons<sup>15</sup>.

The largest absolute energy offsets are caused, at lower energies, mainly by artifacts of the LDF

<sup>15</sup>The median of the  $\theta$  distribution in data is  $\sim 35^\circ$ , as it was mentioned in Sec. 3.3.6. The arrival directions of the events in data for the present analysis is discussed in Sec. 7.1.1.

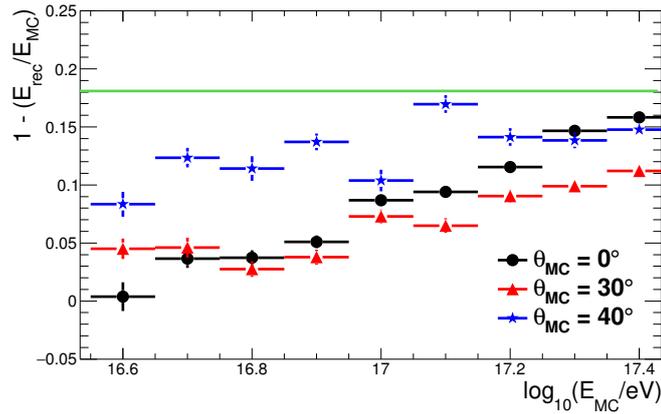


**Figure 4.27:** The relative energy bias in terms of the MC energy for photon-initiated events when no saturation cut is imposed (empty black circles) and when saturated events are excluded (red dots). The error bars correspond to the  $1\sigma$  deviation. In both cases, the events have  $\theta_{MC} = 0^\circ$ .

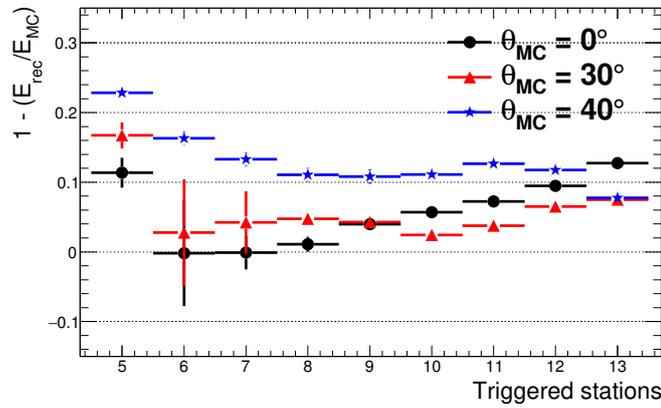
fitting procedure with the slope  $\beta$  as a free parameter. The saturated events at  $E_{MC} < 10^{17}$  eV, though not statistically frequent, show fitted parameters that deviate from the expected averages. In particular, steeper slopes and overestimated  $S_{opt}$ , even though several stations are expected to trigger at these energies. As the primary energy increases, the impact of these outliers is less important, as shown in Fig. 4.27. The non-saturation condition successfully reduces the fluctuations on the energy estimation for low energies, while not affecting the expected averages. At higher energies, the energy offset has a physical origin, being caused by deeply-penetrating showers which produce a smaller signal footprint and  $S_{opt}$  than expected. In this regime, the non-saturation condition does not change the average energy offsets nor their fluctuations.

It is worth mentioning that this discrepancy is also present in hadron-initiated showers due to a mismatch between the predicted muon content by the shower simulations and the Auger data, specially for energies above  $10^{18}$  eV [243]. In the case of the SD-433, the mismatch for proton primaries grows in average from  $< 3\%$  at  $10^{16.5}$  eV to  $< 17\%$  at  $10^{17.5}$  eV, well contained by the SD-433 energy resolution. The mean energy offset for proton primaries is shown in Fig. 4.28. The energy bias for events with  $\theta_{MC} = 40^\circ$  is not as stable as in the other studied cases. Indeed, the number of triggered WCDs tends to be smaller in comparison to the vertical events, which in turn produces an increased discrepancy in the energy reconstruction, as displayed in Fig. 4.29. Since the relation between the energy bias and the number of triggered stations in the case of non-vertical events is stronger than for vertical events, the intrinsic variance of the stations multiplicity produces a larger fluctuations for  $\theta_{MC} = 40^\circ$  than in the other cases.

In summary, the geometrical reconstruction of photon-initiated EAS is suitable at the energies of interest even though the procedure has been specially tuned to describe hadronic events. The saturated events present several issues related to the event reconstruction and, as it is later described in Sec. 5.4, are particularly problematic in terms of the photon-hadron discrimination as well. The standard energy reconstruction shows an intrinsic bias in the case of photon primaries, albeit its effect is, in average, well contained within the energy resolution of the SD-433 for the energies and zenith angles of interest.



**Figure 4.28:** The relative energy bias in terms of the simulated energy for proton primaries with the quoted zenith angles. The SD-433 energy resolution is given by the green line, as found in [206].

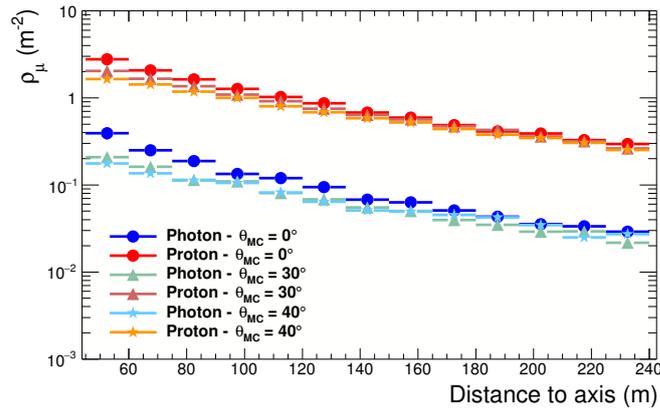


**Figure 4.29:** The relative energy bias against the number of triggered WCDs for proton primaries with the quoted zenith angles.

## 4.5 Features of the MD reconstructed events

Once a simulated event is reconstructed by the SD, the muonic content of the EAS as measured by the MD can be inspected in order to develop criteria sensitive to the mass composition of the primary cosmic ray. Although a detailed fitting procedure has been designed to obtain the MLDF from data, the photon search presented in this thesis uses the muon counting at the single MD station level as the main ingredient, which is performed using a veto of 25 ns and the gap counting strategy (recall Sec. 3.4.5).

The number of muons at the observation level is expected to be around one order of magnitude smaller for photon than proton initiated EAS. Although the soil shielding above the MD stations filter out around half of the muons arriving to the ground, this difference persists. Moreover, depending on the primary energy and zenith, the gap can be even larger. As an example, the muon density measured by the MD stations is shown in Fig. 4.30 in terms of the distance to the shower axis for photon and proton primaries in a large selection of events between  $E_{\text{rec}} = 10^{16.5}$  eV and  $10^{16.7}$  eV. When evaluating the muon signature from the MD stations, the muon density is the preferred magnitude instead of the number of counted muons because it takes into account the different possible detection areas that an MD station can possess and automatically considers



**Figure 4.30:** The average muon density measured with the MD stations in terms of the distance to the shower axis for events at  $10^{16.5} \text{ eV} < E_{\text{rec}} < 10^{16.7} \text{ eV}$ . The primary particles are photon or protons with zenith angles as quoted.

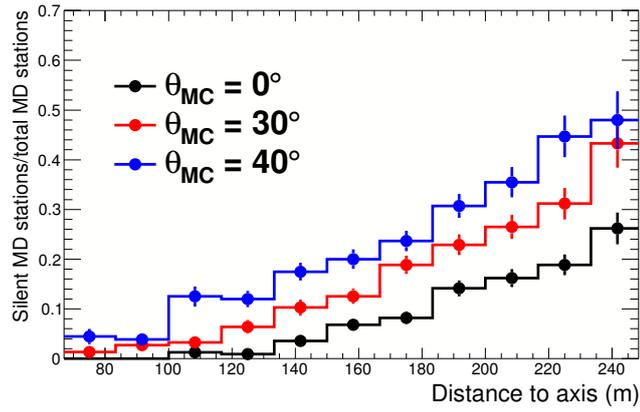
the reduction of the sensitive areas in the case of non-vertical showers. The scarce muon content present in EAS produced by primary photons in comparison to the ones produced by primary protons is clearly visible: the average muon density for proton primaries is at least one order of magnitude higher than for photon primaries, independently of the zenith angle.

Given the distribution of shower cores in the elemental hexagon, it is not possible to probe the entire shower front from the MD point of view (recall that three MD stations are located 750 m from Id 1764 MD station). Thus, distances up to  $\sim 300$  m are shown in Fig. 4.30. In any case, given the energies of interest for this thesis, the expected muon density at larger distances is lower than the resolution of the MD stations, thus imposing a technical limitation for this study. The minimum non-zero muon density that can be measured with a  $60 \text{ m}^2$  MD station is between  $1.7 \times 10^{-2} \text{ m}^{-2}$  and  $2.2 \times 10^{-2} \text{ m}^{-2}$ , depending on the zenith incidence of the EAS.

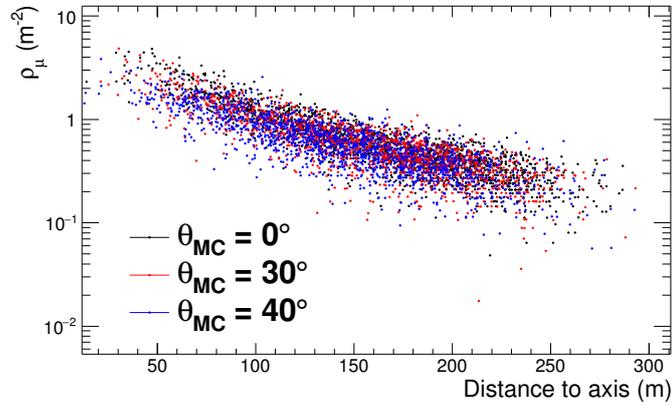
The photon showers, though leaving a muon-poor signature at the observation level, do not always produce a zero reading in the MD stations, i.e. zero counted muons in all the modules of a MD station. On the other hand, a positive number of muons is always detected, even at the lowest energies, in the case of a proton primary. These two aspects, depicted in Fig. 4.31 and Fig. 4.32, indicate that a zero reading at a distance shorter than 300 m from the shower axis is clear signature for a photon primary. However, the selection efficiency for photon showers is rather low, reaching between 30% and 50% at distances around 250 m. This criterium can be loosened to include cases in which the number of muons is positive, but small enough. This simple concept is the starting point of the next chapter and led to the development of the observable  $M_b$ , which is possibly, alongside using the MD stations for photon identification, the most important contribution from this thesis to the Auger community.

In Fig. 4.33 the muon density measured in a MD station in terms of the signal measured in the associated WCD is shown for the same events as before. In light of Fig. 4.30 and Fig. 4.33, the difference in muon content is fairly constant in terms of the distance to the shower axis but also in terms of the signal measured in individual WCD. The latter magnitude is stripped from any geometrical reconstruction systematics, in contrast to the former one.

The signal detected by the WCD adds complementary information to the muon density measured by the MD, as shown in Fig. 4.34, particularly at intermediate muon densities (between  $\sim 0.1 \text{ m}^{-2}$  and  $\sim 0.5 \text{ m}^{-2}$ ). Indeed, for a given measured signal in the WCD, the measured muon

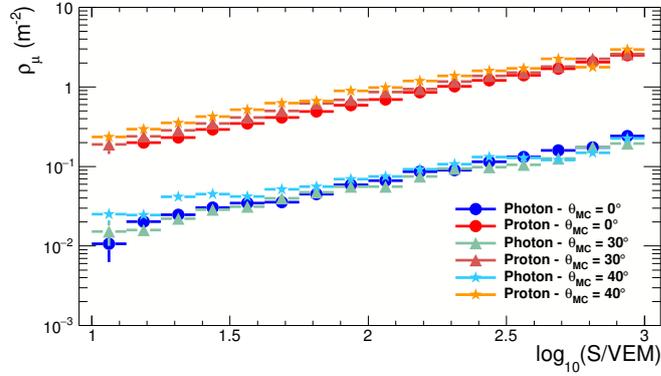


**Figure 4.31:** The ratio between the number of MD stations with a zero reading and the total number of MD stations in a selection of photon events in terms of the distance to shower axis. The energies range between  $E_{\text{rec}} = 10^{16.5}$  eV and  $10^{16.7}$  eV with the quoted zenith angles.

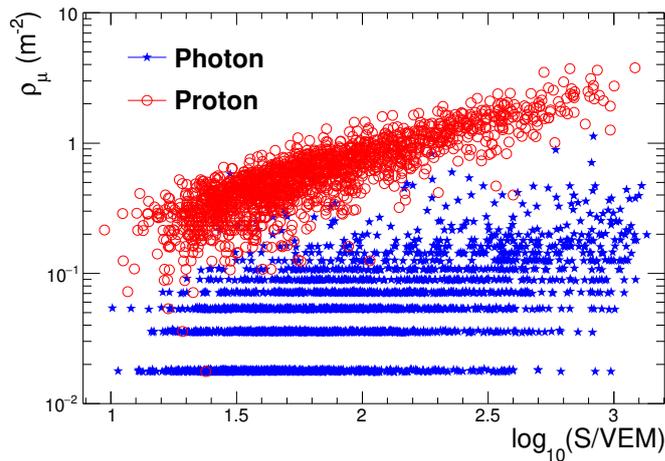


**Figure 4.32:** The measured muon density in terms of the distance to the shower axis for proton primaries with energies ranging between  $E_{\text{rec}} = 10^{16.5}$  eV and  $10^{16.7}$  eV and the quoted zenith angles.

density in the case of photon primaries may be one order of magnitude smaller than for proton primaries, since the lack of muons in the former case may be balanced by the dominant electromagnetic component. This behaviour motivates a combined SD-MD analysis, which is developed in Chapter 6.



**Figure 4.33:** The average muon density measured with the MD stations in terms of the signal in the associated WCD for events at  $10^{16.5} \text{ eV} < E_{\text{rec}} < 10^{16.7} \text{ eV}$ . The primary particles are photon or protons with zenith angles as quoted.



**Figure 4.34:** The muon density measured in an MD station in terms of the signal detected by its associated WCD in an event-by-event basis for events with  $10^{16.5} \text{ eV} < E_{\text{rec}} < 10^{16.7} \text{ eV}$  and  $\theta_{\text{MC}} = 30^\circ$ . The events initiated by photons (protons) are represented by blue stars (red dots).

## Chapter 5

# Photon-hadron discrimination with the observable $M_b$

Simplicity is the ultimate sophistication.

---

Leonardo Da Vinci

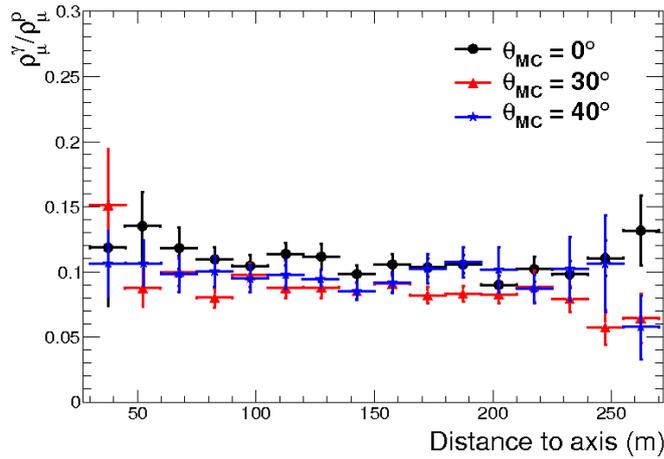
In the search for ultra-high energy photons, it is crucial to define composition-sensitive parameters capable of adequately rejecting the hugely dominant hadronic cosmic-ray flux. As discussed in the previous chapter, the muon density measured by the MD stations can be employed in order to separate photon- from proton-initiated EAS, which are the most photon-like of the possible hadronic primaries in terms of the muon content and the longitudinal development. Thus, a discrimination procedure can be elaborated in the most conservative case by postulating that the cosmic-ray flux has a pure-proton composition. In this context, the EAS initiated by primary protons are considered as the *background*, while the exiguous possible photon component of the flux is thought of as the *signal*.

The silent MD stations located at  $\lesssim 300$  m from the shower axis would represent a clear evidence for a photon primary. Still, photonic EAS usually have a low but non-null muonic component, so the silent station criterium might be too stringent. Also, the scarce muonic component at larger distances from the shower axis in background events, due to the finite MD area, might cause no-muon detection (i.e. events might be tagged as photon primaries even though they are proton-initiated EAS).

One possible way to quantify the muonic footprint in an event-by-event basis is through the MLDF, since it may provide all the required information to distinguish between photon and hadronic primaries in an event-by-event basis, as shown in Sec. 4.5. While this is the most usual approach in mass-composition studies above  $\sim 10^{17.5}$  eV, the MLDF cannot be reconstructed in the energy domain of interest for this photon search without depending on several parameterizations, since only one or two MD stations may be available in any event reconstructed with the SD-433.

Despite the fact that an average difference of one order of magnitude is observed in the muon density between photon and proton primaries, the distance to the shower axis must be taken into account in the discrimination criterium. Without appealing to a cumbersome MLDF procedure, an alternative approach was needed.

The observable  $M_b$  is introduced and described in this chapter as a suitable composition-sensitive observable which makes use of the signals detected in individual MD stations, without



**Figure 5.1:** The ratio between the reconstructed muon density for EAS initiated by photons and protons for events at  $10^{16.5} \text{ eV} < E_{\text{rec}} < 10^{16.7} \text{ eV}$  and three zenith angles, represented by different markers.

needing any MLDF fit.

## 5.1 The observable $M_b$

The ratio between the expected muon density for photon and proton primaries is barely constant with the distance to the shower axis, as deduced from Fig. 4.30. However, the fluctuations of the ratio, as seen in Fig. 5.1, bring some details into consideration. On one hand, the fluctuations are larger closer to the core because of the inherent variations of the point of first interaction between the primary particle and the atmosphere, which mostly affect the hadronic core of the EAS. Thus, different realizations of the same primary would necessarily produce different footprints on ground. This phenomenon is usually referred as to shower-to-shower fluctuations (see Sec. 2.2). On the other hand, the fluctuations on the ratio increase at large distances due to a decrease of the individual muon densities below the resolution of the MD stations (accentuated in the case of photon primaries). In this scenario, the Poissonian uncertainty is the dominant effect on the muonic signature characterization. Therefore, the ratio presents minimal fluctuations at intermediate distances.

These two phenomena rise the necessity of taking into consideration the distance to the shower axis in the photon-hadron discrimination quest. The observable  $M_b$  combines the density of muons and their topological distribution as defined in Eq. 5.1. It is based on the composition-sensitive observable  $S_b$  introduced to exploit the difference in terms of the WCD signal amplitude for different primaries [244], which was successfully used in previous photon studies in Auger [245], specially above  $10^{18} \text{ eV}$  [114].  $M_b$  is calculated through the sum of the mean density of muons  $\rho_\mu^i$  in each MD station  $i$  weighted by its distance to the shower axis  $r_i$ . The muon density is normalized by the expected density for a proton primary  $\rho_\mu^{\text{ref}}$  at the corresponding energy, zenith angle and hadronic model. The reference distance  $r_{\text{ref}}$  is chosen as 250 m, as for the energy estimation of the SD-433. The index  $b$  is a free parameter that must be selected as to maximize the separation between primaries.

$$M_b(E, \theta) = \log_{10} \left( 1 + \sum_i \left( \frac{\rho_{\mu}^i}{\rho_{\mu}^{\text{ref}}(E, \theta)} \right) \times \left( \frac{r_i}{r_{\text{ref}}} \right)^b \right) \quad (5.1)$$

## 5.2 Modelling the $\rho_{\mu}^{\text{ref}}(E, \theta)$ for $M_b$

One of the ingredients needed to calculate the observable  $M_b$  is the expected muon density for proton primaries at  $r_{\text{ref}} = 250$  m. The reference muon density is parametrized in this section for  $E_{\text{rec}} > 10^{16.4}$  eV and  $\theta_{\text{rec}} < 45^\circ$ .

The list of ground particles provided by CORSIKA can be scrutinized in order to have a first insight on the zenithal dependence without introducing the detector and reconstruction systematic effects. A simple model of propagation through the soil is employed to obtain the number of muons at the depth corresponding to a MD station. Each muon is propagated linearly following its original trajectory described by the zenith angle  $\alpha$ . If its kinetic energy on ground  $E$  is greater than the threshold  $E_{\text{th}}$  given by Eq. (5.2) the muon arrives to the underground level [246]. This threshold depends on the mean muon energy loss  $dE/dX$  during its propagation. Since muons can be considered as minimum ionizing particles [247], the Bethe-Bloch equation can be used to obtain that  $dE/dX = 1.6 - 1.8 \text{ MeV g cm}^{-2}$  [75]. On the other hand, the soil density  $\rho_{\text{soil}}$  in the MD Engineering Array location ranges between  $2.28 \text{ g cm}^{-2}$  and  $2.45 \text{ g cm}^{-2}$  [247]. The underground depth  $d$  at which an MD station is buried usually ranges between 2.25 m and 2.50 m.

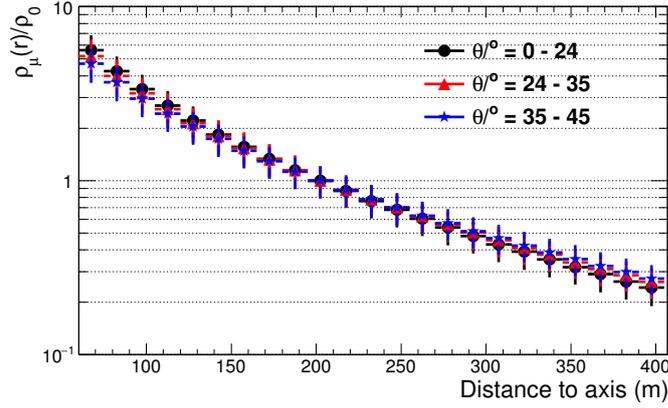
$$E_{\text{th}} = \frac{dE}{dX} \times \rho_{\text{soil}} \times \frac{d}{\cos \alpha} \quad (5.2)$$

For each shower, random values for these three parameters are chosen from uniform distributions according to the corresponding limits. Thus, the typical values for  $E_{\text{th}}$  are  $\sim 0.9 \text{ GeV}$ , in agreement with full detector simulations [213]. The muons that survive this kinetic energy cut conform the muon content at the underground level.

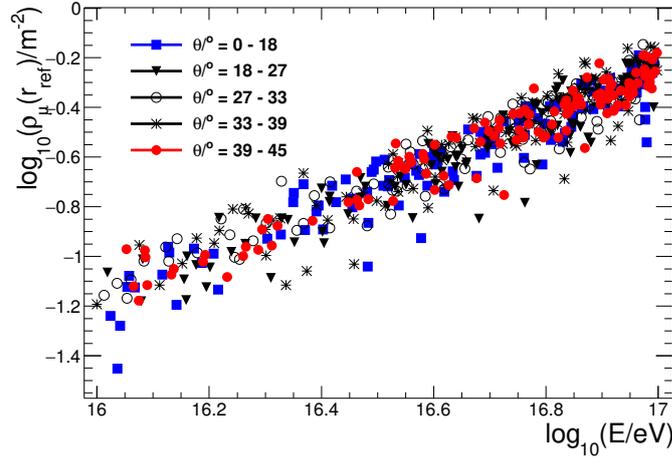
Employing this procedure, the muon density can be estimated at  $\sim 2.25$  m below the surface using the output files from CORSIKA. In Fig. 5.2 the estimated muon density is shown in terms of the distance to the shower axis for proton-initiated EAS with different zenith intervals and a primary energy ranging between  $E_{\text{MC}} = 10^{16.6}$  eV and  $10^{16.8}$  eV. The showers arrive to the observation level at a more mature stage of development as the zenith angle increases, causing a flattening of the shower front and an average increase of the muon density at large distances with respect to the vertical case. However, in the regions near the shower core the vertical showers possess more muons than the more inclined ones because of the diminished attenuation. The transition occurs at  $\sim 200$  m from the shower axis. In consequence, the vertical showers have overall more muons than the non-vertical ones.

The muon density shows a weak dependence on the zenith angle at all distances as shown in Fig. 5.3, particularly at 250 m where the shower-to-shower fluctuations are known to be minimal. The linear dependence in log-log scale between  $\rho_{\mu}^{\text{ref}}$  and the primary energy, which is discussed in the context of the Heitler-Matthews model in Sec. 2.3.1, is verified. In this ideal scenario, the resolution of a  $60 \text{ m}^2$  nominal detection area is enough to sample the muon density at all primary energies of interest.

The reconstruction effects need to be incorporated to obtain a full parameterization of  $\rho_{\mu}^{\text{ref}}$ . The detector layout for this particular study was changed in order to account for the azimuthal



**Figure 5.2:** The muon density at  $\sim 2.25$  m below ground in terms of the distance to the shower axis normalized to  $\rho_0 = \rho_\mu(200)$ . The EAS are generated by proton primaries with energies between  $E_{MC} = 10^{16.6}$  eV and  $10^{16.8}$  eV and with different zenith angles as quoted. The error bars represent the  $1\sigma$  interval. The MLDFs are obtained from the raw showers, without introducing detector and reconstruction effects.

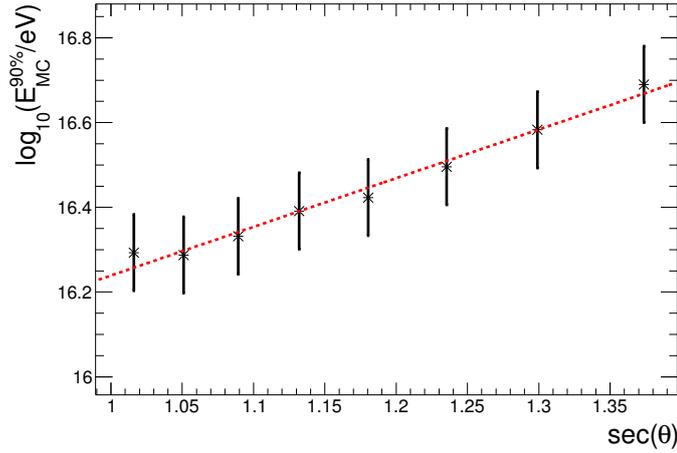


**Figure 5.3:** The muon density at 250 m from the shower axis in terms of the primary energy for showers initiated by protons. Nearly vertical (inclined) showers are represented by blue squares (red dots), while intermediate zenith angles are represented by black markers.

dependence of the muon content and gather high statistics. A ring of 12  $60\text{ m}^2$  MD station, accompanied by a corresponding WCD, at the fixed distance from the shower axis of 250 m was added to the usual detector layout described in Sec. 4.2.2. The addition is performed through the Offline reconstruction machinery after performing the event geometry reconstruction. A quality cut based on the trigger efficiency of the SD array is considered in order to avoid upward fluctuations, i.e. low energy showers with an unusual high muon content. Only showers that trigger the SD more than 90% of the times are considered. The minimum MC energy at which a shower fulfills this requirement in terms of its zenith angle is shown in Fig. 5.4. A linear fit is performed to obtain the parameterized quality cut:

$$\log_{10}\left(\frac{E_{MC}^{90\%}}{\text{eV}}\right) = (15.1 \pm 0.32) + (1.15 \pm 0.27) \times \sec\theta \quad (5.3)$$

After applying the quality cut, the remaining events are shown in Fig. 5.5. The detector



**Figure 5.4:** The MC energy at which the proton-initiated EAS trigger the detector with a probability of at least 90% in terms of the zenith angle. The error bars represent the energy resolution of the SD-433 (see Sec. 3.3.6). The linear fit (red line) represents the quality cut imposed to the sample showers.

effects do not change the linear dependence found before, although the fluctuations in the muon content are increased, specially at low energies. The muon densities corresponding to one and two muons in the MD station are clearly visible as points concentrated at  $\rho_\mu^{\text{ref}} = 10^{-1.8} \text{ m}^{-2}$  and  $\rho_\mu^{\text{ref}} = 10^{-1.45} \text{ m}^{-2}$ . It is worth mentioning that  $\sim 0.03\%$  of the events have a MD station out of the 12 with a null measured muon density, happening at  $E_{MC} < 10^{16.5} \text{ eV}$  and  $\theta_{MC} > 30^\circ$  which is at the verge of the full efficiency threshold. Since the proposed parametrization naturally considers only non-null muon density samples, the existence of null readings could lead to a systematic overestimation of the reference muon density in non-vertical events, which in turn may cause a background event to have an underestimated  $M_b$  and thus being tagged as a photon candidate. This aspect is discussed in Sec. 6.5.

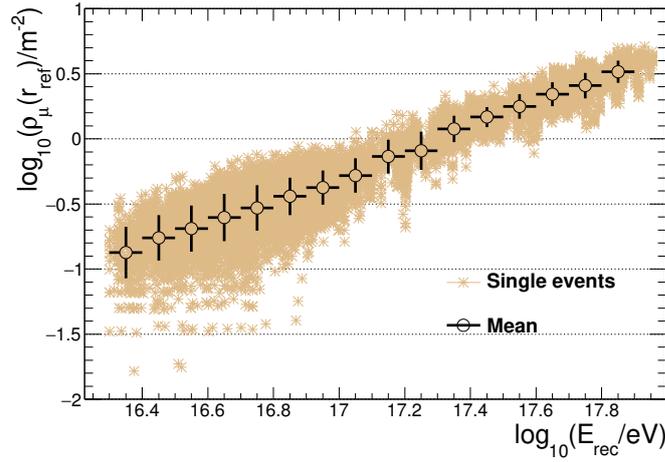
Motivated by the linear dependence in double logarithmic scale, the model of Eq. 5.4 is fitted to the event samples in seven different intervals of zenith angles. The parameters  $a(\theta)$  and  $b(\theta)$  are slowly varying functions of the secant of the zenith angle as shown in Fig. 5.6 which are fitted by a linear ansatz. The results of these fits are shown in Eq. 5.5 and Eq. 5.6 leading to the completion of the  $\rho_\mu^{\text{ref}}$  parameterization. Following the discussion on the Heitler-Matthews model in Sec. 2.3.1, the total number of muons arriving to the ground is expected to depend on the primary energy through a power-law with an index between 0.82 and 0.94 [132]. Indeed, the index  $b(\theta)$  takes values between 0.89 and 0.97 for the zenith angles of interest.

$$\log_{10} \left( \rho_\mu^{\text{ref}} \cdot \text{m}^{-2} \right) = a(\theta) + b(\theta) \times \log_{10} \left( \frac{E_{\text{rec}}}{\text{eV}} \right) \quad (5.4)$$

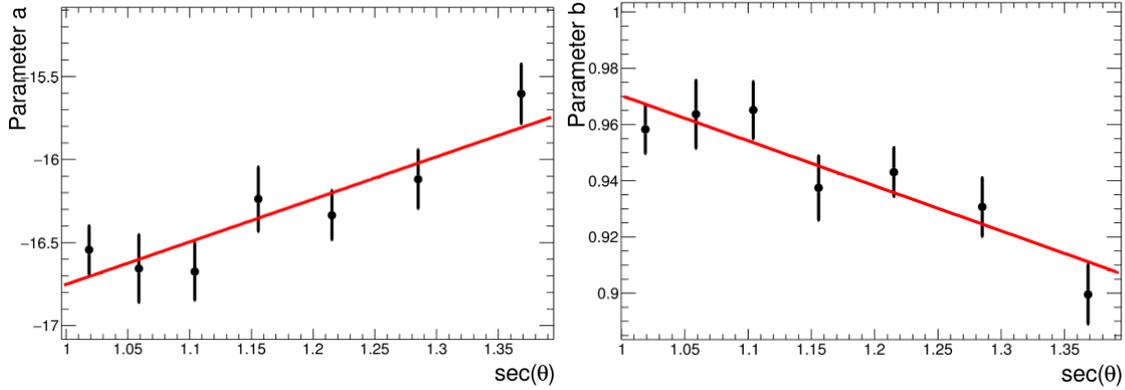
$$a(\theta) = (-19.3 \pm 0.7) + (2.56 \pm 0.55) \times \sec\theta \quad (5.5)$$

$$b(\theta) = (1.13 \pm 0.04) - (0.16 \pm 0.03) \times \sec\theta \quad (5.6)$$

An independent set of events was produced to test the  $\rho_\mu^{\text{ref}}$  parameterization. The number of muons predicted by the model of Eq. 5.4 is compared to the number of muons measured in a  $60 \text{ m}^2$  MD station at 250 m from the shower axis. This is shown in Fig. 5.7. The parameterization is fairly accurate, having an average offset of  $\sim 22\%$  from the measured number of muons at



**Figure 5.5:** The simulated reference muon density measured by the MD stations in terms of the reconstructed energy. Proton primaries are considered following a  $\sin\theta \cos\theta$  distribution for  $\theta_{MC} < 45^\circ$ . The error bars represent the  $1\sigma$  interval.



**Figure 5.6:** The parameters of Eq. 5.4 in terms of the zenith angle. Linear dependences are fitted with  $\chi_{\text{red}}^2 = 1.06$  and  $\chi_{\text{red}}^2 = 1.04$ , respectively, both represented with red lines.

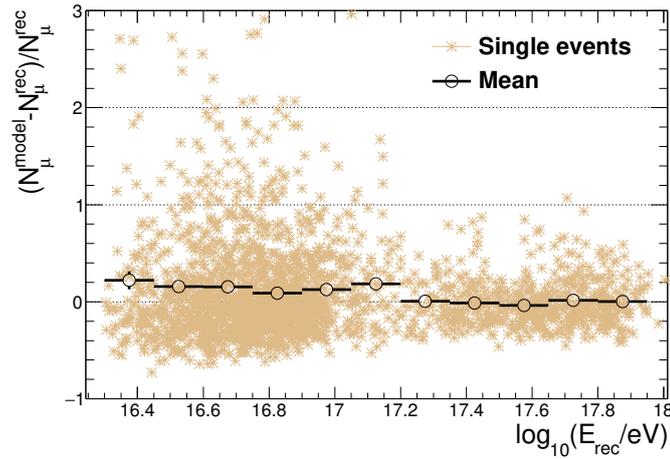
$E_{\text{rec}} = 10^{16.4}$  eV decreasing to negligible average values above  $E_{\text{rec}} = 10^{17.2}$  eV. The resolution of the model is 9% at  $E_{\text{rec}} = 10^{16.4}$  eV increasing to 2% at  $E_{\text{rec}} = 10^{17.8}$  eV.

The model fails to describe the muon-poor events that were also mentioned in relation with Fig. 5.5. As it was pointed out before, if the shower core is too far from a MD station, the measurement of the number of muons would have large Poissonian fluctuations, indicating that in this scenario, a discrimination between photon and proton primaries is less likely to be effective.

### 5.3 Separation power of $M_b$

The observable  $M_b$  reflects the muon content of an EAS. Provided that discrete positions of the shower front can be probed, it is expected that the separation between photon and proton primaries decreases with the information availability. However, some features of the EAS are retained in the  $M_b$  distributions.

The total number of muons at 2.25 m below ground in vertical EAS between  $E_{MC} = 10^{16.6}$  eV



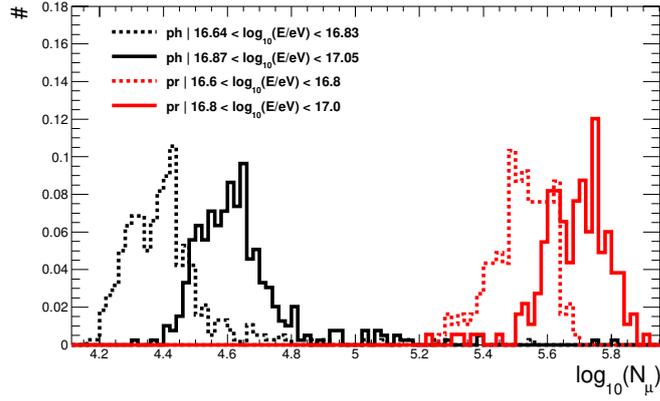
**Figure 5.7:** The difference between the measured number of muons in a  $60 \text{ m}^2$  MD station at 250 m from the shower axis  $N_\mu^{\text{rec}}$  and the predicted value from the model of the  $\rho_\mu^{\text{ref}}$  parameterization  $N_\mu^{\text{model}}$  in terms of the reconstructed energy. Beige markers represent the readings from single MD stations while empty black markers represent the average values.

and  $10^{17}$  eV are shown in Fig. 5.8, where the same procedure as described in Sec. 5.2 is summoned in order to obtain this magnitude from the CORSIKA output. The energy intervals for photon EAS are chosen in accordance to the discussion of the energy offset in Sec. 4.4.3, i.e. higher MC energies are considered to balance the expected energy underestimation. There is a small fraction of the photon showers composing the long tail towards the muon-rich (i.e. proton-like) regime, which can be understood as an effect of the early stages of the shower development: if one of the first interactions is photonuclear, the subsequent development would be more similar to a proton shower<sup>1</sup>. This effect is also responsible for the skewness towards the proton regime. On the other hand, the EAS produced by protons have a also skewness towards the photon distribution, although much less pronounced. In particular, the showers with less muons have several features identical to a photon-initiated EAS, such as the longitudinal profile or the lateral distribution of particles on ground. These extraordinary cases comprehend the hadronic events that could contaminate any primary photon detection. Thus, most of the efforts of the discrimination procedure are aimed at rejecting them, as it is largely discussed in Chapter 6.

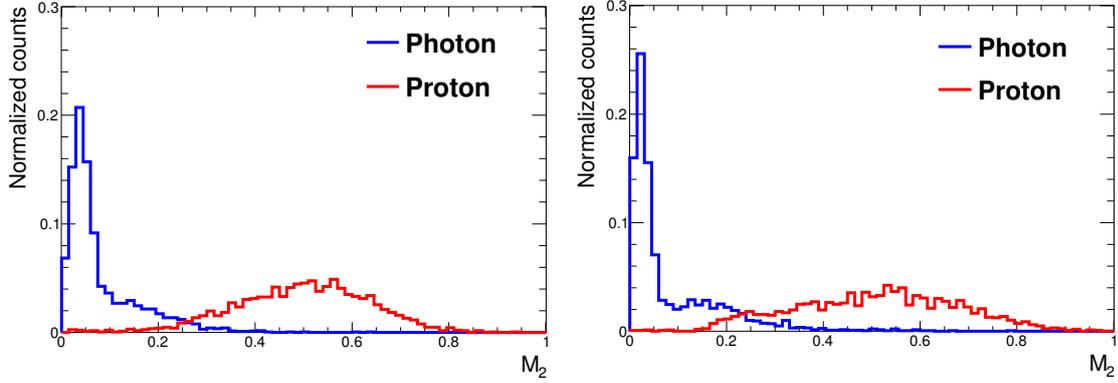
The features of the muon content at the underground level obtained from CORSIKA are seen in the distributions of  $M_b$  as well. In Fig. 5.9, the  $M_2$  distributions for photon and proton events between  $E_{\text{rec}} = 10^{16.9}$  eV and  $10^{17.1}$  eV are shown for different zenith angles. Note that  $M_b$  is calculated with signals measured by the four MD stations. The distributions reflect the features described before: a sharp photon distribution close to  $M_2 = 0$  with a long tail towards positive values and a wider proton distribution extending above  $M_2 \simeq 0.2$ . These are the distinctive features of the signal and background distributions of  $M_b$  that are seen throughout this thesis, with blends produced by selection and quality cuts.

Several improvements can be made in order to minimize the presence of background events with values of  $M_b$  compatible with a photon origin, as seen in Fig. 5.9 in the region below  $M_2 = 0.2$ . An indirect way to measure the overlap between the distributions of a certain observable in different scenarios (or, in this case, different primary particles) is by means of the *merit factor*. A modified version of the merit factor is used to take into account the  $M_b$  distributions asymmetries,

<sup>1</sup>This occurs with a larger probability for higher primary energies, being  $\sim 0.5\%$  ( $\sim 2\%$ ) at  $10^{17}$  eV ( $10^{17.8}$  eV).



**Figure 5.8:** The total number of muons per shower at the underground level (histograms normalized to the number of entries). The primary MC energies for protons (in red) are chosen between  $10^{16.6}$  eV and  $10^{17}$  eV, while for photons (in black) similar intervals are selected accordingly (see text for details). The zenith angle is fixed at  $\theta_{MC} = 0^\circ$ .



**Figure 5.9:** The  $M_b$  distributions for  $b = 2$  according to Eq. 5.1 for photon (blue histograms) and proton (red histograms) events with energies between  $E_{rec} = 10^{16.9}$  eV and  $10^{17.1}$  eV. The zenith angles are  $\theta_{MC} = 0^\circ$  (left) and  $\theta_{MC} = 40^\circ$  (right). Note that  $M_b$  is calculated with signals measured by the four MD stations.

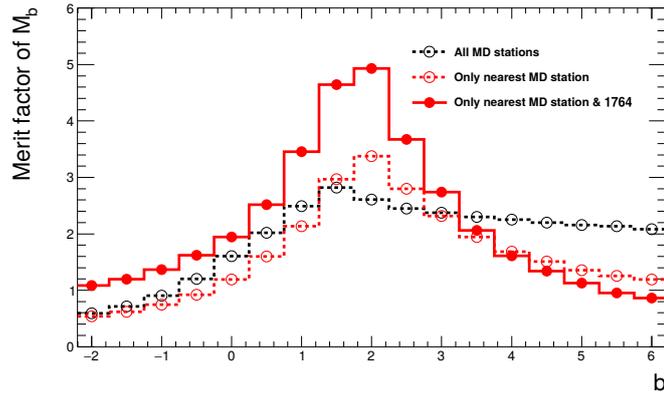
like in previous  $S_b$  studies [248]. Medians are used instead of the means and quantiles at 84% and 16% define the dispersion of the distribution of  $M_b$  as seen in Eq. 5.7. The merit factor  $\eta$  is then defined by Eq. 5.8.

$$\sigma_{68}^2 [M_b] \equiv \left( \frac{M_b^{84} - M_b^{16}}{2} \right)^2 \quad (5.7)$$

$$\eta = \frac{|\text{Med}(M_b^\gamma) - \text{Med}(M_b^p)|}{\sqrt{\sigma_{68}^2(M_b^\gamma) + \sigma_{68}^2(M_b^p)}} \quad (5.8)$$

The separation power can be assessed in a variety of scenarios depending on the topological distribution of the MD stations providing the signals for the  $M_b$  calculation. Thus, for this particular analysis a dedicated event production with EAS cores distributed in a rectangular area containing the whole central hexagon has been done. Also, the MC zenith angle is fixed to  $\theta_{MC} = 0^\circ$  to decouple this dependence from the topological analysis. The extension of this analysis to non-vertical events is discussed in Sec. 5.6.

The optimal value of  $b$  should be chosen according to the degree of separation between photon



**Figure 5.10:** The merit factor  $\eta$  of the observable  $M_b$  in terms of  $b$ . The events are selected according to their reconstructed energy between  $E_{\text{rec}} = 10^{16.9}$  eV and  $E_{\text{rec}} = 10^{17.1}$  eV at  $\theta_{\text{MC}} = 0^\circ$ . Different cases for the  $M_b$  calculation are explored (see text for references).

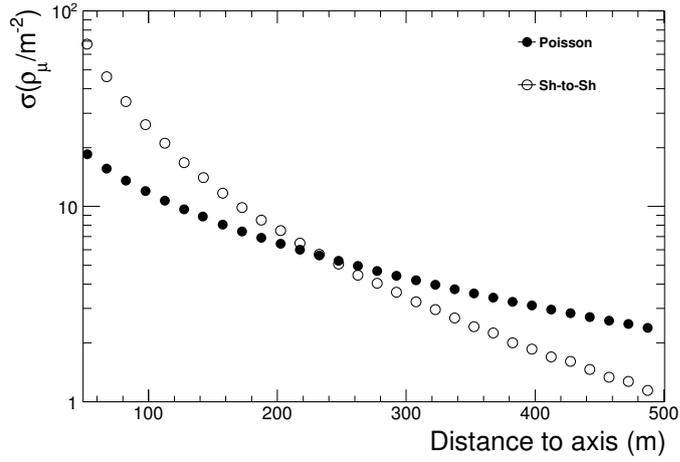
and proton showers quantized by  $\eta$  in terms of the primary energy and zenith. A sweep with steps of 0.01 in the value of  $b$  between  $-2$  and  $6$  is performed. The corresponding values of  $\eta$  are then averaged in bins of  $0.5$ . The evolution of  $\eta$  with the value of  $b$  is shown in Fig. 5.10 in different scenarios for energies between  $E_{\text{rec}} = 10^{16.9}$  eV and  $10^{17.1}$  eV. It is important to note that in view of the particular layout used in this study (see Fig. 4.3), where not every WCD is paired with a corresponding MD station, the hottest WCD (i.e. nearest to shower core) may not coincide with the hottest MD station. Therefore the three considered scenarios are: i) the case in which the signals from the four available MD stations are used (black markers), ii) the case where only the hottest MD station is used irrespective if it coincides with the hottest WCD (red empty markers), and iii) the case where the hottest MD station and the hottest WCD are in the same position (red filled markers). This last condition is fulfilled if the core is inside the elemental hexagon around the central position of the array, resembling a T5 condition. Therefore, this latter case would be recreated in the SD-433 for the KT MD station with Id 1764.

When using only the nearest MD station the separation is increased both if the hottest MD station coincides or not with the hottest SD station. This can be understood in view of the inherent Poissonian fluctuations in the number of muons. Due to the finite sampling size of the detector the number of muons is dominated by Poissonian uncertainties for large distances (i.e. low number of muons). The shower-to-shower fluctuations produced by the shower development dominate over the Poissonian fluctuations of the number of muons at distances shorter than approximately half of the spacing<sup>2</sup>, as it can be seen in Fig. 5.11. Thus, MD stations nearer than this distance threshold provide the most valuable information in terms of discrimination, while the further MD stations mainly add noise<sup>3</sup>. The consequence of using the nearest MD station is an increase of the merit factor of  $M_2$  from  $2.60$  to  $3.38$ . When using only one MD station close enough to the core (full red markers in Fig. 5.10), the merit factor of  $M_2$  reaches a value of  $4.93$ , representing an improvement of  $\sim 90\%$  with respect to the scenario in which all MD stations are used for the  $M_2$  calculation.

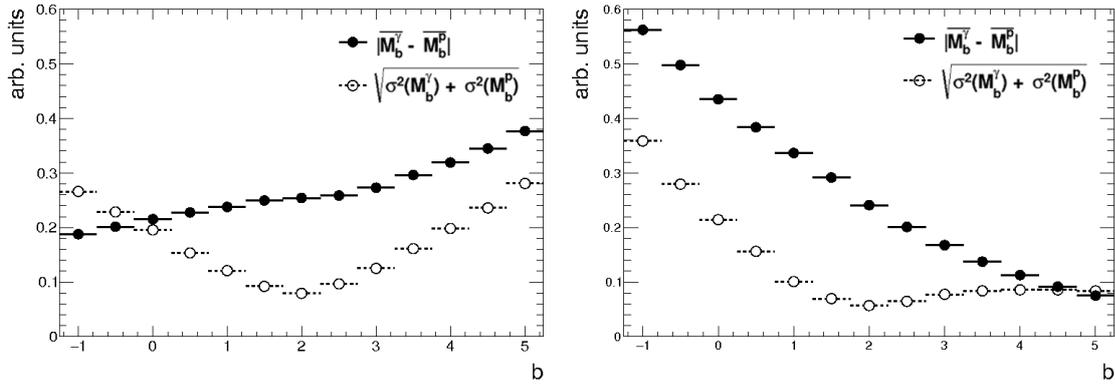
The best separation from the point of view of the topological distribution of accepted events is attained if they are located in the elemental hexagon around the central MD station (i.e. around

<sup>2</sup>In this context, the Poissonian fluctuations are defined as the square root of the average number of muons in a sampling area of  $60 \text{ m}^2 \times \cos\theta$ , whereas the shower-to-shower fluctuations in the number of muons is mathematically obtained from the RMS of the average muon density at a given distance from the shower axis.

<sup>3</sup>This is strictly true for the energies studied in this section. For a higher primary energy, the Poissonian fluctuations are expected to dominate at larger distances [249]. Then, this quality cut might be relaxed to contain further MD stations as well, although the reduced area of some of the other MD stations might also play a negative role.



**Figure 5.11:** The uncertainty contributions in the number of muons per  $60 \text{ m}^2$  at  $2.25 \text{ m}$  depth for proton showers of  $E_{MC} = 10^{17} \text{ eV}$ . The inherent Poissonian fluctuations due to the finite sampling area are represented by filled markers. The shower-to-shower fluctuations are represented by empty markers.

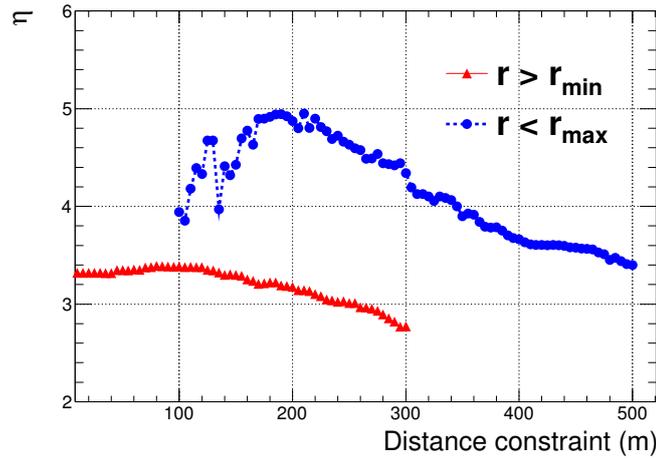


**Figure 5.12:** The evolution of the numerator (full markers) and denominator (empty markers) of the merit factor in terms of  $b$  for two different scenarios: (left) when using only the nearest MD station without any distance constraints and (right) when considering only events with the nearest MD station in the elemental hexagon.

the KT position). This is the main reason for the choice of the area to toss the shower cores, back in Chapter 4.

Focussing on the calculation of  $M_b$  by using only the nearest MD station, the  $b = 2$  is selected by a minimum of the deviation term (denominator of Eq. 5.8). On the other hand, the difference of the medians (numerator of Eq. 5.8) has a monotonic tendency with  $b$ . If no distance cut is placed on the nearest MD station, there is a higher probability for it to be at  $r > 250 \text{ m}$  and hence the ratio  $r/r_{ref} > 1$  for most of the cases. On the contrary, if the nearest MD station is allowed to be only in the elemental hexagon, then  $r/r_{ref} < 1$  for all events. This is represented in Fig. 5.12 for these two possible scenarios.

The MLDF depends on the distance in a non-linear way (see the ansatz of Eq. 3.13). However, it can be approximated by a power-law as seen before in Fig. 4.30. For example, the slope of a power-law fit between  $100 \text{ m}$  and  $500 \text{ m}$  for vertical photon and proton showers at  $E_{MC} = 10^{17} \text{ eV}$  are  $(-1.98 \pm 0.18)$  and  $(-1.96 \pm 0.06)$  respectively. On the other hand, the maximum separation is attained for values of  $b \sim 2$  in all the considered scenarios. Remarkably, his value reflects



**Figure 5.13:** The merit factor of  $M_2$  calculated with MD stations satisfying a minimum (maximum) required distance to the shower axis with red (blue) markers. The events have energies between  $E_{\text{rec}} = 10^{16.9}$  eV and  $E_{\text{rec}} = 10^{17.1}$  eV and  $\theta_{\text{MC}} = 0^\circ$ .

qualitatively the power-law slope of the MLDF at the relevant distances of the SD-433.

The separation power obtained by pure muon counting without introducing information from the position of the MD stations is reflected by  $b = 0$ . In this case, the merit factor is less than half of the maximum possible, enforcing the importance of a compound parameter that considers the muon radial dependance.

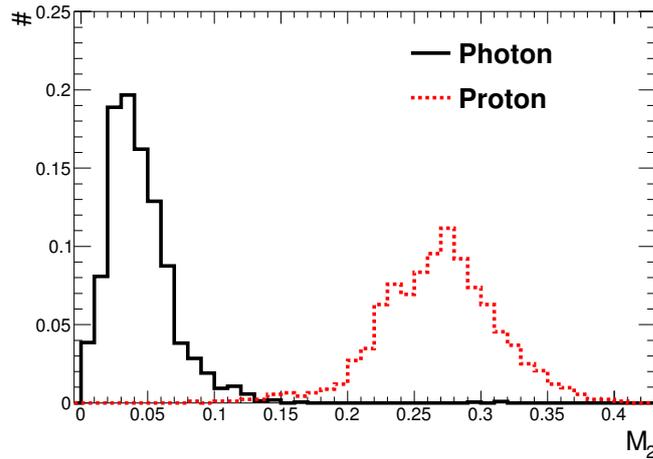
A similar merit factor can be obtained by placing a constraint of  $r \lesssim 200$  m to the nearest MD station as deduced from Fig. 5.13. A more stringent maximum allowed distance leads to an errant behaviour due to the low statistics and a looser requirement suffers from the same limitations as explained before. In a complementary fashion, a minimum required distance does not lead to a significant improvement of the separation. On the contrary, as  $r_{\text{min}}$  increases, the separation decreases monotonically, indicating that the valuable information lies in regions closer to the core, as mentioned before. In summary, by asking for the shower core to lie in the elemental hexagon, which is a usual requirement for mass-composition studies, the best separation is obtained. Therefore, this condition is imposed from now on.

It is worth noting that this high merit factor comes at the price of a reduced acceptance for  $M_b$ , which is the elemental hexagon of the SD-433. Although a composition-sensitive parameter from a MLDF, for example  $\rho_\mu(250)$ , gives a smaller merit factor of 2.76 at this fixed energy and zenith angle, it could be used in a more general scenario in which a larger acceptance is desired.

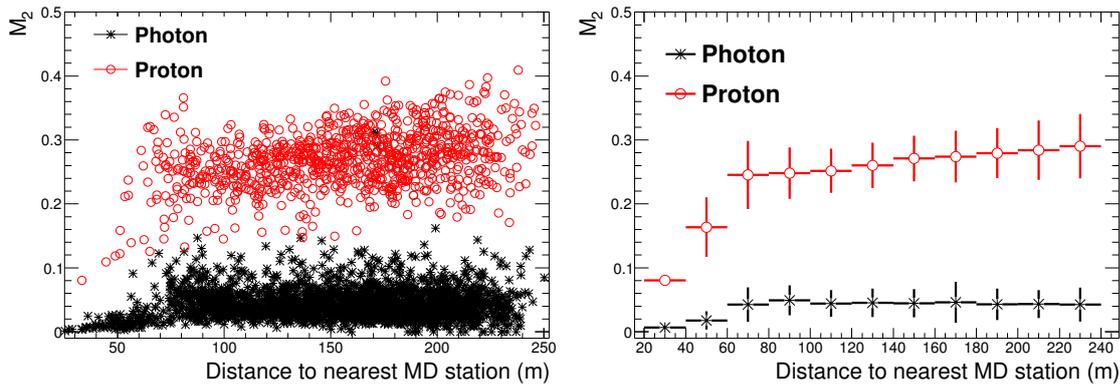
## 5.4 Saturation effects on $M_b$

By calculating  $M_b$  with only the nearest MD station (hence, converting the sum into a single term in Eq. 5.1) and requiring the coincidence of the hottest SD and MD station, thus verifying a T5 condition in the central hexagon of the SD array, the maximum separation possible is attained.

The distributions for  $M_2$  for vertical events with energies between  $E_{\text{rec}} = 10^{16.9}$  eV and  $10^{17.1}$  eV, calculated under the aforementioned conditions, for proton and photon primaries are shown in Fig. 5.14. It can be seen that  $\sim 1\%$  of proton events are contained by the photon distribution. These



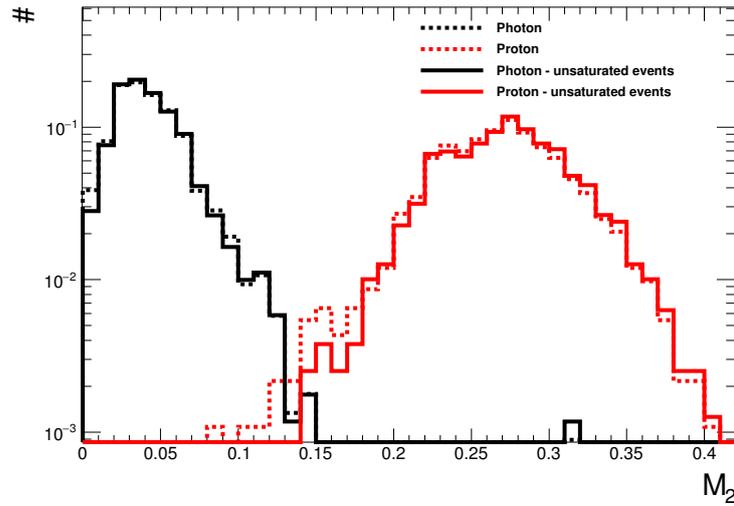
**Figure 5.14:** The  $M_2$  distributions for vertical photon (black solid line) and proton (red dotted line) events by using the nearest MD station when it is also the nearest SD station. The events are selected according to their reconstructed energy between  $E_{\text{rec}} = 10^{16.9}$  eV and  $E_{\text{rec}} = 10^{17.1}$  eV. The distributions are normalized to the number of entries.



**Figure 5.15:**  $M_2$  in terms of the distance to the nearest MD station for vertical photon (black stars) and proton (red circles) primaries in an event-by-event basis (left) and in mean values with  $1\sigma$  deviation (right). The energies range between  $E_{\text{rec}} = 10^{16.9}$  eV and  $E_{\text{rec}} = 10^{17.1}$  eV.

events have their cores extremely close to the MD station, as it is seen in Fig. 5.15. Thus its distance to the shower axis, which in turn defines the value of  $M_2$ , is comparable to the reconstructed core position uncertainty (see Sec. 4.4.1). For example, the proton event with the lowest  $M_2$  (seen in the bottom left corner of Fig. 5.15, left), has an  $M_2 = 0.08$ . If the distance to the simulated core position is used instead,  $M_2^{\text{MC}} = 0.183$ , which is similar to the expected values for proton events.

Since the hottest WCD is low-gain saturated in all these events, a quality cut on this condition can be imposed. In this case, 80% of the events survive the cut independently of the primary type. The  $M_2$  distributions before and after this cut are shown in Fig. 5.16. Although this quality cut leads to a small increase in the merit factor (leading to a final value of 5.06), all the proton events that could be clearly misclassified as photon events (entries of the red dotted histogram with  $M_2 < 0.14$ ) are rejected without changing the overall shape of the distributions. A few photon events can be promptly identified inside the proton distribution. These events are caused by muon-rich EAS and are not accessible for discrimination by the current analysis. The proton-like signature of the shower development is not only seen in the muon footprint, but also in the LDF features and the longitudinal profile, conforming an upper limit to the signal efficiency of the

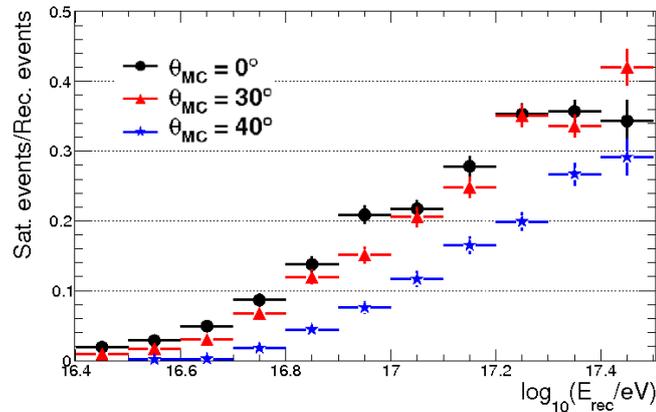


**Figure 5.16:** The  $M_2$  distributions for vertical photon (black lines) and proton (red lines) events including events with a low-gain saturated SD station (dotted style) and excluding them (filled line). The events are selected according to their reconstructed energy between  $E_{\text{rec}} = 10^{16.9}$  eV and  $10^{17.1}$  eV. The distributions are normalized to the number of entries.

discrimination procedure.

This quality cut may be too stringent for higher primary energies. For example,  $\sim 40\%$  of the proton events at  $E_{\text{rec}} = 10^{17.5}$  eV are low-gain saturated, as seen in Fig. 5.17. Nevertheless, the event saturation probability is less than  $\sim 20\%$  for energies below  $E_{\text{rec}} = 10^{17}$  eV that would compose the bulk of the data, as it described in Chapter 7. Note that this estimation is based on events located inside the elemental hexagon around the central MD station. If the events are distributed along larger areas, and in particular at farther distances from the nearest WCD, the event saturation probability decreases.

The quality cut on the event saturation status is equivalent to a distance cut, since the mean distance to the saturated SD station is  $\sim 85$  m. Another possible quality cut could be based on a constant minimum allowed distance to the shower axis. The requirement of  $r_{\text{nearest}} > 50$  m, which is the inner radius cut used in the simulation procedure, filters out only the most problematic



**Figure 5.17:** The ratio between the number of saturated events (i.e. events with a low-gain saturated WCD) and the number of reconstructed events in terms of the reconstructed energy for events produced by primary protons with the quoted zenith angles.

events, while leaving  $\sim 98\%$  of events. In terms of the merit factor, this approach does not lead to a significant improvement with respect to the saturation quality cut if  $r_{\text{nearest}} < 120$  m, while for larger distances the separation decreases abruptly as shown in Fig. 5.13. As it was described before, the inclusion of saturated events in the analysis may also introduce artificial fluctuations in several reconstructed observables at the energies of interest for this thesis. Therefore, the standard saturation cut is adopted for the rest of the analysis.

## 5.5 Systematic effect of the reconstructed energy

The events considered in the previous sections have a fixed MC energy and zenith angle and were used for the sole purpose of describing the features of the observable  $M_b$  and the impact of the different selection and quality cuts. A thorough background rejection estimation must include the systematic effects from the primary energy reconstruction. The distributions of  $M_b$  can be obtained for both primaries under the assumptions that led to the best separation power, as described in Sec. 5.3 and 5.4.

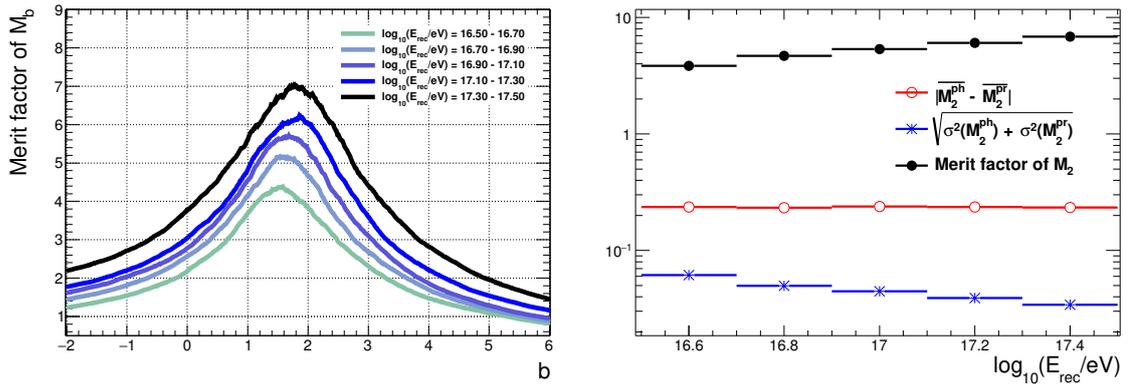
The evolution of  $\eta$  with the value of the index  $b$  is shown for different  $E_{\text{rec}}$  intervals in Fig. 5.18, left. The merit factor exhibits a plateau between  $b = 1.5$  and  $b = 2$ . In particular, the merit factor of  $M_2$  is at most below 11% of the maximum possible. This maximum deviation is reached in the first  $E_{\text{rec}}$  bin. Since the depth of the shower maximum decreases with the primary energy (i.e. the shower reaches its maximum development at higher altitudes), the shower front is flatter when it arrives to the observation level for lower primary energies. Therefore, considering that the value of  $b$  reflects qualitatively the slope of the MLDF (see Sec. 5.3), it is expected that the best separation is attained, in average, for smaller values of  $b$  as the primary energy decreases. In the case of vertical events, this behaviour has a relatively low impact on the separation, so that  $M_b$  with  $b = 2$  permits a suitable discrimination, even at  $E_{\text{rec}} = 10^{16.5}$  eV.

The average merit factor of  $M_2$  in terms of the primary energy is shown in Fig. 5.18, right. The separation power (solid black markers) grows with the primary energy due to an average decreasing behaviour of the parameter fluctuations (star blue markers) and a rather stable difference of the parameter medians (empty red markers) in terms of the primary energy<sup>4</sup>. The trend of these two parameters are actually determined by the proton events alone. This can be deduced from the shape of the  $M_b$  distributions: both the median and the dispersion is larger for proton than photon events. As an example, the median of the  $M_2$  distributions for photon and proton primaries in terms of the reconstructed energy is shown in Fig. 5.19. Remarkably, the ratio between the median for proton and photon primaries is around  $\sim 12\%$  above  $E_{\text{rec}} = 10^{16.5}$  eV, reflecting the one order of magnitude difference in the muon content expected between photon- and proton-initiated EAS.

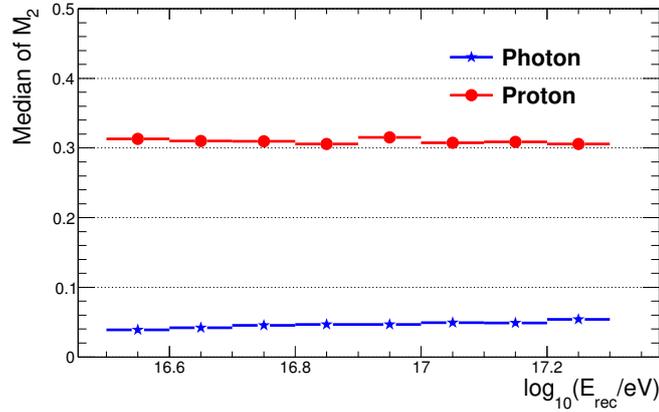
### 5.5.1 Background rejection and signal efficiency

Although the merit factor is a feasible measurement of the discrimination power of an observable, the most important aspect that must be studied is its background rejection capabilities, since in

<sup>4</sup>This is a direct consequence of the energy-sensitive normalization of the muon density  $\rho_{\mu}^{\text{ref}}$  used in the calculation of  $M_2$ . In this way, the observable exhibits a weak dependence on the reconstructed energy.



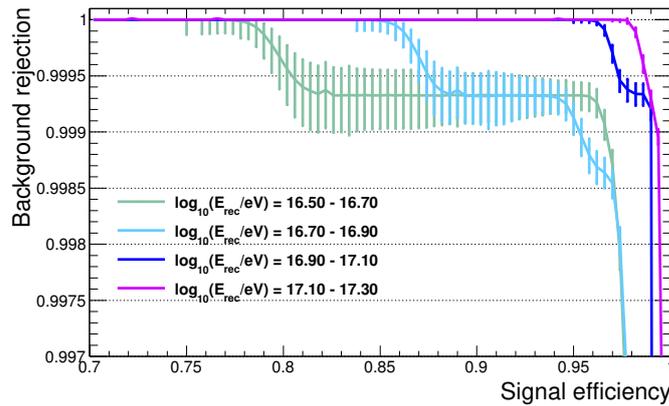
**Figure 5.18:** (Left) The merit factor of the observable  $M_b$  between photon and proton primaries in terms of  $b$ , where the line colors represent intervals of reconstructed energies. (Right) The evolution of the merit factor of  $M_2$  with the reconstructed primary energy (solid black markers), which is calculated as the ratio between the difference of the medians of  $M_2$  (empty red markers) and the fluctuation term (star blue markers).



**Figure 5.19:** The median of the  $M_2$  distributions for photon (blue stars) and proton (red dots) primaries in terms of the reconstructed energy.

practice the hadronic background will be hugely dominant over any photon component in the cosmic-ray flux. For this reason, virtually a full background rejection is needed in order to detect a hypothetical tiny ultra-high energy photon flux. The *benchmark background rejection* employed in photon searches is 99.90%, which means that a straight-forward cut in the observable phase space, such as in Fig. 5.16, can be placed in order to reject 99.90% of the proton events (and thus, any heavier primaries). In a realistic scenario, some photon events would also get rejected. The ratio between the photon events surviving this cut (and thus, being tagged as photon candidates by the detection procedure) and the total number of photon events is defined as *signal efficiency*.

An adequate background rejection estimation is achievable given sufficient event statistics (i.e. at least  $> 10^4$  events in each energy and zenith bin) which is usually a technical and computational challenge. One possible way to circumvent this issue is by describing the  $M_2$  distributions with a tailor-made phenomenological model. While this could be done for vertical events at a fixed energy (see Appendix B), that proposed model suffers from severe problems in different scenarios. Therefore, the *bootstrapping* technique is employed to assess the background rejection and the signal efficiency for different primary energies [250]. In this way, the original  $M_2$  distributions play the role of probability density functions. Thus, for each energy bin, a certain number of samples



**Figure 5.20:** The background rejection in terms of the signal efficiency for different energy bins and  $\theta_{MC} = 0^\circ$ . The error bars correspond to the binomial confidence intervals according to the Clopper-Pearson method [239]. Interpolating lines are added to guide the eye.

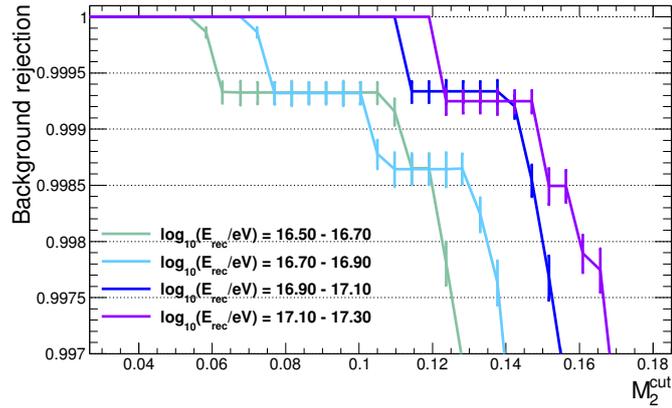
(the same for each primary to avoid a bias due to different statistics) is extracted randomly from the original distributions. Then, a sweep in values of  $M_2^{\text{cut}}$  is done between 0 and 0.7 in steps of  $10^{-3}$ . Consequently, the number of proton events verifying  $M_2^p > M_2^{\text{cut}}$  define the background rejection and the number of photon events verifying  $M_2^\gamma < M_2^{\text{cut}}$  define the signal efficiency. When  $M_2^{\text{cut}} = 0.7$ , the *bootstrapping realization* ends and another one starts with a new random selection of samples from the original distributions.

The average background rejection after  $10^5$  realizations is shown in terms of the signal efficiency in Fig. 5.20 for different reconstructed energy intervals. The error bars correspond to the binomial confidence intervals according to the Clopper-Pearson method [239]. The benchmark rejection and a signal efficiency above 95% (99%) can be achieved with relative ease at reconstructed energies between  $10^{16.5}$  eV and  $10^{16.9}$  eV (between  $10^{16.9}$  eV and  $10^{17.3}$  eV).

The step structure of the background rejection profiles in the first and second energy bin is generated by the presence of one proton event close to the photon  $M_2$  distribution in each bin. In both cases, the events have cores at the edge of the acceptance region (i.e. of the elemental hexagon area) at around 200 m. Moreover, it can be seen that the inherent Poissonian fluctuations on the number of muons for proton showers between  $E_{MC} = 10^{16.5}$  eV and  $10^{16.9}$  eV dominate over the intrinsic shower-to-shower fluctuations for distances larger than 200 m. In the case of vertical events, this effect has a minor impact, in the light of the extraordinary discrimination power shown in Fig. 5.20.

In the case of the first event at  $E_{\text{rec}} = 10^{16.65}$  eV, it shows distinctive hadronic features, such as a shallow  $X_{\text{max}} = 666 \text{ g cm}^{-2}$  and a flat LDF slope of  $-2.33$ . Thus, a complementary LDF-related observable may be enough to reject it. On the contrary, the second event at  $E_{\text{rec}} = 10^{16.84}$  eV is produced by a shower with an extremely poor muon content. Its muon density at 200 m and the total number of high energy muons<sup>5</sup> is at  $2.97\sigma$  and  $5.33\sigma$  respectively from the expected averages, which is correlated to the deep photon-like  $X_{\text{max}} = 814 \text{ g cm}^{-2}$ . The existence of muon-poor showers originated by primary protons, that could contribute to an irreducible background to the photon searches, were previously studied [251]. These showers exhibit a strong electromagnetic behaviour due to a high-energy leading  $\pi^0$  in the first interaction of the development, in particular carrying more than  $\sim 60\%$  of the primary energy.

<sup>5</sup>In this context, the high energy muons are those capable of reaching the MD, typically with kinetic energy on ground larger than  $\sim 1$  GeV, as explained in Sec. 5.2.



**Figure 5.21:** The background rejection in terms of  $M_2^{\text{cut}}$  for different energy bins and  $\theta_{\text{MC}} = 0^\circ$ . The error bars correspond to the binomial confidence intervals according to the Clopper-Pearson method [239]. Interpolating lines are added to guide the eye.

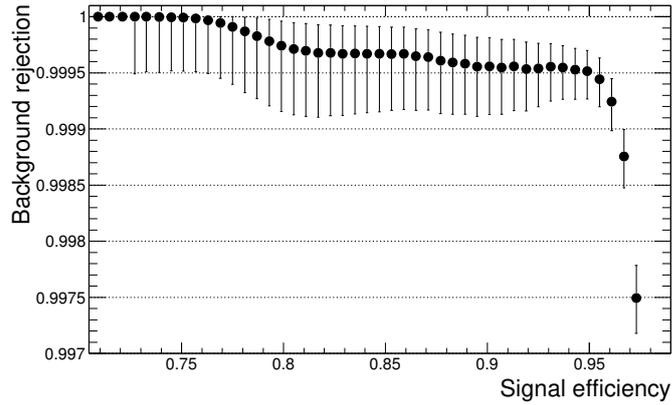
The background rejection is represented in terms of  $M_2^{\text{cut}}$  in Fig. 5.20. It can be seen that a background rejection of at least 99.90% can be reached with a constant cut on  $M_2 = 0.1$  for energies between  $E_{\text{rec}} = 10^{16.5}$  eV and  $10^{16.9}$  eV, while a more relaxed cut on  $M_2 = 0.14$  can be placed at higher energies.

In order to estimate a realistic background rejection (and its corresponding cut on  $M_2$ ) integrated over an energy range, the energy spectrum of cosmic rays must be taken into account. The rationale behind the bootstrapping procedure previously described can be applied again, but to two  $M_2$  distributions conformed by events with different energies. Prior to a bootstrapping realization, these distributions are built by selecting events with a probability proportional to  $E_{\text{rec}}^{-\gamma}$ , i.e. following a power-law spectrum as measured in data. The spectral index  $\gamma$  assuming a predominantly light primary composition was found to be  $\gamma = (3.06 \pm 0.13)$  for energies above  $E_{\text{rec}} = 10^{16.6}$  eV in [141]. The index  $\gamma$  is selected in each realization according to a Gaussian distribution with  $\mu = 3.06$  and  $\sigma = 0.13$ . The process of selecting a sample of events with a probability  $E_{\text{rec}}^{-\gamma}$  from a pool of events with different energies is referred as to *spectral shuffling*. In this way, each bootstrapping realization is performed with two  $M_2$  distributions conformed mainly by low energy events and produced by a spectral shuffling realization.

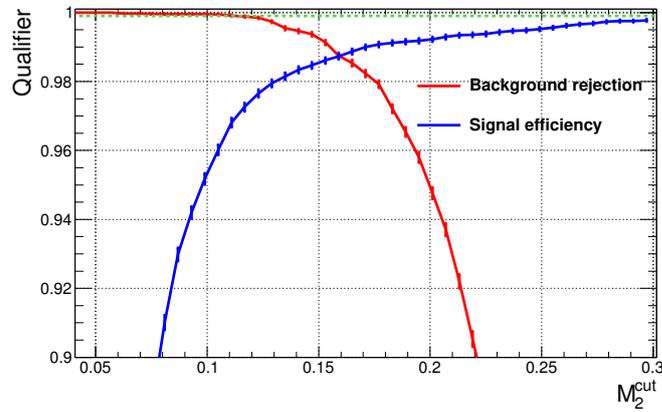
This dual procedure is repeated  $5 \times 10^4$  times. The background rejection in terms of the signal efficiency is shown in Fig. 5.22, where it can be seen that a signal efficiency of  $\sim 96\%$  is achieved at the benchmark rejection, which can be thought of as a weighted average of the estimations for low and high energy regime estimated from Fig. 5.20. The two outlier events produce again a step structure in the background rejection. The average background rejection and signal efficiency in terms of the corresponding  $M_2^{\text{cut}}$  are shown in Fig. 5.23. The benchmark can be reached at  $M_2^{\text{cut}} = 0.105$  in the spectral region between  $E_{\text{rec}} = 10^{16.5}$  eV and  $10^{17.3}$  eV.

## 5.6 Application of $M_b$ to non-vertical events

The atmospheric attenuation is stronger in EAS with an inclined incidence, which produces several features in the event trigger efficiency and the reconstructed observables. The shower front gets



**Figure 5.22:** The background rejection in terms of the signal efficiency integrated in an energy range between  $E_{\text{rec}} = 10^{16.5}$  eV and  $10^{17.3}$  eV. The events are selected according to the measured cosmic-ray spectrum with an index  $\gamma = (3.06 \pm 0.13)$ , as found in a previous study [141].

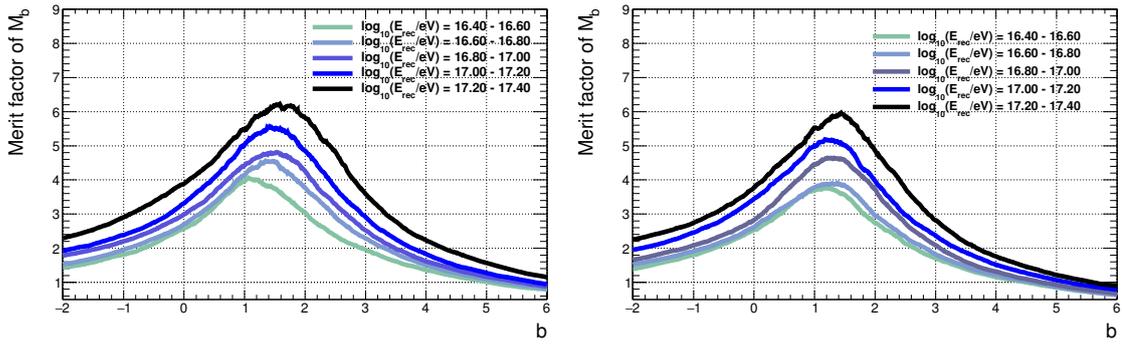


**Figure 5.23:** The background rejection (red line) and the signal efficiency (blue line) in terms of the cut on the observable  $M_2$  integrated in an energy range between  $E_{\text{rec}} = 10^{16.5}$  eV and  $10^{17.3}$  eV. The events are selected according to the measured cosmic-ray spectrum with an index  $\gamma = (3.06 \pm 0.13)$ , as found in a previous study [141]. The green line denotes the benchmark rejection level of 99.90%.

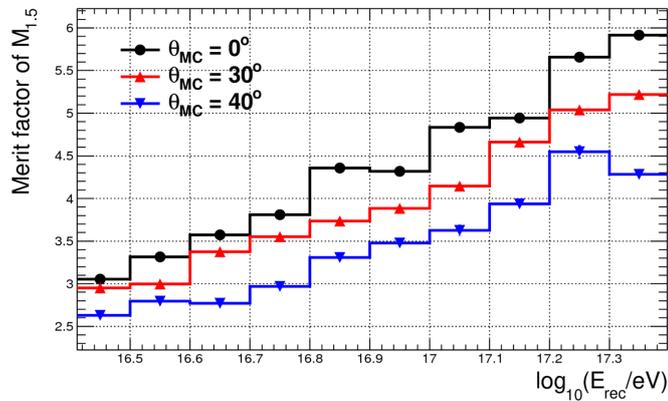
flatter as the zenith angle increases, provided that the EAS arrives to the observation level after reaching its maximum development. In the same way, the MLDF slope decreases slowly with the zenith angle (recall Fig. 5.2). Another effect to take into account is the decrease of the number of expected muons detected in the MD stations when comparing non-vertical to vertical events (recall Fig. 4.30), at least for distances close enough to the shower axis.

The foundations of the discrimination with the  $M_b$  procedure remain untouched even in the case of inclined incidence. The nearest MD station is the only one used for the  $M_b$  calculation. Only events without any low-gain saturated WCD and with its reconstructed core located in the elemental hexagon around the central MD station are considered for the discrimination procedure. The evolution of the merit factor of  $M_b$  with the value of  $b$  for non-vertical events is shown in Fig. 5.24. The maximum separation is reached, in average, for slightly smaller values of  $b$  as the zenith angle increases or the primary energy decreases, in accordance to the MLDF slope trend described before.

In view of this behaviour, an index dependent on both the energy and the zenith angle would be an efficient addition to the method. However, the first energy bin in Fig. 5.24 is affected by



**Figure 5.24:** The merit factor of the observable  $M_b$  between photon and proton primaries in terms of  $b$  for  $\theta_{MC} = 30^\circ$  (left) and  $\theta_{MC} = 40^\circ$  (right), where the line colors represent intervals of reconstructed energies ordered according to increasing energy.

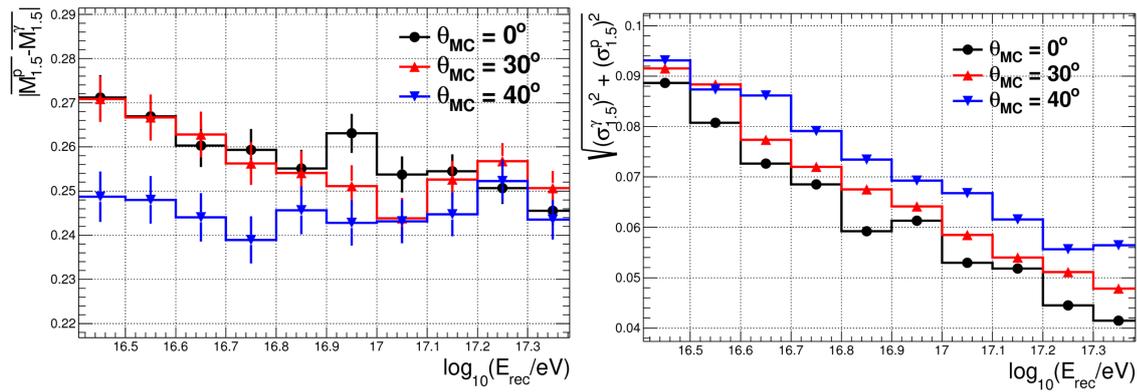


**Figure 5.25:** The merit factor of  $M_{1.5}$  between proton and photon primaries in terms of the reconstructed energy for different zenith angles, represented by marker styles.

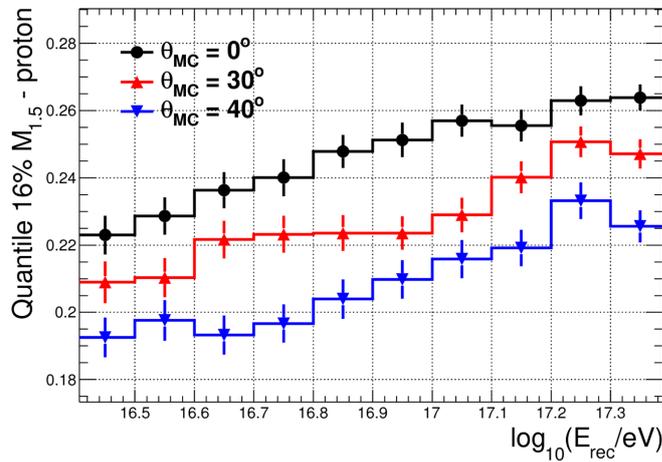
upward fluctuations because the SD array is not fully efficient, specially at  $\theta_{MC} = 40^\circ$ . Therefore, a choice of  $b < 2$  is more suitable for extending the discrimination procedure to non-vertical events. For the sake of simplicity, a value of  $b = 1.5$  is adopted for the rest of the analysis.

In Fig. 5.25 the merit factor of  $M_{1.5}$  is shown at different energies and zenith angles. By choosing a fixed value of  $b = 1.5$ , the merit factor of  $M_b$  is  $\sim 25\%$  less than the maximum possible at the first energy bin, as seen qualitatively in Fig. 5.24. However, the events in the SD-433 data below  $10^{16.6}$  eV are mainly vertical events (see Sec. 7.1.1), reducing in practice the impact of this discrepancy at higher zenith. At energies above  $E_{rec} = 10^{17.2}$  eV, the difference between the merit factor of  $M_{1.5}$  and the maximum possible decreases to  $\sim 15\%$ . From the point of view of the angular dependence, the separation decreases steadily with the zenith angle. This is the consequence of a two-fold effect: the decreasing difference of muon content between protons and photons and, mainly, the increasing fluctuations of the observable, as displayed in Fig. 5.26. The width of the  $M_{1.5}$  for proton primaries (which dominate the fluctuations term in the merit factor calculation) grows asymmetrically as the zenith angle increases, due to the enhanced presence of events with a lower muon content. This is translated in a decrease of the 16% quantile of the  $M_{1.5}$  distribution as seen in Fig. 5.27, which in turn impacts on the increase of the observable fluctuations (see Eq. 5.7).

The second aspect to take into account, linked with the previous discussion, is the impact of the Poissonian fluctuations on the separation power. Because of the reduced number of muons



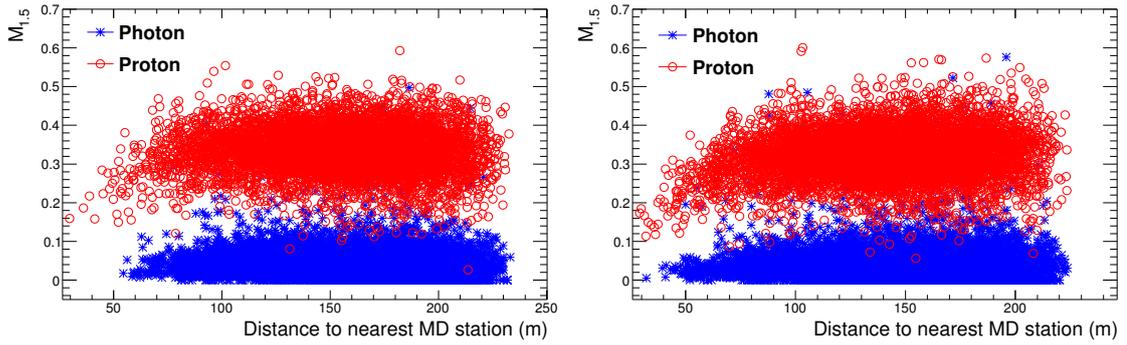
**Figure 5.26:** (Left) The absolute difference between the median of  $M_{1.5}$  for photon and proton primaries in terms of the reconstructed energy. (Right) The sum of the fluctuations of  $M_{1.5}$  for photon and proton primaries in terms of the reconstructed energy. The different MC zenith angles are represented by marker styles.



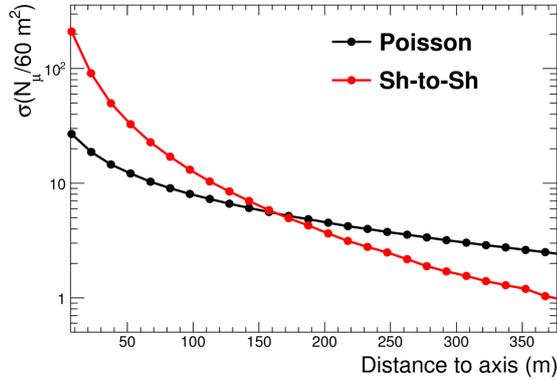
**Figure 5.27:** The evolution of the 16% quantile of the  $M_{1.5}$  distribution for proton primaries in terms of the reconstructed energy. The different MC zenith angles are represented by marker styles.

that reach the observation level in the case of inclined incidence, the distance constraint needs to be reviewed. This is displayed in Fig. 5.28 for  $\theta_{MC} = 30^\circ$  and  $\theta_{MC} = 40^\circ$ . Two kinds of outliers can be found in these plots. The first kind is exclusively due to the Poisson uncertainty in the low number of muons and it is present at large distances. As an example, it can be seen in Fig. 5.29 that a threshold of  $\sim 200$  m can be imposed in order to avoid a large impact from this systematic effect. Therefore, an additional quality cut, which imposes a more stringent distance requirement, comes out from this discussion: the nearest MD station needs to be at most 200 m away from the shower axis. It is worth mentioning that this is not a replacement of the original elemental hexagon condition, but a refinement of it that can be thought of as pruning the vertexes of the elemental hexagon. Therefore, this requirement is imposed from now on.

After applying this extra condition, the second kind of outlier can be addressed. They have a physical origin and, as in the vertical case, impose a limitation to the sensitivity of the  $M_b$  discrimination method. These proton outliers have several features similar to the photon events, as it is seen in Fig. 5.30 for both  $\theta_{MC} = 30^\circ$  and  $\theta_{MC} = 40^\circ$  together. Most of the proton outliers are compatible with a deep photon-like shower development. Another remarkable detail is that most



**Figure 5.28:**  $M_{1.5}$  in terms of the distance between the nearest MD station (used for its calculation) and the shower axis. Photon and proton events are symbolized by blue stars and red circumferences, respectively. The selected events have reconstructed energies between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV and (left)  $\theta_{\text{MC}} = 30^\circ$  or (right)  $\theta_{\text{MC}} = 40^\circ$ .



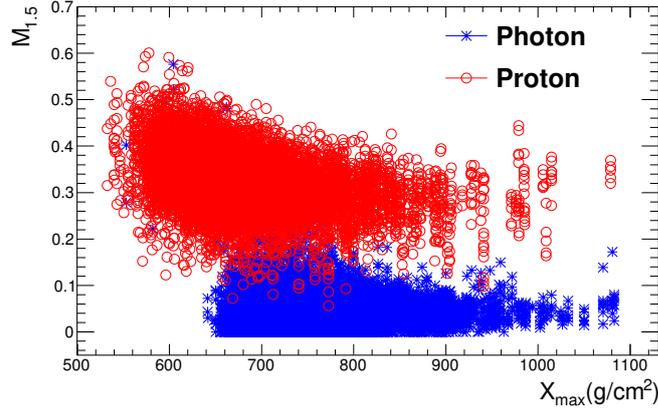
**Figure 5.29:** The uncertainty contributions in the number of muons per  $60 \text{ m}^2$  at  $2.25 \text{ m}$  depth for proton showers with MC energy between  $E_{\text{MC}} = 10^{16.6}$  eV and  $10^{16.7}$  eV. The inherent Poissonian fluctuations due to the finite sampling area are represented by black dots. The shower-to-shower fluctuations are represented by red dots.

of the separation between the two samples comes from the  $M_{1.5}$ , considering that  $X_{\text{max}}$  is one of the most powerful observable in terms of separation in all current photon searches [114].

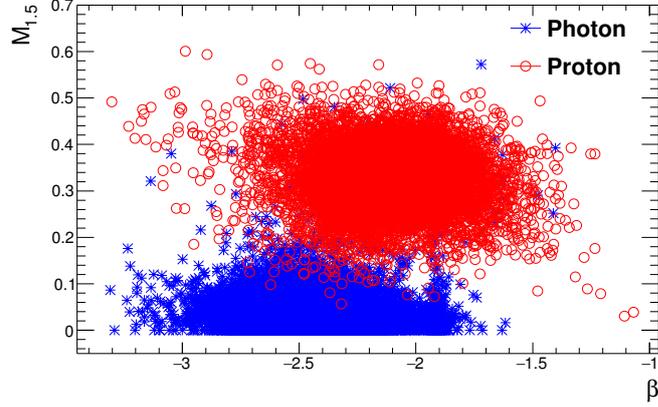
At the same time, these outlier events present a LDF slope that may serve as a hint to distinguish them from real photon events, as it can be seen in Fig. 5.31. However, a few difficulties are readily spotted. The minimum reconstructed energy at which a truthful LDF reconstruction can be achieved is higher. Moreover, the events showed in Fig. 5.31 have  $E_{\text{rec}} > 10^{16.6}$  eV, a higher limit than in Fig. 5.30. Even in this case, the photon events around a value of  $\beta \sim -1.9$  are actually fitted with the usual fixed slope model since the number of triggered stations is not large enough. Apart from these issues, the LDF may add some valuable information to complement  $M_b$ . This approach is discussed in the next chapter.

### 5.6.1 Background rejection and signal efficiency

After extending the  $M_b$  procedure to non-vertical events, the background rejection and signal efficiency in this scenario can be addressed. It is worth noting that the usage of  $b = 1.5$  instead of



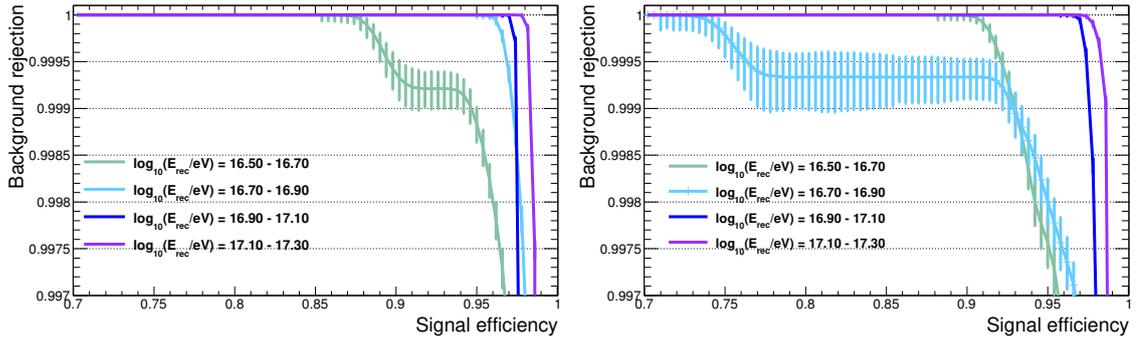
**Figure 5.30:**  $M_{1.5}$  in terms of the  $X_{\max}$  of the shower originating the events, after imposing the requirement for the MD station to be nearer than 200 m from the shower axis. The events are selected according to their reconstructed energy between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV. Events with  $\theta_{\text{MC}} = 30^\circ$  and  $\theta_{\text{MC}} = 40^\circ$  are shown together.



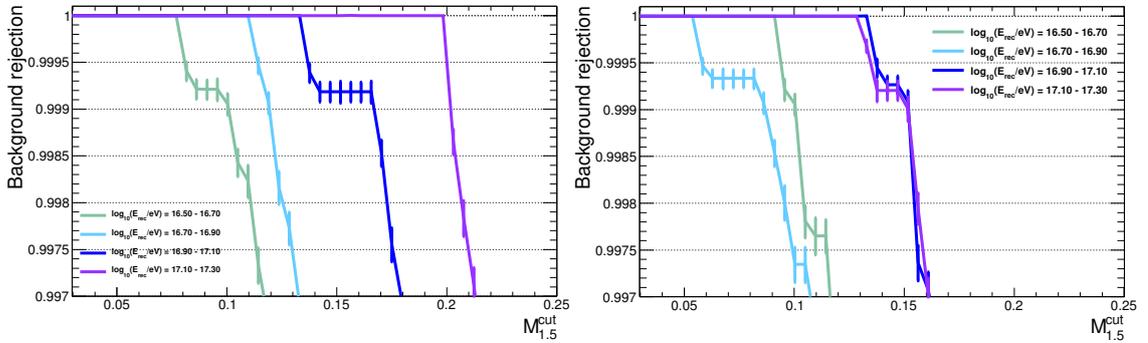
**Figure 5.31:**  $M_{1.5}$  in terms of the  $\beta$ , after imposing the requirement for the MD station to be nearer than 200 m from the shower axis. The events are selected according to their reconstructed energy between  $E_{\text{rec}} = 10^{16.6}$  eV and  $10^{17.4}$  eV. Events with  $\theta_{\text{MC}} = 30^\circ$  and  $\theta_{\text{MC}} = 40^\circ$  are shown together.

$b = 2$  and the additional quality cut of  $r_{\text{nearest}} < 200$  m does not change the overall picture of the studies presented in the vertical case in Sec. 5.5.1. As a matter of fact, one of the proton outliers in that analysis corresponded to an event landing at  $\sim 210$  m from the shower axis.

For completeness, and following the same procedure as explained before in Sec. 5.5.1, the background rejection in terms of the signal efficiency for  $\theta_{\text{MC}} = 30^\circ$  and  $\theta_{\text{MC}} = 40^\circ$  are shown in Fig. 5.32. In the first (second) case, a signal efficiency of at least 95% (93%) can be reached at the benchmark rejection. Just as in the vertical case, the step structure is caused by single outliers in the samples. In particular, the early drop of the background rejection for the second energy bin (cyan line) for  $\theta_{\text{MC}} = 40^\circ$  is produced by an event which already appeared in Fig. 5.30 and Fig. 5.31:  $M_{1.5} = 0.056$ ,  $\beta \simeq -2.3$  and  $X_{\max} = 772 \text{ g cm}^{-2}$ . The background rejection is represented in terms of  $M_{1.5}^{\text{cut}}$  in Fig. 5.33 for both zenith angles, where it can be seen that the benchmark is reached with an energy-independent cut on  $M_{1.5} = 0.1$  for energies between  $E_{\text{rec}} = 10^{16.5}$  eV and  $10^{17.3}$  eV in the case of  $\theta_{\text{MC}} = 30^\circ$ . However, due to the outlier mentioned before, a cut on  $M_{1.5} = 0.08$  is needed in the case of  $\theta_{\text{MC}} = 40^\circ$ . The more stringent cut on the observable  $M_b$  needed for events



**Figure 5.32:** The background rejection in terms of the signal efficiency for different energy bins and (left)  $\theta_{MC} = 30^\circ$  or (right)  $\theta_{MC} = 40^\circ$ . The error bars correspond to the binomial confidence intervals according to the Clopper-Pearson method [239]. Interpolating lines are added to guide the eye.

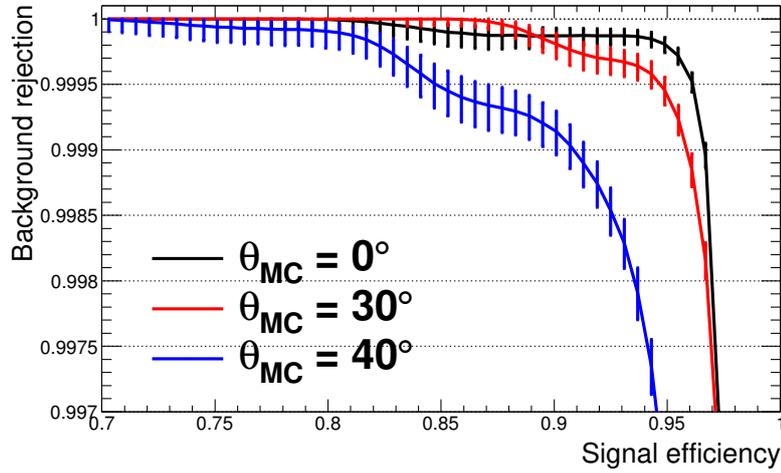


**Figure 5.33:** The background rejection in terms of  $M_{1.5}^{cut}$  for different energy bins and with MC zenith angles of (left)  $\theta_{MC} = 30^\circ$  or (right)  $\theta_{MC} = 40^\circ$ . The error bars correspond to the binomial confidence intervals according to the Clopper-Pearson method [239]. Interpolating lines are added to guide the eye.

at  $\theta_{MC} = 40^\circ$ , and thus a smaller signal efficiency at the benchmark, is a systematic behaviour that appears once and again in several analyses, pointing out the impact of the decreasing number of average muons in the MD stations at the energies and distances of interest for this photon search.

As it is clearly deduced from Fig. 5.33, the signal efficiency provided by a constant cut in the observable  $M_{1.5}$  at all energies can be enhanced by including an energy and zenith dependence in the calculation, i.e. designing a  $M_{1.5}^{cut}(E, \theta)$ . This sophisticated approach is developed within the framework of the multivariate analysis, which is vastly discussed in Chapter 6.

The spectral shuffling is again summoned to merge events with different primary energies in a realistic way. After  $5 \times 10^4$  realizations with events between  $E_{rec} = 10^{16.4}$  eV and  $10^{17.4}$  eV, a signal efficiency of 97%, 95% and 91% can be attained at the benchmark rejection for the fixed zenith angles of  $\theta_{MC} = 0^\circ$ ,  $\theta_{MC} = 30^\circ$  and  $\theta_{MC} = 40^\circ$ , as shown in Fig. 5.34. As mentioned before, the change of the index  $b$  and the additional distance constraint in the case of vertical events do not lead to any significant change in the background rejection capabilities of the  $M_b$  reflecting the robustness of the observable. The values for  $M_{1.5}^{cut}$  to reach the benchmark for energies following a spectral distribution are, in each case, 0.14, 0.11 and 0.09. The relevant quantities of the discrimination with  $M_b$  at fixed zenith angles are summarized in Tab. 5.1.



**Figure 5.34:** The background rejection in terms of the signal efficiency integrated in an energy range between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV. The events are selected according to the measured cosmic-ray spectrum with an index  $\gamma = (3.06 \pm 0.13)$ , as found in a previous study [141]. The events have the fixed MC zenith angles as quoted.

$\theta_{\text{MC}} (^{\circ})$	Efficiency @ 99.90% rejection	$M_{1.5}^{\text{cut}}$
0	97%	0.14
30	95%	0.11
40	91%	0.09

**Table 5.1:** The signal efficiency at the benchmark rejection and the corresponding value of  $M_{1.5}^{\text{cut}}$  for events with the quoted zenith angles. The events are selected according to the reconstructed energy between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV following the measured cosmic-ray spectrum, as found in a previous study [141]. Only unsaturated events with a reconstructed core position lying within the elemental hexagon are considered.

## Chapter 6

# Multivariate analysis combining SD and MD

Genius is one percent inspiration,  
ninety-nine percent perspiration.

---

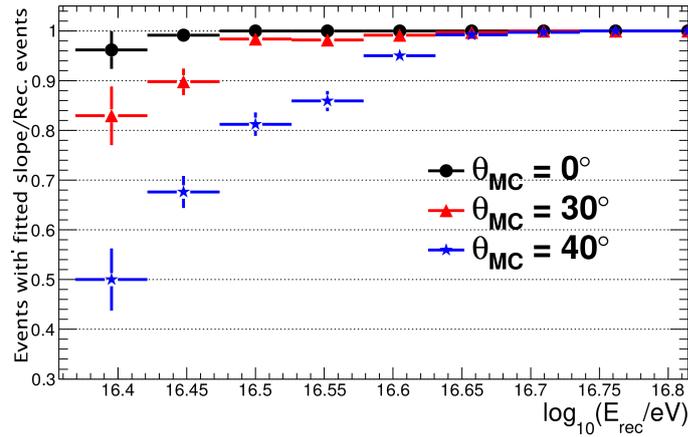
Thomas A. Edison

The muon content of the EAS, quantified by the observable  $M_b$ , successfully discriminates between photon and proton primaries with an unprecedented efficiency at the benchmark background rejection of 99.90%. Complementary information can be extracted from the electromagnetic component of the EAS as well, that in turn defines the shape of the LDF, at least close to the shower axis. As discussed in Sec. 5.6, some background events that are compatible with a photon primary from the point of view of  $M_{1.5}$  may be exposed by a hadronic-like LDF. Although the bulk of the separation power actually comes from  $M_{1.5}$ , this approach may lead to a valuable complementary tool to reject, at least, a part of the background contamination.

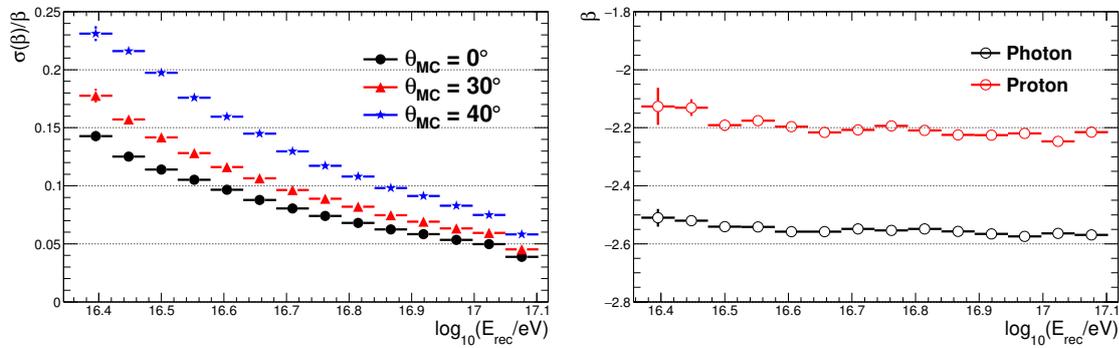
The event reconstruction procedure is responsible for the estimation of the core position, the arrival direction, the event energy and, optionally, the LDF slope. While the official choice is to perform the event reconstruction with a fixed LDF slope (i.e. using an a priori parametrization of the slope in terms of the other reconstructed parameters, see Sec. 3.3.4), if an event has enough triggered WCDs it is possible to add the LDF slope in the fitting procedure. This requirement is easily fulfilled for energies above  $E_{\text{rec}} = 10^{16.7}$  eV, as inferred from Fig. 6.1 for proton primaries (the same behaviour is seen for photon primaries). However, the event footprint may not be large enough to perform a LDF slope reconstruction in an event-by-event basis at lower energies and higher zenith angles.

In principle, the restriction imposed by the fitted LDF slope requirement may be included in the discrimination procedure as a quality cut like the non-saturation condition described before. However, even for events below  $E_{\text{rec}} = 10^{16.7}$  eV, the LDF slope is fitted with an increasing uncertainty, as seen in Fig. 6.2, left, for proton primaries (similar results are obtained for photon primaries). Unfortunately, the uncertainty at these energies is similar to the expected average separation between photon and proton events, as displayed in Fig. 6.2, right, for the specific case of  $\theta_{\text{MC}} = 30^\circ$ .

Nevertheless, the LDF shape has attracted much attention in the context of the photon/hadron discrimination quest during the last years. Particularly, a quantity arising from the expected



**Figure 6.1:** The ratio between the number of events for which it is possible to fit the LDF slope and the total reconstructed events. The events are initiated by proton primaries.



**Figure 6.2:** (Left) The relative uncertainty of the fitted LDF slope for proton-initiated events in terms of the reconstructed energy. Events with different MC zenith angles are represented by marker styles. (Right) The average fitted LDF slope in terms of the reconstructed energy for photon (black) and proton (red) primaries at a fixed MC zenith angle of  $\theta_{MC} = 30^\circ$ .

energy offset in the case of photon events with respect to hadronic events has been extensively explored. This observable, called  $F_\gamma$ , compares the SD signal at a fixed distance of 1000 m expected for a photon primary,  $S_{1000|\gamma}$ , with the energy estimator coming from the standard reconstruction procedure,  $\langle S_{1000} \rangle (E_\gamma, \theta)$  [252].  $F_\gamma$  has been applied in the data analysis of hybrid events above  $10^{18}$  eV, since it requires an unbiased estimation of the primary energy  $E_\gamma$  (i.e. a calorimetric measurement provided by the FD) [114]. The parameter  $S_{1000|\gamma}$  is obtained by fitting the events with a dedicated parametrization of the LDF slope for photon primaries above  $10^{18}$  eV.

The underlying idea the observable  $F_\gamma$  is that of a comparison between two different models: one specially designed to describe photon-initiated EAS and an equivalent one for hadronic primaries. In this sense, a model describing the LDF slope in photon events is expected to work better for photon- than proton-initiated EAS. On the contrary, the standard LDF slope parametrization is thus expected to fit better hadronic events than photon events. In the first part of this chapter, the observable  $Q$  is defined and developed around this principle. In the second part, a multivariate analysis combining  $M_b$  with  $Q$  is presented, which leads to the completion of the discrimination procedure developed in this thesis.

## 6.1 Extracting information from the LDF

The fitting of the LDF is accomplished by maximizing a likelihood function. Given a set of triggered WCDs represented by an index  $i$ , the signals  $S_i$  measured by each of them at a distance  $r_i$  are the main ingredients in this maximization. The likelihood function can be written as

$$L = \prod_i f_P(n_i, \mu_i) \times \prod_i f_G(n_i, \mu_i) \times \prod_i F_{\text{sat}}(n_i, \mu_i) \times \prod_i F_{\text{zero}}(n_i, \mu_i) \quad (6.1)$$

and thus, the log likelihood function can be written as

$$l = \sum_i \ln f_P(n_i, \mu_i) + \sum_i \ln f_G(n_i, \mu_i) + \sum_i \ln F_{\text{sat}}(n_i, \mu_i) + \sum_i \ln F_{\text{zero}}(n_i, \mu_i) \quad (6.2)$$

with  $n_i$  the effective number of particles detected in the WCD  $i$  and  $\mu_i$  the corresponding theoretical expectation. Each term of Eq. 6.2 describes the contributions from WCDs with specific statuses:  $f_P$  is the Poissonian probability density function describing small signals,  $f_G$  is the Gaussian probability function describing large signals,  $F_{\text{sat}}$  describes the saturated signals by means of a complementary error function and  $F_{\text{zero}}$  describes the expectation from silent WCDs through a sum of Poisson probability functions [187]. The theoretical expectation  $\mu_i$  is calculated by means of the LDF slope model provided through the `Offline` configuration files.

By itself, the value of the likelihood function at its maximum,  $L_{\text{max}}$ , does not have any physical meaning. However, the ratio of  $L_{\text{max}}$  obtained in two different scenarios is a feasible way to quantify how much the observed data  $n_i$  is compatible with the model prediction  $\mu_i$  in each scenario. In other words, this ratio is a measure of the goodness-of-fit of the models to the data. In the context of the photon search, and following the discussion on  $F_\gamma$ , the fixed LDF slope model for photons and hadronic events would be the contending models in the aforementioned comparison. If the maximum value of the likelihood function in each case is, respectively,  $L_\gamma$  and  $L_h$ , the observable  $Q$  defined as:

$$Q = \log_{10} \left( \frac{L_\gamma}{L_h} \right) \quad (6.3)$$

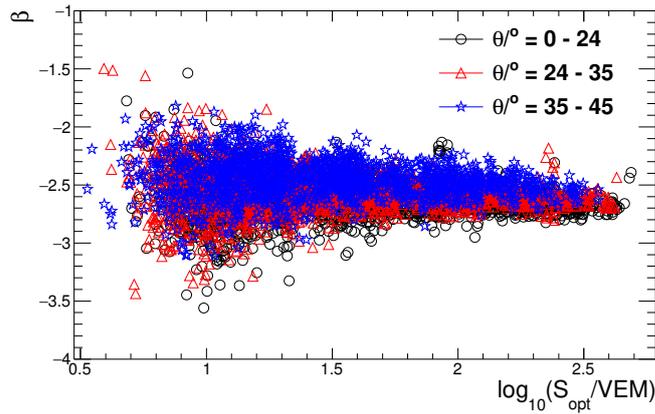
measures how much an event is compatible with a photon origin hypothesis with respect to the hadronic origin. An event yielding a negative value for  $Q$  is then interpreted as to be more compatible with a hadronic LDF model than with a photon dedicated one. In this scenario, the null hypothesis (being “the observed event has a hadronic origin”) can not be rejected.

### 6.1.1 Parameterization of the photon LDF slope

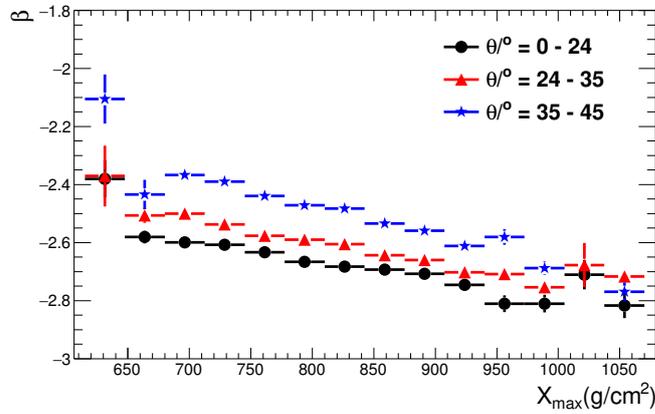
The first step towards the implementation of the observable  $Q$  in the discrimination problem is the determination of the photon-dedicated LDF model. The hadronic LDF model dependence, as explained in Sec. 3.3.4, is also assumed for photon primaries:

$$\beta(S_{\text{opt}}, \theta) = a_0(\theta) + a_1(\theta) \times \log_{10}(S_{\text{opt}}) \quad (6.4)$$

The events considered for this study have their shower cores randomly tossed in the central hexagon of the SD array in order to accomplish an unbiased sampling of the shower footprint. On the contrary, if the events considered for the parametrization were tossed exclusively in the



**Figure 6.3:** The fitted LDF slope in terms of the energy estimator  $S_{\text{opt}}$  for photon events with MC energies between  $E_{\text{MC}} = 10^{16}$  and  $10^{18}$  eV and zenith angles in the quoted intervals, represented by different markers.

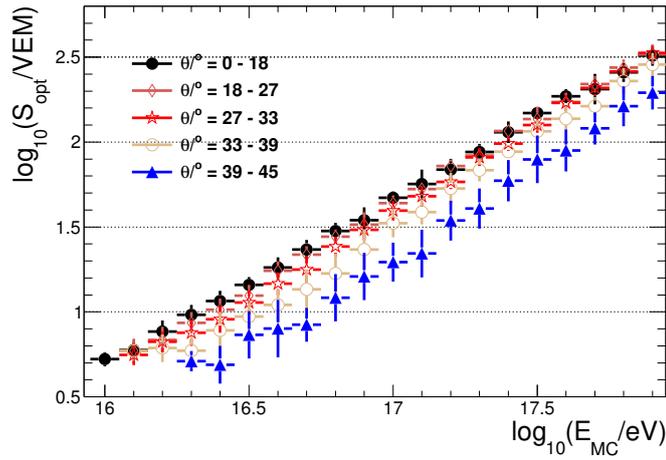


**Figure 6.4:** The fitted LDF slope in terms of the depth of the shower maximum for photon events with MC energies between  $E_{\text{MC}} = 10^{16}$  and  $10^{18}$  eV and zenith angles in the quoted intervals, represented by different markers.

elemental hexagon, as required by the  $M_b$  discrimination procedure, certain types of footprints (i.e. collections of  $(r_i, S_i)$ ) would be favored. Additionally, saturated events are not considered for the analysis, not only due to the  $M_b$  requirement, but also due to the artifacts of the reconstruction in this case (see Sec. 4.4). Lastly, the reconstruction is made with an ideal SD array as used for the simulation (recall Sec. 4.2.2) to avoid any artifacts, specially at high energies, from the incompleteness of the second crown of WCDs.

The parameter  $\beta$  is shown in terms of  $S_{\text{opt}}$  in Fig. 6.3 for events with MC energies between  $E_{\text{MC}} = 10^{16}$  eV and  $10^{18}$  eV, with MC zenith angles following a  $\sin\theta \cos\theta$  distribution. A small fraction of events are visibly departed from the mean behaviour (e.g., the vertical events at  $\log_{10}(S_{\text{opt}}) \simeq 1.9$  and  $\beta \simeq -2.2$ ). This kind of events are produced by EAS with an unusual high muon content, thus having a proton-like development. In turn, they can be spotted by looking at the depth of the maximum development in Fig. 6.4. The first two bins are populated by these shallow showers. As the shower maximum is reached at larger depths, the LDF slope increases in absolute value.

It can be seen that  $\beta$  shows a weak dependence on  $S_{\text{opt}}$ , at least for the considered angles  $\theta_{\text{MC}} < 45^\circ$ . At low energies, the fluctuations of  $\beta$  are increased due to the compound effect from the



**Figure 6.5:** The average  $S_{\text{opt}}$  in terms of the MC energy for photon events at the quoted MC zenith intervals.

trigger efficiency and the high uncertainty of the fit, as mentioned before. On the other hand, the constraint on the event saturation status may bring a bias at high energies. Indeed, the saturated events tend to have a flatter LDF than unsaturated events due to the lever arm effect imposed by the underestimation of the signal at the nearest WCD to the shower axis. Therefore, when these events are filtered out, the steeper LDFs are favoured.

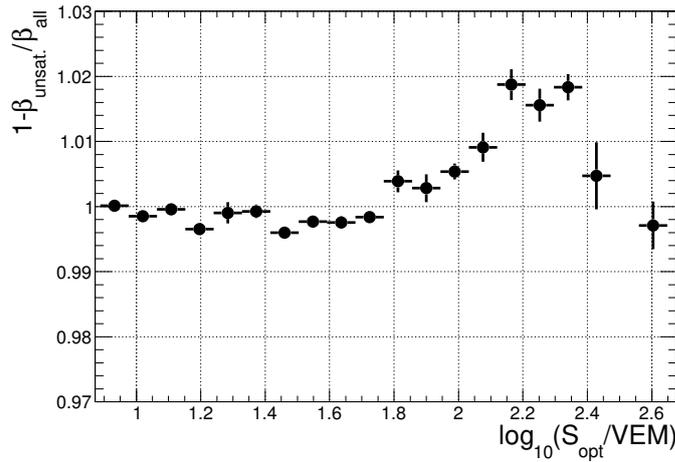
Although a fit of the functional form in Eq. 6.4 may seem straight-forward, a selection of the  $S_{\text{opt}}$  interval with some criteria must be made in order to reduce the impact of the two aforementioned bias sources. To reduce the trigger efficiency effect at low energies, the same condition as in the  $\rho_{\mu}^{\text{ref}}$  calculation is imposed (this time, for photon primaries). Events with a MC energy for which the SD array has a trigger probability less than 90% are discarded. Qualitatively, this can be translated in a minimum required  $S_{\text{opt}}$  as displayed in Fig. 6.5. Considering the trigger efficiency studies presented in Sec. 4.3.1, a constraint of  $\log_{10}(S_{\text{opt}}) = 1$  is a fair choice.

On the other hand, the upper boundary of the interval of  $S_{\text{opt}}$  to perform the parametrization is defined by the absence of saturated events, as mentioned before. The bias between the average LDF slope for unsaturated events and all events is shown in Fig. 6.6. Indeed, the saturation effect kicks in at  $\log_{10}(S_{\text{opt}}) \simeq 1.8$  and is responsible for the increase of the bias above 1% at  $\log_{10}(S_{\text{opt}}) \simeq 2.1$ . Thus, this value is imposed as the upper boundary of the slope parametrization, which in practical terms means that it is optimized for energies up to  $E_{\text{rec}} \simeq 10^{17.4}$  eV and  $E_{\text{rec}} \simeq 10^{17.7}$  eV in events with  $\theta_{\text{MC}} < 18^\circ$  and  $39^\circ < \theta_{\text{MC}} < 45^\circ$ , respectively.

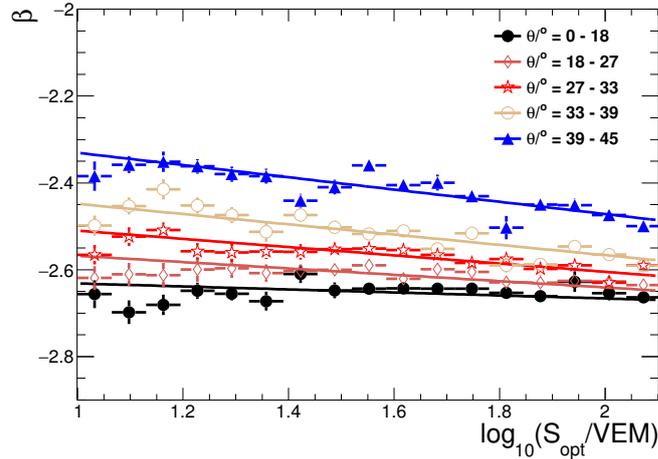
Having set the interval of  $S_{\text{opt}}$  for the parametrization, the linear fits presented in Eq. 6.4 can be performed, as shown in Fig. 6.7. The values for the coefficients of the linear dependence are shown in Tab. 6.1 with the associated  $\chi_{\text{red}}^2$  in each zenith interval. Finally, the angular dependence of the coefficients is fitted in Fig. 6.8, leading to:

$$\alpha_0(\sec\theta) = (-3.79 \pm 0.08) + (1.18 \pm 0.07) \times \sec\theta \quad (6.5)$$

$$\alpha_1(\sec\theta) = (0.258 \pm 0.042) - (0.299 \pm 0.036) \times \sec\theta \quad (6.6)$$



**Figure 6.6:** The bias between the average LDF slope in unsaturated events against that obtained from all events in terms of  $S_{\text{opt}}$ . The events have MC zenith angles between  $\theta_{\text{MC}} = 0^\circ$  and  $\theta_{\text{MC}} = 45^\circ$  following a  $\sin\theta\cos\theta$  distribution.



**Figure 6.7:** The average  $\beta$  in terms of  $S_{\text{opt}}$  and the corresponding linear fits according to Eq. 6.4.

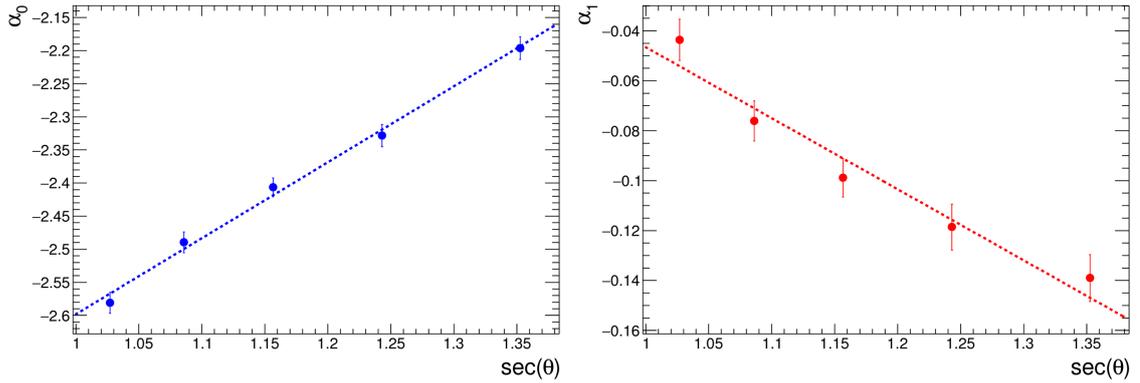
## 6.1.2 Testing the photon LDF model

The predicted slope from the parametrization,  $\beta_{\text{model}}$ , can be compared with the fitted LDF slope from the reconstruction procedure  $\beta_{\text{MC}}$ . In particular, consistency checks can be done by comparing both parameters for the same set of events used to obtain the parametrization. Fig. 6.9, left, shows the bias between  $\beta_{\text{model}}$  and  $\beta_{\text{MC}}$  in terms of the reconstructed energy for  $\theta_{\text{MC}} < 45^\circ$ . Overall, the bias is well contained within a 1% band. The average bias along all the energies of interest is  $\sim 0.1\%$ . There is a slight increase of the bias at low energies, which is conformed by non-vertical events, as deduced from Fig. 6.9, right. Indeed, those events present biases up to  $\sim 2\%$ . However, events at  $\theta > 40^\circ$  and  $E_{\text{rec}} < 10^{16.6}$  eV are statistically uncommon, as it is described in Sec. 7.1.1.

The events considered for the parametrization and for the consistency checks discussed before were tossed in a rectangular area containing the central hexagon of the SD array. However, the discrimination procedure requires a smaller area, particularly the elemental hexagon around the central position. Therefore, an independent set of events with fixed MC zenith angles were

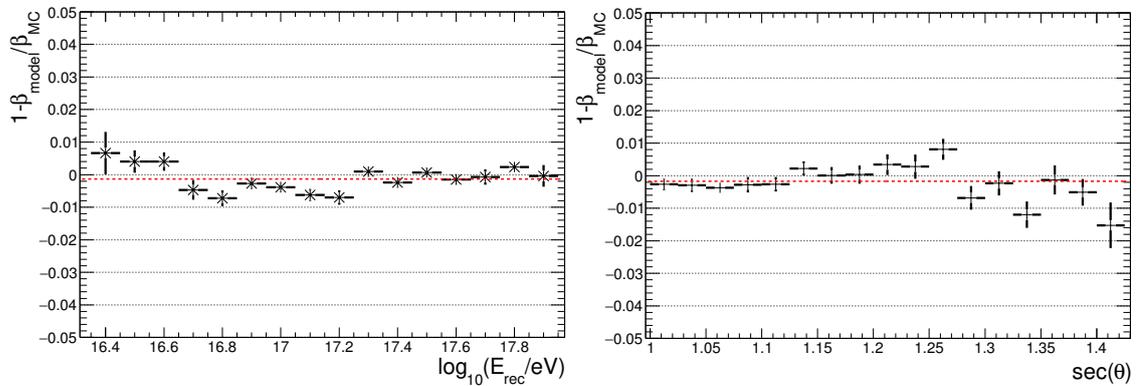
Zenith interval ( $^{\circ}$ )	$\alpha_0$	$\alpha_1 \times 10^2$	$\chi_{\text{red}}^2$
0 – 18	$-2.58 \pm 0.02$	$-4.37 \pm 0.82$	2.01
18 – 27	$-2.49 \pm 0.02$	$-7.61 \pm 0.80$	1.14
27 – 33	$-2.41 \pm 0.01$	$-9.89 \pm 0.77$	1.96
33 – 39	$-2.33 \pm 0.02$	$-11.9 \pm 0.92$	2.10
39 – 45	$-2.20 \pm 0.02$	$-13.9 \pm 0.94$	2.38

**Table 6.1:** The fitted values for the linear coefficients of Eq. 6.4 in each zenith interval.

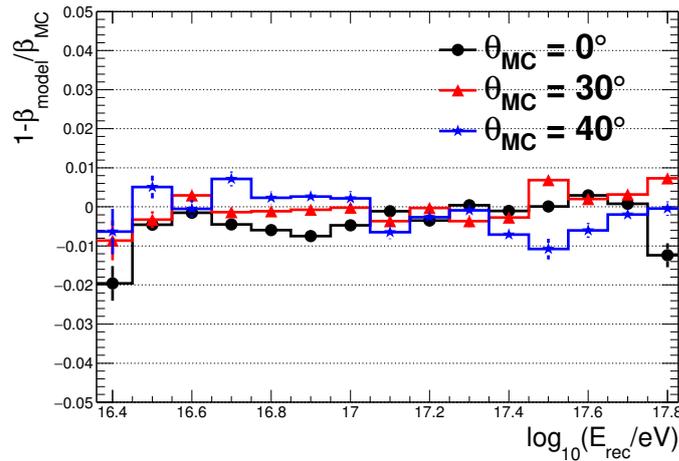


**Figure 6.8:** The linear coefficients of Eq. 6.4 in terms of the  $\sec\theta$ . The superimposed dotted lines represent linear fits with  $\chi_{\text{red}}^2 = 0.71$  and  $\chi_{\text{red}}^2 = 1.28$  respectively.

produced to test the parametrization in this scenario. The corresponding biases are shown in Fig. 6.10. The results are consistent with the previous discussion. In the vertical case, the presence of the shallow proton-like showers mentioned before appear as a bias of  $\sim 2\%$  in the first energy bin. A similar bias is seen at the last energy bin due to the opposite case of deeply-penetrating showers. It is worth noting that, even in these extreme cases, the average bias is one order of magnitude less than the expected difference in the LDF slope between photon and proton events. Thus, the bias of the presented parametrization has a negligible impact on the separation power.



**Figure 6.9:** The bias between the predicted LDF slope,  $\beta_{\text{model}}$ , and the fitted LDF slope,  $\beta_{\text{MC}}$ , in terms of the reconstructed energy (left) and the zenith angle (right). The red lines represent the average between all bins.



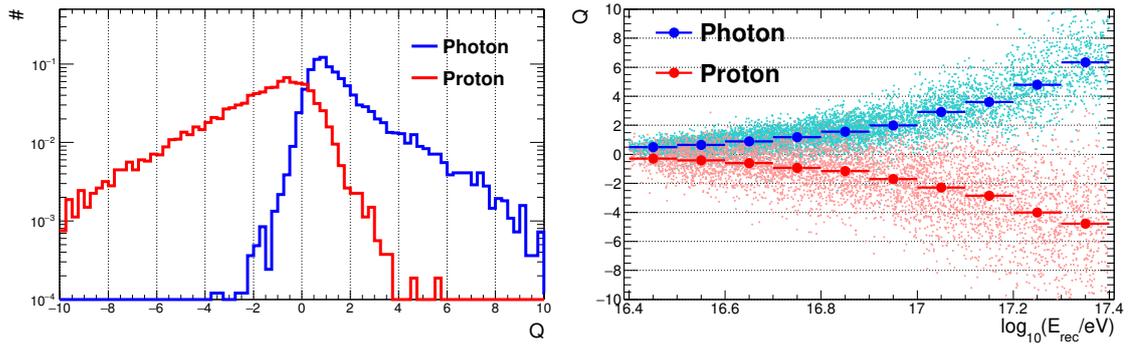
**Figure 6.10:** The bias between the predicted LDF slope,  $\beta_{\text{model}}$ , and the fitted LDF slope,  $\beta_{\text{MC}}$ , in terms of the reconstructed energy for different fixed MC zenith angles, as quoted. The events produced by EAS impinging on the elemental hexagon, as required by the  $M_b$  discrimination procedure.

### 6.1.3 Performance of the observable $Q$

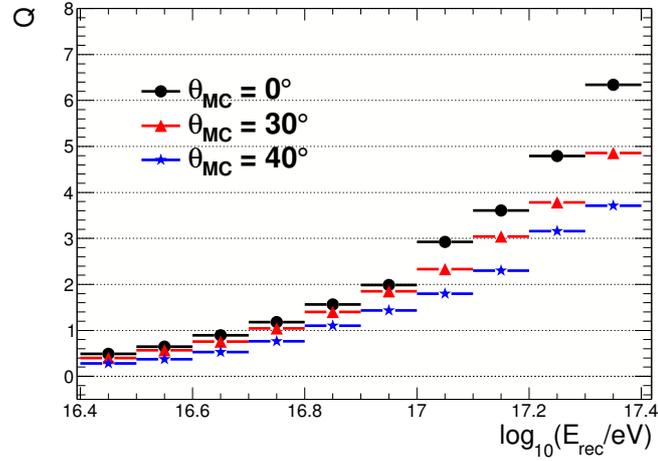
The event reconstruction is now performed thrice for each event, regardless of the primary particle type: with  $\beta$  as a free parameter, with a fixed  $\beta$  from the (standard) hadronic model and with a fixed  $\beta$  from the dedicated photon model. However, the most suitable reconstruction is the only one taken into account to extract parameters such as the reconstructed energy, the geometry or impose the necessary quality cuts. In other words, the reconstruction with the fixed LDF slope by the photon (hadron) model is taken as the reference when analysing signal (background) events.

The quality cuts coming from the  $M_b$  discrimination procedure, namely the elemental hexagon with  $r_{\text{max}} < 200$  m requirement plus the non-saturation event condition, are also applied to the analysis with the observable  $Q$  for consistency. The  $Q$  distributions for photon and proton primaries with a fixed  $\theta_{\text{MC}} = 0^\circ$  and energies between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.5}$  eV are shown in Fig. 6.11, left. Both the photon-initiated EAS with a high muon content and the proton-initiated EAS with an early electromagnetic-dominated development are clearly seen as outliers. In general, negative values of  $Q$  are expected for hadronic primaries and viceversa. In Fig. 6.11, right, the average  $Q$  for both types of primaries are shown in terms of the energy for  $\theta_{\text{MC}} = 0^\circ$ . While  $Q$  is considered in this photon search as a secondary observable (i.e. empowering the discrimination of  $M_b$  in a complementary fashion), the separation offered by  $Q$  itself increases with the reconstructed energy, meaning that  $Q$  may be a feasible observable to greatly enhance the photon search at energies above  $10^{17}$  eV. The observable  $Q$  is dependent on the zenith angle, as displayed in Fig. 6.12 for photon events. A similar behaviour is seen for the case of proton primaries, albeit with the opposite sign. However, for the energies of interest for this photon search, the dependence is rather weak.

The additional information provided by the observable  $Q$  can be combined with the observable  $M_b$  resulting in a novel combined analysis between the SD and the MD, as shown in Fig. 6.13 for events with  $\theta_{\text{MC}} = 30^\circ$ . The background events with  $M_{1.5} \sim 0.1$  may well be rejected by considering the value of  $Q$ . This can be represented by a diagonal cut in the  $M_{1.5} \cup Q$  space. It can be seen qualitatively that this kind of multidimensional cut may not enhance in a notorious way the background rejection. Nevertheless, the signal efficiency could be increased for a fixed



**Figure 6.11:** (Left) The  $Q$  distributions for photon and proton events with  $E_{\text{rec}} > 10^{16.4}$  eV, normalized to the number of entries. (Right) The mean values of  $Q$  in terms of the reconstructed energy for photon and proton events. In both cases, photon (proton) events are represented by blue (red) lines or markers. All events have  $\theta_{\text{MC}} = 0^\circ$ .



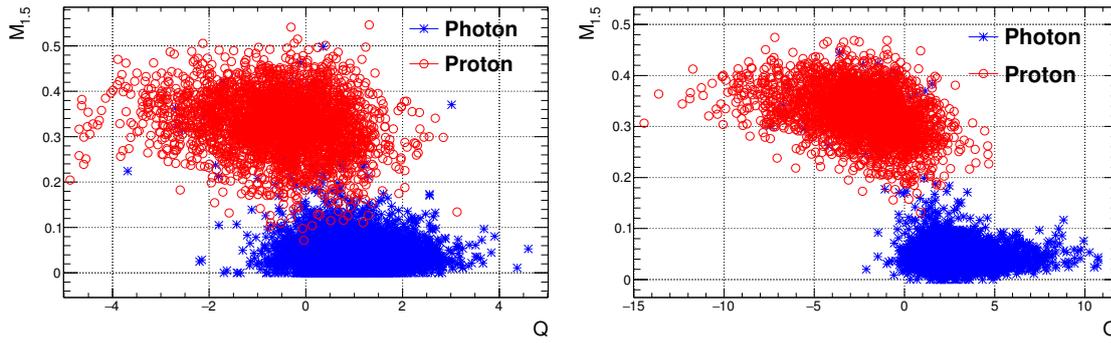
**Figure 6.12:** The mean values of  $Q$  in terms of the reconstructed energy for photon events and fixed MC zenith angles, represented by different markers.

background rejection with respect to the  $M_{1.5}$  criterium alone. In other words, the observable  $Q$  adds stability and robustness to the observable  $M_{1.5}$ . These features are better described in the framework of the multivariate analysis, as explained in the next section.

## 6.2 Linear discriminant analysis

As it is already clear from the previous discussions, all the photon search studies have one element in common: the search for tiny signals in huge data sets. Therefore, the efforts are put in the design of observables, just as  $M_b$  and  $Q$ , that may effectively isolate signal events while rejecting as many background events as possible.

The discrimination between signal and background events with only one observable may be carried out with a tailor-made procedure as explained in Chapter 5. Even in that simple case, the two meta-parameters, namely the primary energy and zenith angle, were considered in the procedure in different ways: the primary energy was given a realistic spectrum-like distribution and the zenith angle was discretized. When adding a second observable, this kind of tailor-made



**Figure 6.13:**  $M_{1.5}$  vs  $Q$  for photon (blue stars) and proton events (red circles) for a fixed MC zenith angle of  $\theta_{\text{MC}} = 30^\circ$ . The events have a reconstructed energy between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{16.9}$  eV (left) or between  $E_{\text{rec}} = 10^{16.9}$  eV and  $10^{17.4}$  eV (right).

procedure would be extremely cumbersome and prone to biases.

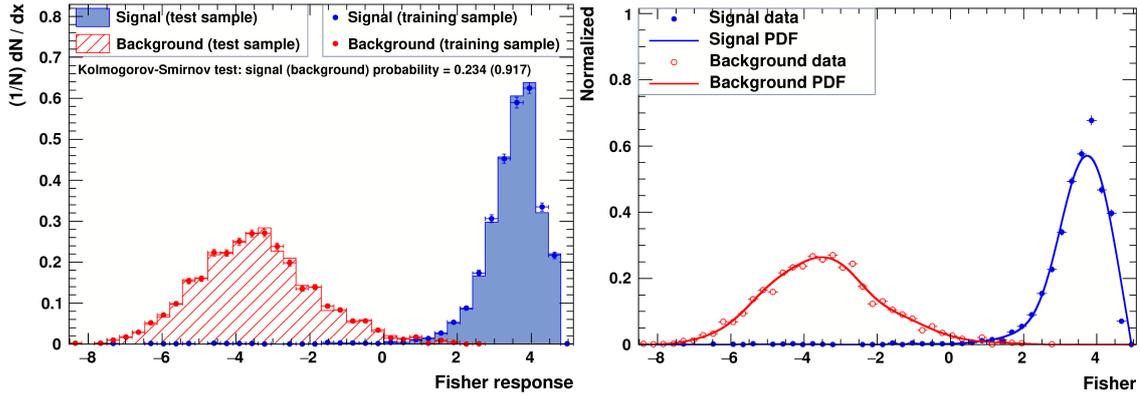
In the photon search analyses, the problem is that of two *classes* of events (background and signal), in which the maximum of the available information needs to be extracted from each event in order to *classify* it as an element of either class. The misclassification chance needs to be minimal as to avoid, as much as possible, background tagged as signal events. This statistical problem is tackled by the *multivariate classification* methods [253]. Based on machine learning techniques, they have become a fundamental ingredient to most analyses, specially in the field of High-Energy Physics.

There is a wide variety of multivariate analysis (MVA) methods that are based in linear or non-linear combination of the input observables. In all cases, the output from these methods is a single classifier parameter. While the non-linear methods are the most powerful, the events lose their identity in the process (i.e. the output events are entangled combinations of the input events). In the case of linear methods, the output events are linear combinations (with known coefficients) of the input events, giving the possibility to interpret the results in terms of the known input observables.

Within the linear MVA, the *Fisher discriminant analysis* is one of the most popular methods [254]. Its base concept is quite intuitive: given an event  $i$ , in which  $\{x^k\}$  observables are measured, the “Fisher coefficients”  $\{F^k\}$  that define the output parameter  $F_i$  as in Eq. 6.7 are obtained. The Fisher coefficients are calculated as to maximize the distance between these distributions while minimizing the variance within each of them. Therefore, maximizing the separation between the two classes by employing all the available information. The output of the method is composed by two distributions on the parameter  $F$ , corresponding to signal and background.

$$F_i = F^0 + \sum_{k=1}^m F^k \cdot x_i^k \quad (6.7)$$

If only two observables are used in the analysis (i.e. if  $m = 2$ ),  $\{F^k\}$  can be obtained analytically. At higher dimensions (i.e. a larger number of observables or classes), the problem should be numerically solved. There is a broad variety of software to perform MVA. The TMVA toolkit [255] is the one chosen for this thesis, since it hosts a large variety of multivariate classification algorithms and is integrated into the ROOT analysis framework, which is also the basis of the Offline development.



**Figure 6.14:** (Left) The distributions of the Fisher response ( $F_i$  in Eq. 6.7) for the training and testing sub-sets of events. (Right) The spline of elemental functions that approximates the probability density function for background and signal events. In both panels, red (blue) lines and markers represent the background (signal) events.

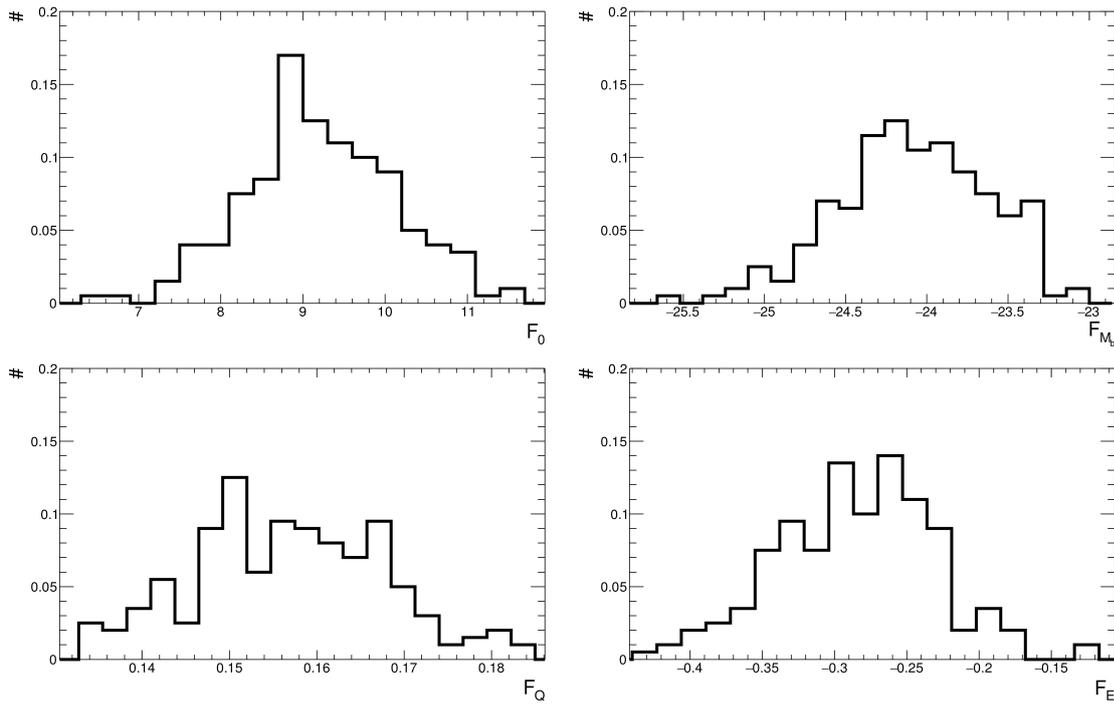
### 6.2.1 Performance of the Fisher discriminant

The photon search presented in this thesis uses two discrimination observables,  $M_b$  and  $Q$ . It is customary to include the meta-parameters in the Fisher analysis as well, although they do not add separation power. Applying the same criterium as before, the Fisher coefficients are first calculated for a set of signal and background events following a spectrum and with a fixed MC zenith angle.

The background and signal events surviving the selection and quality cuts are re-sampled by applying the bootstrap method and the spectral shuffling in the same way as described in Sec. 5.5. Contrarily to that case, the re-sampled events are accumulated until, at least,  $10^5$  background and signal events are obtained. 200 repetitions of the dual procedure are normally taken to accomplish this condition. Both sets of  $10^5$  events are then used as the input ingredients for the Fisher analysis.

A single MVA run is splitted in two parts: the training and the testing phase. During the training, a fraction of the input events is selected randomly in order to find the coefficients  $\{F^k\}$ . In this thesis, this fraction is chosen as 10% (i.e. the training is performed with  $10^4$  background and signal events) to avoid memory consumption problems. It is worth noting that exactly the same number of events for background and signal is used for the training phase. In the testing phase, the rest of the data is projected into the Fisher space (i.e. the  $x_i^k$  are transformed into  $F_i$ ). As an example, some of the outputs from a single run in TMVA are displayed in Fig. 6.14. In the left panel, the distributions of  $F$  are shown for signal and background events, where it can be seen that the distributions for both phases do not have significant differences. In the right panel, the underlying probability density function is estimated by the toolkit as a spline of polynomial and exponential functions. While the background events are fairly well described, the signal events distribution is too peaked for this simple approach. In any case, it can be noted that the Fisher distributions have a striking resemblance to the  $M_b$  ones: a wide background and a sharp signal distributions.

Independent TMVA runs do not necessarily produce the same output, since in each run the training is performed with randomly selected events. Therefore, 200 runs are done in order to find the average transformation defined by  $\{F^k\}$ . As an example, the distributions for each coefficients are displayed in Fig. 6.15 for the case of  $\theta_{MC} = 30^\circ$ . It is interesting to note that the coefficient associated with the observable with the most separation power,  $M_b$ , has the smallest relative



**Figure 6.15:** The distributions of the four Fisher coefficients as obtained in 200 runs of the Fisher discriminant analysis for events with energies between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV following a spectrum  $E^{-(3.06 \pm 0.13)}$ . The MC zenith angle is fixed at  $\theta_{\text{MC}} = 30^\circ$ .

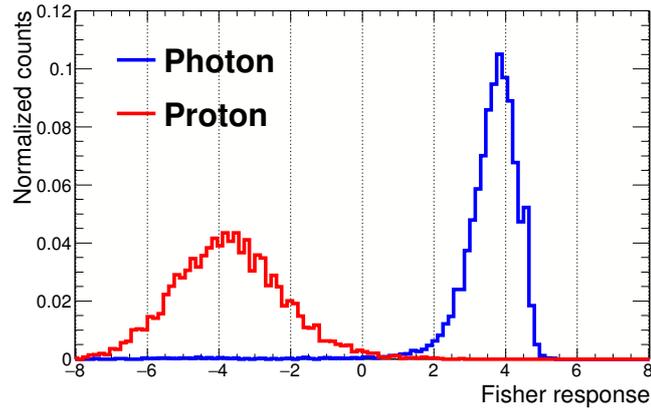
fluctuations, with a mean value of  $-24.1$  and a relative dispersion of  $\sim 1.9\%$ . The coefficient related to the observable  $Q$  has a mean value of  $0.157$  with a relative dispersion of  $\sim 7\%$ . The other two coefficients, which actually do not add any separation, have larger relative dispersions.

The original sample of background and signal events with  $\theta_{\text{MC}} = 30^\circ$  after the quality cuts, projected into the Fisher space, are shown in Fig. 6.16. Note that all events are included in Fig. 6.16, i.e. without applying the spectral correction. The outstanding separation is clear. The photon events that present a hadronic shower development are visible at negative values of the Fisher response. On the other hand, there is one clear proton outlier at  $F \sim 2.8$ , as displayed in Fig. 6.17. This outlier corresponds to an event with  $M_{1.5} = 0.07$  and  $Q = -0.04$  with  $E_{\text{rec}} = 10^{16.6}$  eV. The depth of the shower maximum is  $X_{\text{max}} = 712 \text{ g cm}^{-2}$  which is compatible with the expectation for a photon primary (see Fig. 5.30). Moreover, the measured muon density at the MD station, located at  $r = 117 \text{ m}$  from the shower axis, is  $0.125 \text{ m}^{-2}$  which is more than seven times smaller than the average of  $0.926 \text{ m}^{-2}$  at this distance, energy and zenith.

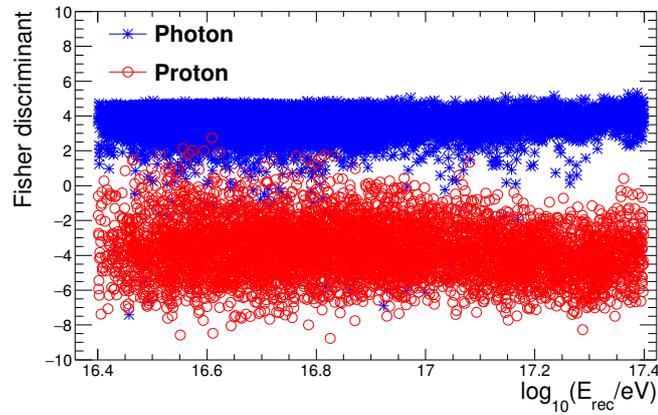
From the evolution of the Fisher response with the reconstructed energy, it can be deduced that an energy-dependent cut can be applied. Indeed, the separation gets larger for higher energies. However, the simplest approach of a constant cut for all energies is considered for this thesis, while an exhaustive examination of more elaborated algorithms are left for future studies.

## 6.2.2 Background rejection and signal efficiency

The Fisher transformation is tested by a second spectral shuffle of the original events (i.e., the events showed in both Fig. 6.16 and 6.17). The background rejection in terms of the signal



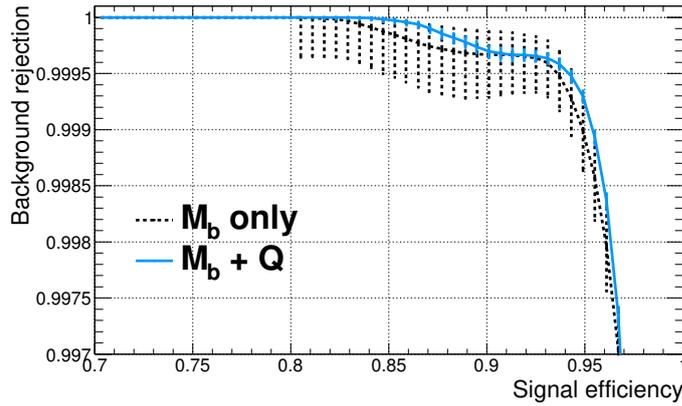
**Figure 6.16:** The Fisher response of the set of events passing the quality cuts with energies between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV and  $\theta_{\text{MC}} = 30^\circ$ , without any spectral re-sampling.



**Figure 6.17:** The Fisher response in terms of the reconstructed energy for the events passing the quality cuts, without any spectral re-sampling. The zenith angle is fixed at  $\theta_{\text{MC}} = 30^\circ$ .

efficiency may then be addressed employing the same procedure as described in Sec. 5.5.1, with the only difference that here the sweep is done in values of  $F_{\text{cut}}$ . The results are shown in Fig. 6.18 at a fixed zenith angle of  $\theta_{\text{MC}} = 30^\circ$ , compared to the same curve obtained by the discrimination method using only  $M_b$ . Note that in both scenarios the same events are used and the same number of spectral shuffling realizations are done in each case. Also, note that the events used for the  $M_b$ -only scenario are reconstructed with the fixed hadronic LDF slope, which explains the small differences in the curve compared to the one presented in Fig. 5.34.

The addition of the observable  $Q$ , through the Fisher machinery, leads to a small increase in the signal efficiency at a fixed background rejection. Particularly, the signal efficiency at the benchmark is 95.4% compared to 94.8% with the case of using only  $M_b$ . Nevertheless, the biggest improvement, as foreseen at the beginning of this chapter, is on the stability of the discrimination procedure. Indeed, the fluctuations of the curve are notoriously smaller than in the case of  $M_b$ -only. Although the combination of two independent observables, moreover, coming from different detectors, leads to a more robust discrimination method, it must be noted that the background outliers with electromagnetic-dominated shower development would still be present in the  $M_b + Q$  scenario. This denotes the physical origin of these outliers and excludes the hypothesis of outliers being created by artifacts in the presented method. Nevertheless, as deduced from Fig. 5.34, the



**Figure 6.18:** The background rejection in terms of the signal efficiency for events with energies between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV after applying the spectral shuffling. The blue (dotted black) line represents the results with  $M_b+Q$  in a Fisher analysis (only  $M_b$ ) is used. The same base events are used in both scenarios. The zenith angle is fixed at  $\theta_{\text{MC}} = 30^\circ$ .

increase of the information availability leads to a decrease of the impact of these outliers. Hence, the background rejection drops from 1 to 0.9996 at a higher signal efficiency than when using only the observable  $M_b$ .

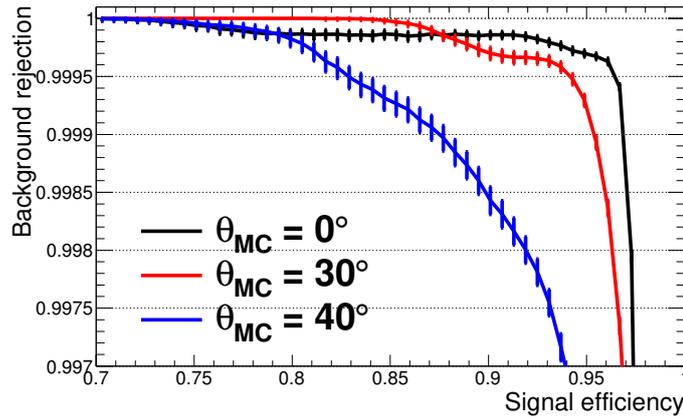
Lastly, the background rejection in terms of the signal efficiency for three fixed MC zenith angles are shown in Fig. 6.19 after performing  $2 \times 10^4$  spectral shuffling realizations. The benchmark rejection is reached at a 96.9%, 95.4% and 87.5% signal efficiency for  $\theta_{\text{MC}} = 0^\circ$ ,  $\theta_{\text{MC}} = 30^\circ$  and  $\theta_{\text{MC}} = 40^\circ$ , respectively. The corresponding values of  $F_{\text{cut}}$  are 1.5, 2.1 and 1.9.

On the other hand, the signal efficiencies at the benchmark rejection when using only  $M_b$  are 96.8%, 94.8% and 87.0%. The addition of  $Q$  mainly impacts non-vertical events, meaning that in the vertical case, the observable  $M_b$  may be already enough to reject most of the background while keeping the available (i.e. not muon-rich) photons. In average, an improvement of  $\sim 0.5\%$  signal efficiency is obtained in the  $M_b + Q$  scenario. Also, as seen in Fig. 6.18, the combined procedure leads to a decrease in the fluctuations in all cases. It is worth noting that the fluctuations are mostly of artificial origin since they reflect the limited available statistics, both in terms of the number of unique events and the number of spectrum shuffling realizations.

### 6.3 Parameterization of the Fisher coefficients

The three sets of background and signal events employed in the last section have a fixed MC zenith angle. For each of them, the background rejection was analyzed with dedicated Fisher coefficients, i.e. calculated on-the-run specially for each set of events. In this fashion, the separation power is addressed despite of the angular dependence. However, in a more realistic scenario, the events present a  $\sim \sin\theta \cos\theta$  zenith distribution. Therefore, a parametrization of the Fisher coefficients in terms of the zenith angle is mandatory.

The Fisher coefficients for each fixed MC zenith angle averaged over 200 MVA realizations are displayed in Tab. 6.2, where standard deviations are shown as uncertainties. The relative dispersion is  $< 10\%$  in almost all cases.



**Figure 6.19:** The background rejection in terms of the signal efficiency for events with energies between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV after applying the spectral shuffling. The colors represent different MC zenith angles as quoted.

$\theta_{\text{MC}}(^{\circ})$	$F^0$	$F^{M_b}$	$F^Q$	$F^E$
0	$(8.97 \pm 0.96)$	$(-27.7 \pm 0.6)$	$(0.235 \pm 0.010)$	$(-0.209 \pm 0.057)$
30	$(9.63 \pm 0.84)$	$(-24.3 \pm 0.4)$	$(0.159 \pm 0.010)$	$(-0.308 \pm 0.049)$
40	$(11.5 \pm 0.70)$	$(-18.6 \pm 0.3)$	$(0.110 \pm 0.010)$	$(-0.494 \pm 0.042)$

**Table 6.2:** The Fisher coefficients averaged over 200 MVA runs for background and signal events between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV after applying the spectral shuffling and corresponding to three fixed MC zenith angles. The presented uncertainties are the standard deviations of each parameter.

As shown in Tab. 6.2,  $F^0$  and  $F^{M_b}$  increase with the zenith angle, while  $F^Q$  and  $F^E$  show a decreasing trend. The quoted MC zenith angles are actually equidistant in  $\sec\theta$  and thus a linear dependence can be imposed as a first approximation:

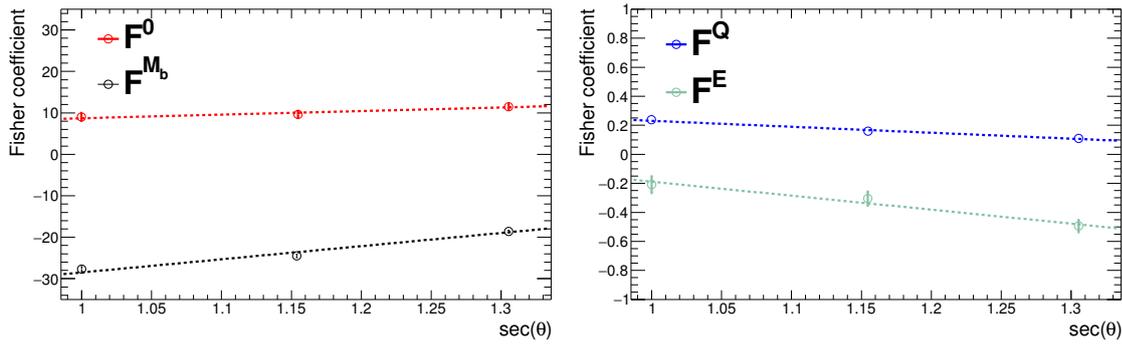
$$F^k(\theta) = a^k + b^k \times \sec\theta \quad (6.8)$$

where the index  $k$  represents the observables and the offset. The coefficients calculated by series of MVA runs at fixed zenith angles seem compatible with this simple linear dependence. Moreover, no drastic change in any of the coefficients is observed in the zenith interval spanned by these fixed MC zenith angle libraries, which actually compose more than 80% of the field of view in  $\theta < 45^{\circ}$ . The linear fits are shown in Fig. 6.20 and the fitted parameters,  $a^k$  and  $b^k$ , are displayed in Tab. 6.3. Since for  $F^0$  the uncertainties are similar to the rate of change in terms of the  $\sec\theta$ , the mean value between the three estimations is taken.

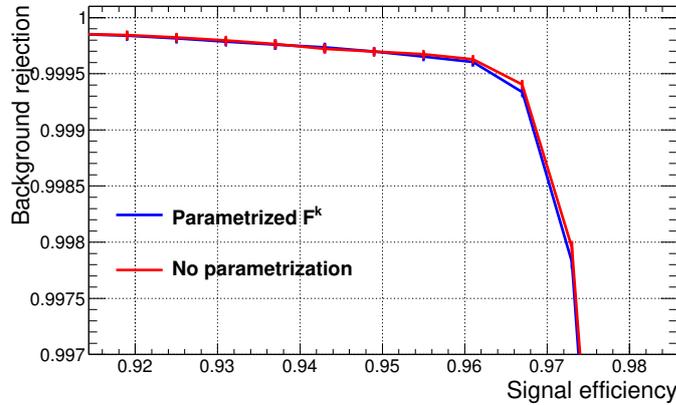
The background rejection when calculating the Fisher coefficients in a dedicated way (i.e. performing the 200 MVA realizations during runtime) can be compared to the case in which the

	$F^0$	$F^{M_b}$	$F^Q$	$F^E$
$a$	$(9.72 \pm 0.04)$	$(-59.3 \pm 0.5)$	$(0.630 \pm 0.005)$	$(0.832 \pm 0.022)$
$b$	—	$(30.7 \pm 0.5)$	$(-0.400 \pm 0.005)$	$(-1.01 \pm 0.02)$

**Table 6.3:** The fit parameters of the linear dependence between the Fisher coefficients and  $\sec\theta$  calculated for events between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV after applying the spectral correction.



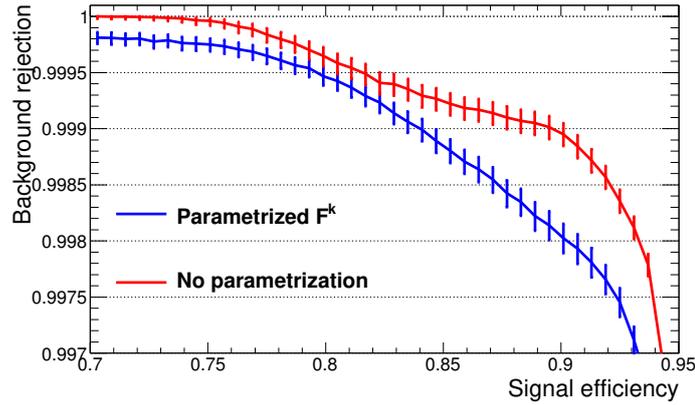
**Figure 6.20:** The Fisher coefficients in terms of the MC zenith angle with linear fits superimposed. The coefficients have been grouped in pairs for better visualization.



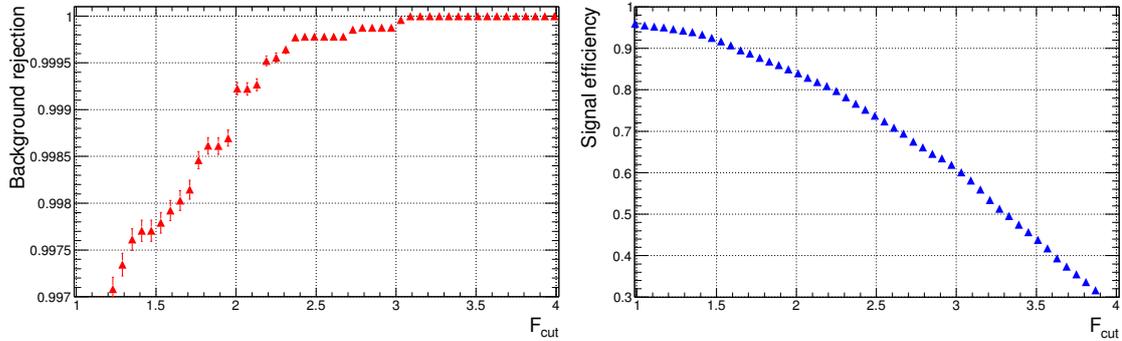
**Figure 6.21:** The background rejection in terms of the signal efficiency for the set of events at  $\theta_{MC} = 0^\circ$ . The red curve represents the case in which the Fisher coefficients are calculated on-the-run (i.e. the values from the first row of Tab. 6.2). The blue curve represents the results obtained by employing the Fisher coefficients as calculated from Eq. 6.8.

Fisher coefficients are parametrized by Eq. 6.8. A consistency check can be seen in Fig. 6.21 for the particular case of  $\theta_{MC} = 0^\circ$ . Since these sets of events are used to design the parametrization, it is not surprising that the discrepancy is less than  $\sim 0.5\%$  in the signal efficiency for a fixed rejection, and in particular, that no change is found at the benchmark.

The parametrization must be tested with an independent set of events with a continuous energy and angular distribution. For this purpose, a set of background and signal events with energies between  $E_{rec} = 10^{16.4}$  eV and  $10^{17.4}$  eV and  $\theta < 45^\circ$ , distributed as  $\sin\theta\cos\theta$ , is introduced into the Fisher analysis. The average Fisher transformation is obtained and, as before, the background rejection in the Fisher space is addressed. It must be noted that, since the zenith angle in this case is variable, it is included as a second meta-parameter in the Fisher machinery. On the other hand, the background rejection can be also studied by using the parametrized transformation obtained before. The comparison between these two scenarios is displayed in Fig. 6.22. Indeed, the parametrization does not lead to a full rejection as in the case of a dedicated Fisher calculation. Moreover, the background rejection drawback of using the parametrized coefficients could be as large as 0.1%. In particular, at the fixed benchmark rejection, the signal efficiency falls from 89% to 84%, meaning that the parametrization is not sensitive enough to the signal at such high background rejection. The cut on the parameter  $F$  that is needed to achieve the benchmark can be extracted from Fig. 6.23, left. Indeed, a value of  $F_{cut} = 2$  is enough to reject, at least, 99.90% of



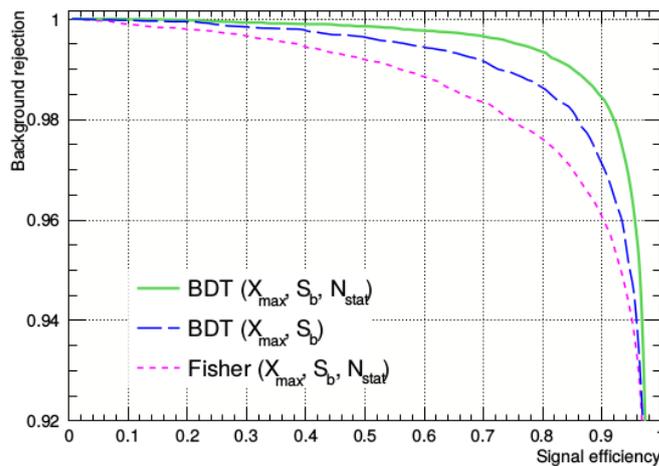
**Figure 6.22:** The background rejection in terms of the signal efficiency for events with a distributed as  $\sin\theta \cos\theta$  for  $\theta < 45^\circ$ . The red curve represents the case in which the Fisher coefficients are calculated on-the-run. The blue curve represents the results obtained by employing the Fisher coefficients as calculated from Eq. 6.8.



**Figure 6.23:** The background rejection (left) and signal efficiency (right) in terms of the cut on the parameter  $F$ , calculated with the parametrization of Eq. 6.8. The selected events have energies between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV after the spectral shuffling. The zenith angles are distributed as  $\sin\theta \cos\theta$  with  $\theta < 45^\circ$ .

the background. For completeness, the signal efficiency in terms of  $F_{\text{cut}}$  is displayed in Fig. 6.23, right.

Even though the model of Eq. 6.3 is rather simple and the fitted parameters are estimated with a poor data set (only three fixed MC zenith angles), the performance of the parametrization is remarkable. Although there is much room for improvement, the current status of the Fisher parametrization would represent the best tool for the photon studies in the Auger Collaboration, both in terms of the signal efficiency and the background rejection. For comparison, the background rejection in terms of the signal efficiency obtained in the last photon search paper of the Auger Collaboration is shown in Fig. 6.24, where the observable  $S_b$  is related to the signals measured by the WCDs (i.e. an observable sensitive to the LDF steepness) and  $N_{\text{stat}}$  refers to the number of triggered WCDs. The events have  $E_{\text{rec}} > 10^{18}$  eV and  $\theta < 65^\circ$  with a mixed-composition background of 50% proton and 50% iron primaries. The background rejection obtained at a 50% of signal efficiency, using the Fisher algorithm, is  $\sim 99\%$ , while the benchmark rejection of 99.90% is nearly out of reach. In the case of the photon search procedure developed in this thesis, the background rejection at 50% signal efficiency would be easily 99.99% only limited by the lack of statistics, leading to an improvement of two orders of magnitude in terms of the background



**Figure 6.24:** The background rejection efficiency against the signal efficiency for different MVA algorithms and observables from [114]. The events have  $E_{\text{rec}} > 10^{18}$  eV and  $\theta < 65^\circ$  and the curves are obtained with a background composition of 50% proton and 50% iron.

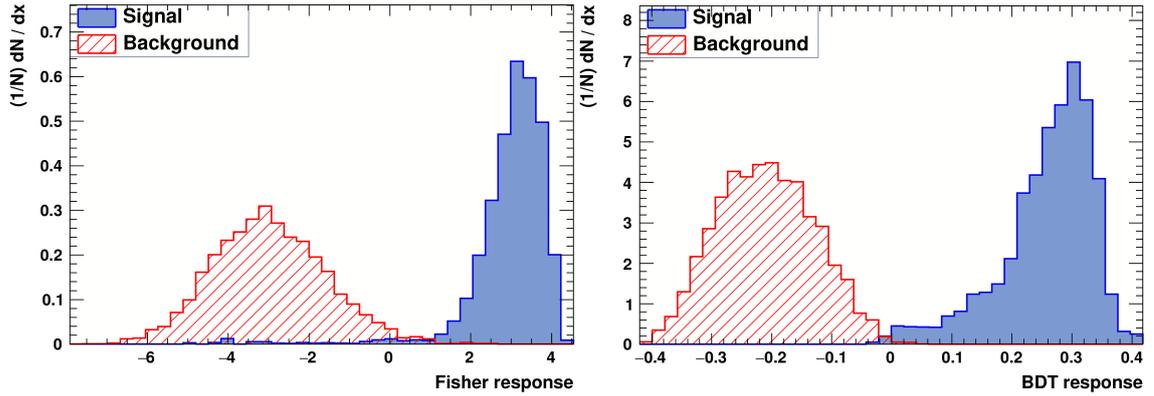
contamination. Moreover, the  $M_b + Q$  procedure is developed with a pure proton background, which constitutes a more disfavoured scenario in terms of the muon content.

## 6.4 Non-linear multivariate analysis

The Fisher analysis is chosen in this thesis as the MVA algorithm due to its simplicity and its intuitive approach. In the majority of studies, the non-linear algorithms easily outperform the linear Fisher discriminant. This is clearly seen in Fig. 6.24, where BDT stands for *Boosted Decision Trees*, a popular non-linear algorithm. In the particular case of this photon search, the BDT gives a better separation, as seen in Fig. 6.25 for a single MVA run with events between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{17.4}$  eV and  $\theta < 45^\circ$ , distributed as  $\sin\theta \cos\theta$ . In both cases, the zenith angle is included as a fourth parameter in the MVA and the events follow a spectral distribution as usual. The usage of a non-linear method such as BDT may lead to a great enhancement of this analysis. However, several aspects must be thoroughly studied. Indeed, the BDT algorithm is based upon a careful selection of internal parameters, e.g. the number of decision trees in the forest, the depth of each of them, etc. In the example showed in Fig. 6.25, the default parameters provided by TMVA were used, only for the purpose of a qualitative comparison to the Fisher analysis.

The better performance leads directly to a higher signal efficiency for a fixed background rejection, as displayed in Fig. 6.26. For this particular MVA run, the benchmark rejection is reached with 87% signal efficiency when using Fisher, compared to  $> 98\%$  when using BDT, implying that in the latter case it may be possible to attain an almost complete separation of signal events from background. Lastly, an algorithm based on artificial neural networks is also shown in Fig. 6.26. Although it is a complex non-linear algorithm, its performance is not better than the simple Fisher analysis.

In the examples showed before, four observables were included in the TMVA package. From the discussion of Sec. 5.6, the depth of the shower maximum may also add composition-sensitive information to the  $M_b + Q$  method. Although it is difficult to measure the longitudinal profile of the showers at the energies of interest, even with the HEAT telescopes, a thought experiment in



**Figure 6.25:** The MVA responses to background (red) and signal events (blue) using a linear Fisher analysis (left), as employed in this thesis, and using a Boosted Decision Trees algorithm (right). In both cases, the same events are used for training and testing. The events are selected according to their reconstructed energy between  $10^{16.4}$  eV and  $10^{17.4}$  eV as  $E_{rec}^{-\gamma}$  and with zenith angles distributed as  $\sin\theta\cos\theta$  for  $\theta < 45^\circ$ .

which  $X_{\max}$  is also included in the MVA is discussed in Appendix C.

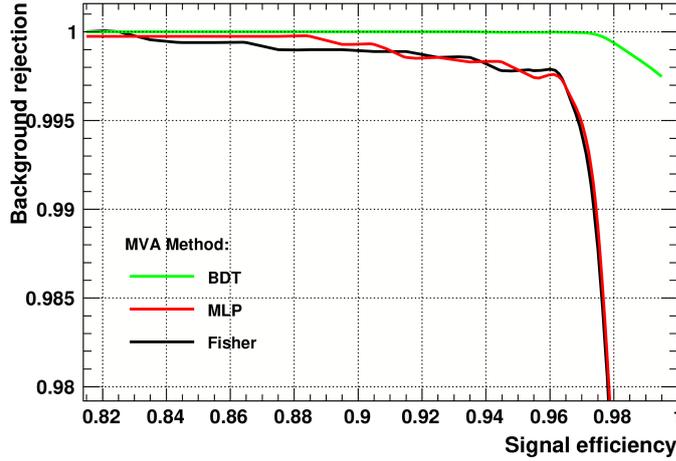
## 6.5 Discussion on possible systematic effects

Although the discrimination procedure developed in this thesis is sustained on robust foundations, there are systematic effects that inevitably affect the observables in an event-by-event basis. While an extensive study is needed to properly assess the impact of each effect, a summary of the possible sources of uncertainty is presented in this section, as a prelude for future studies on this topic.

As mentioned a few times before, the mass composition of the primary particles (the background) has been considered as pure-proton to develop the procedure. While this is not incompatible with the current measurements that indicate a predominantly light composition at the energies of interest for this thesis [128], the cosmic-ray flux may well have a non-negligible presence of nuclei with  $A > 1$ . Therefore, a larger muon content may be present in data. Mustering the Heitler-Matthews model (recall Sec. 2.3.1), for example  $\sim 25\%$  more muons are expected for an EAS generated by a carbon primary ( $A = 12$ ).

The EAS simulations performed for this thesis were produced with QGSJetII-04 as the hadronic model for high-energy interactions. Nowadays, there is not a consensus about which hadronic model describes data more trustworthily. Thus, MC studies need to assess this intrinsic uncertainty and its impact on the corresponding observables. The three most popular models, being Epos-LHC, SYBILL 2.4c and the chosen one, predict a different number of muons. The differences between them are less than  $\sim \pm 15\%$  [256].

Apart from the choice of the hadronic model, there is a general agreement that all the hadronic models underestimate the muon content of the EAS. This issue is usually addressed as the “muon deficit problem”. The muon deficit has been investigated in Auger for energies above  $6 \times 10^{18}$  eV [243, 232]. In that energy domain, it has been estimated that the simulations may have a deficit of at least  $\sim 30\%$  in the muon content, in comparison to data. Unfortunately, no estimations have been reported for the energies of interest for this photon search.



**Figure 6.26:** The background rejection against the signal efficiency using Fisher (black) or a BDT analysis (green). The red curve corresponds to a class of artificial neural networks called Multilayer Perceptron (MLP). In the three cases, the same events are used for training and testing. The events are selected according to their reconstructed energy between  $10^{16.4}$  eV and  $10^{17.4}$  eV as  $E_{\text{rec}}^{-\gamma}$  and with zenith angles distributed as  $\sin\theta \cos\theta$  for  $\theta < 45^\circ$ .

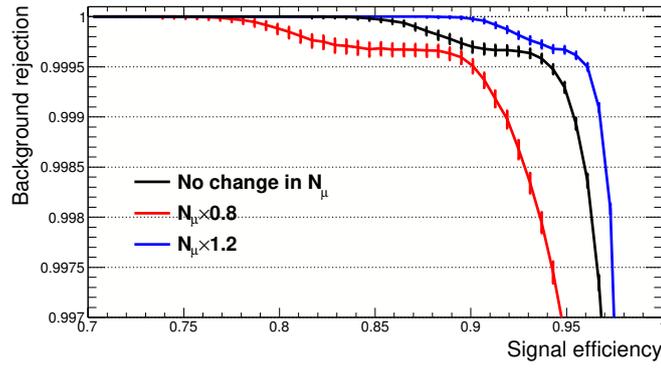
From the point of view of the detector itself, the choice of an inhibition window of 25 ns in the muon counting of the MD stations (recall Sec. 3.4.5) leads to an undercounting in the number of muons, which has been estimated to range between 15% and 25% for the distances of interest (see Fig. 3.27, right).

All these systematic effects would lead to an overall increase of the measured muon density in the MD stations in data with respect to the MC study presented so far. Thus, larger values of  $M_b$  would be expected in data. Indirectly, the enhancement of the muon component in the EAS would also affect the slope of the LDF, making them even less similar to the expectation from electromagnetic-dominated EAS. In turn, this would lead to an increase in the absolute value for  $Q$ . Therefore, a larger Fisher absolute value would be expected, and thus, a larger background rejection, compared to the MC studies of this thesis, could be attained. By rescaling the measured number of muons (hence, redefining the values for  $M_b$ ) a first estimation of the impact of the muon under- or overestimation can be addressed. This is displayed in Fig. 6.27 for a fixed zenith angle of  $\theta_{\text{MC}} = 30^\circ$ , where the number of muons are changed by  $\pm 20\%$ . Indeed, if muons are underestimated in simulations, the separation in data would be greatly improved, with a signal efficiency changing from 95.4% to 96.8%. It must be noted that the values of  $Q$ , under a rescaling of the muon content, is also perturbed, leading to an even greater signal efficiency. On the other hand, an overestimation of the muon number in simulations would lead to a signal efficiency of 91.9%.

Different systematic effects may also emerge from the procedure itself. In the calculation of  $M_b$ , a change in the reference muon density  $\rho_\mu^{\text{ref}}$  may also lead to the detection of false positives. The variation of the observable  $M_b$  due to variations of the reference muon density can be calculated as:

$$\Delta M_b = (1 - 10^{-M_b}) \times \frac{\Delta \rho_\mu^{\text{ref}}}{\rho_\mu^{\text{ref}}} \quad (6.9)$$

Since  $M_b \sim 0.1..0.6$ , a change of  $\pm 10\%$  in  $\rho_\mu^{\text{ref}}$  would lead to a variation of 0.02..0.07 in the expected values of  $M_b$  for both background and signal events. In practice, a decrease of  $\pm 0.03\%$



**Figure 6.27:** The background rejection against the signal efficiency in the case of a rescaling of the measured number of muons by  $-20\%$  (red),  $+20\%$  (blue) and without change (black). The events are selected according to their reconstructed energy between  $10^{16.4}$  eV and  $10^{17.4}$  eV as  $E_{\text{rec}}^{-\gamma}$  and with a fixed MC zenith angle of  $\theta_{\text{MC}} = 30^\circ$ .

of the signal efficiency at the benchmark rejection is seen when  $\rho_\mu^{\text{ref}}$  is fluctuated  $\pm 10\%$ .

The spectrum shuffling by itself is based upon the hypothesis that the photon component of the cosmic-ray flux follows the same spectral shape as the observed flux. This could not be the case if a photon flux is produced by a transient source or exotic phenomena beyond the Standard Model. Increasing the spectral index for the signal to  $\gamma = -2$ , while keeping the measured index for the background, an increased in the signal efficiency from  $95.4\%$  to  $95.6\%$  at the benchmark is found for  $\theta_{\text{MC}} = 30^\circ$ , which is negligible compared to the impact of a rescaling in the number of muons.

There are additional effects known to appear in data that are not entirely simulated. One of them is the “direct light effect” that arises when particles enter the WCD with a direction where the beam of Cherenkov photons points directly onto the PMTs. These particles produce larger signals than particles which deposit the same energy in the detector but produce a beam of photons which are scattered isotropically around the WCD before reaching the PMTs. Therefore, it is expected that “fake muon signals” may appear in data. This effect is partially corrected for with a dedicated Offline module (see Appendix A). Additionally, atmospheric muons, that are not part of the EAS, may leave a detectable trace in the WCD, even when the WCD stations that are triggered exclusively due to these accidental muons are removed from the event reconstruction. In terms of the discrimination problem, this case imposes a conservative scenario, since a signal event may appear as a hadronic event (through the SD-related observable  $Q$ ), while a background event would be unchanged. However, by employing two complementary observables estimated with different detectors, the impact of these effects is expected to be negligible.

In summary, the conditions in which the discrimination procedure of this thesis has been designed comprise nearly the worst case scenario in terms of systematic effects, reinforcing the robustness of this work.



## Chapter 7

# Towards the search for ultra-high energy photons in data

The scientific man does not aim at an immediate result. He does not expect that his advanced ideas will be readily taken up. His work is like that of the planter - for the future. His duty is to lay the foundation for those who are to come, and point the way.

---

Nikola Tesla

Through the chapters of this thesis, the problem of the discrimination between photon- and hadron-initiated events with energies  $E_{\text{rec}} > 10^{16.4}$  eV and  $\theta < 45^\circ$  has been extensively investigated. The combination of the MD information, carried by the observable  $M_b$ , and the LDF information, through the observable  $Q$ , led to a multiparametric discrimination method whose outstanding performance is unmatched by any other procedure in the Auger Collaboration. All these efforts were conducted while keeping in mind the idea of studying the data reconstructed with the SD-433, with which an accurate event reconstruction can be attained at the energies of interest, and the 60 m<sup>2</sup> MD station data at the center of it.

As commonly encountered in any field, data acquired by the detectors may show features that are not present in the MC studies done beforehand. In the first part of this chapter, the available data of the SD-433 suitable for this photon search is scrutinized and the characteristics of the reconstructed events are discussed. In the second part, the exposure of the SD-433, given the necessary selection and quality cuts for the discrimination procedure, is calculated. This is the key element to obtain the two most important results of this thesis: the upper limits to the photon flux that would be imposed if no events compatible with a photon origin are found in data and the sensitivity to a photon signal, if present, in the cosmic-ray flux. To conclude the chapter, the future expectations of this work are described.

## 7.1 Data selection

The SD-433 array has been operational since January 2013 and the engineering array of the MD, particularly the Id 1764 MD station at the center of the SD-433, was finished at the end of 2014. Since then, several tests related to the overall performance of the MD stations have been conducted, particularly aimed at evaluating minor changes in the front-end electronics design and setting the most efficient calibration parameters such as the discrimination threshold for the muon counting and the gain of the PMTs. On October 19, 2015, the calibration phase of the MD ended. Therefore, data acquired after this day may be considered to perform physics analysis.

As explained in Sec. 3.4.3, the optic system of the design of the MD engineering array is based on a PMT. In the framework of the AugerPrime upgrade, the MD design was upgraded, replacing the PMTs for SiPMs for the forthcoming UMD. After several successful laboratory tests, the new electronic and optic system was installed in the Id 1764 MD station for testing purposes on October 17, 2016.

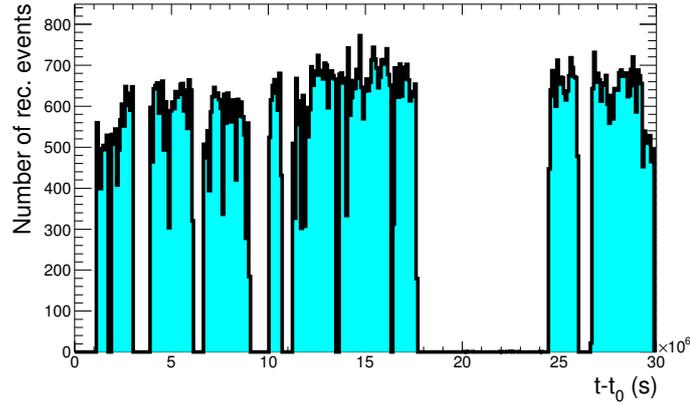
The MD simulation in the Offline framework represents, as the time of this thesis, the base MD design. Thus, the simulated electronic response is based on the PMT characteristics, such as the cross-talk probabilities, the quantum efficiency, the shape of the SPE pulses, etc. While the SiPM adaptation is under development, all the MC studies presented in this thesis were done using the former stable version. Therefore, the two mentioned dates determine the feasible data period to perform this photon search, while its extension including the SiPM-based design is left for future studies.

The data is reconstructed with the Offline framework using a module sequence as detailed in Appendix A. It has some minor changes from the module chain used for the reconstruction of simulated events. A few modules responsible for small corrections from an ideal detector are added, such as changes in the GPS position of the WCDs, known time delays in the trigger communications in certain WCDs, etc. On the other hand, some modules related to the MC studies, such as the *PositionRejector*, are not needed anymore. From the point of view of the LDF, the reconstruction is performed with the fixed (hadronic) LDF slope model, which is the official choice in the case of data reconstruction.

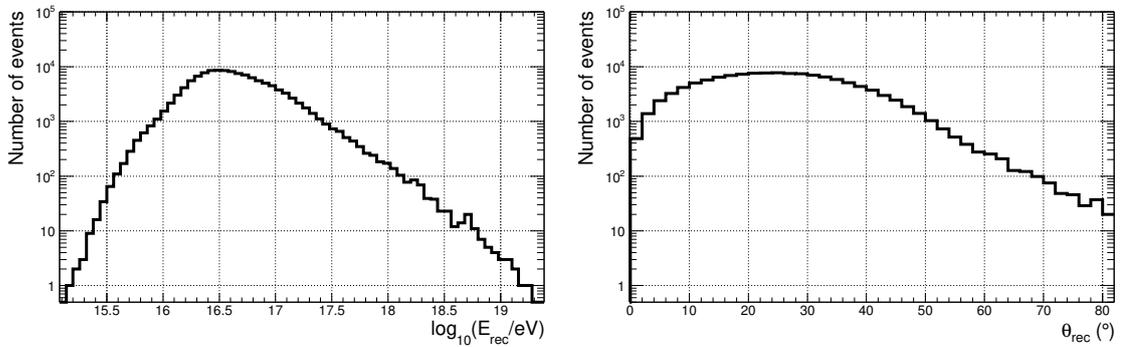
### 7.1.1 General features of the data

The features of the reconstructed events, such as the energy and the geometry, are addressed in this section. While the reconstruction procedure itself is well-understood and developed through the years to be almost free from biases, some detector effects are unavoidable. Each WCDs could be non-functional at any given time due to failures in the telecommunications, the batteries, the electronics, etc. The downtime of the WCDs may cause a notorious loss of statistics. However, this does not affect the discussion of the general features seen in data. The proper assessment of this problem is actually crucial in the calculation of the exposure to be discussed in Sec. 7.2.

The GPS timestamp of the first event of the period,  $t_0 = 1129291205$  (i.e. October 19, 2015, 11:59 UTC), is taken as the time origin for all the studies of this chapter. The selected period ends at  $t_{\text{last}} = 1160740813$  (i.e. October 17, 2016, 11:59 UTC). Unless stated otherwise, time-dependent magnitudes are averaged in a daily basis (i.e. the time bin width represents one day). The number of reconstructed events is displayed in Fig. 7.1, where no selection or quality cuts have been



**Figure 7.1:** The number of reconstructed events by the SD-433 in terms of the elapsed time since the GPS  $t_0 = 1129291205$  on October 19, 2015. The last day of the selected data period is October 17, 2016.

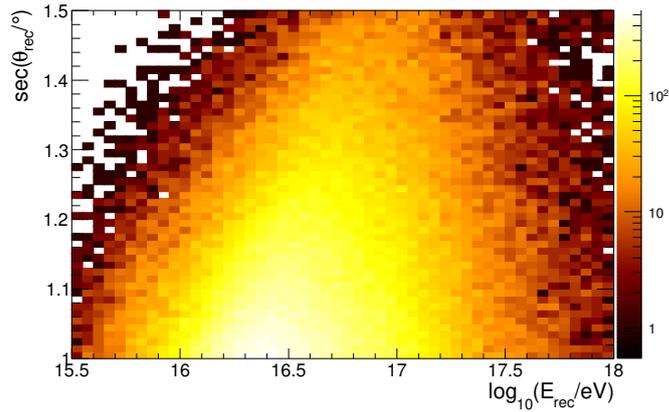


**Figure 7.2:** The distribution of reconstructed energy (left) and zenith angle (right) for events in the selected data period. No selection or quality cuts are applied to the events.

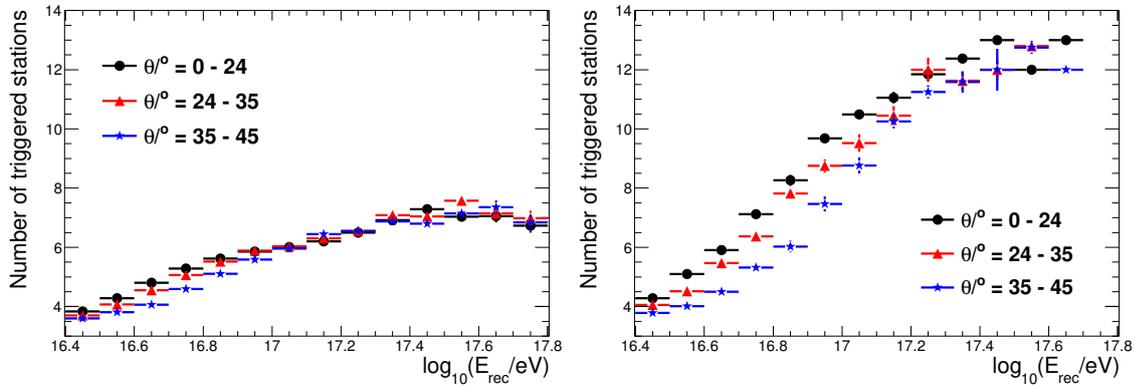
applied. It can be seen that the data acquisition is not stable in time, having some gaps in the first few months. In the most stable periods,  $\sim 560$  events are acquired per day. There is a long downtime from May 11 to July 28, 2016. As it is described in Sec. 7.2.1, the lack of events in this period could have been due to some kind of failure in the communication with all the WCDs of the SD-433 hexagon.

The reconstructed energy distribution of all the events in the selected period is shown in Fig. 7.2, left. Below the maximum of the distribution at  $E_{\text{rec}} = 10^{16.5}$  eV the number of events decreases with the reconstructed energy due to the trigger inefficiency of the array. Above this energy, the SD-433 is fully efficient and the decrease of the events is explained by the steeply falling spectrum of the cosmic rays. It is worth noting that this energy threshold is compatible with the estimation done previously from MC studies (see Sec. 4.3.1). In Fig. 7.2, right, the zenith angle distribution is shown for the same events, where a plateau around  $\theta_{\text{rec}} = 20^\circ$  is seen. Due to the atmospheric attenuation of the particles, the events with smaller reconstructed energy tend to have a more vertical incidence, as deduced from Fig. 7.3. The average zenith angle at  $E_{\text{rec}} = 10^{16.4}$  eV is  $\theta_{\text{rec}} \simeq 25^\circ$  while at  $E_{\text{rec}} = 10^{17.5}$  eV is  $\theta_{\text{rec}} \simeq 41^\circ$ .

As discussed in the previous chapters, the footprint of the EAS on the ground depends on the primary energy and the zenith angle. Fig. 7.4, left, displays the number of triggered WCDs in terms of the reconstructed energy for three zenith bins. The footprint size is indeed larger



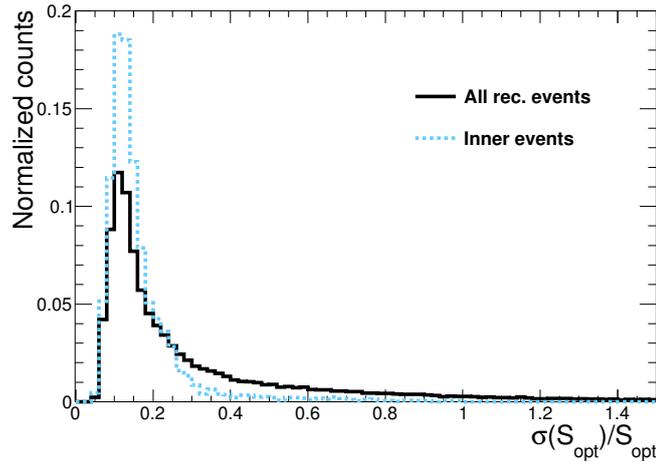
**Figure 7.3:** The two-dimensional distribution of the  $\sec\theta_{\text{rec}}$  and  $E_{\text{rec}}$  of the data events in the selected period. No selection or quality cuts are applied to the events.



**Figure 7.4:** The number of triggered WCDs against the reconstructed energy when no geographic constraint is imposed to the shower cores (left) and when taking only events well-contained within the SD-433 (right). The different markers represent zenith angle bins.

as the primary energy increases. At energies below  $E_{\text{rec}} \simeq 10^{17}$  eV, the atmospheric attenuation dominates over the geometrical shrinking of the array, as described from the MC studies in Sec. 4.3.1. However, above  $E_{\text{rec}} \simeq 10^{17.4}$  eV the number of triggered WCDs starts to retreat. This is actually an artifact from the lack of selection cuts imposed in the events. Indeed, at these energies, the SD-750 is at full efficiency and thus, data starts to be acquired with this array. In this case, the shower cores can be well outside the SD-433 area but, since the EAS contains enough particles, it can be reconstructed with a few peripheral stations from the SD-433 as well. Therefore, the number of triggered WCDs (from the SD-433 alone) begins to decrease as the energy increases. The kink in the profile is thus zenith dependent. The one corresponding to nearly-vertical events has a break at  $E_{\text{rec}} \simeq 10^{17.4}$  eV while for  $\theta_{\text{rec}} > 35^\circ$  (blue stars) the break is present at  $E_{\text{rec}} \simeq 10^{17.6}$  eV, pointing out the trigger efficiency dependence on the zenith angle (in this case, of the SD-750). When asking for events well contained in the SD-433 hexagon as in Fig. 7.4, right, this feature vanishes and the expected dependence is completely recovered. All the 13 WCDs are, on average, triggered for energies above  $E_{\text{rec}} \simeq 10^{17.5}$  eV if  $\theta_{\text{rec}} < 45^\circ$ , while this happens above  $E_{\text{rec}} \simeq 10^{17}$  eV in the case of the first crown of WCDs.

At the same time, the quality of the reconstructed events is affected by the position of the shower core. The shower cores located outside the SD-433 are reconstructed, in average, with a larger uncertainty in  $S_{\text{opt}}$  (hence, with a less reliable reconstructed energy). As an example, the



**Figure 7.5:** The relative uncertainty in  $S_{\text{opt}}$  for reconstructed events by the SD-433 without any geographic constraints (full black line) and for events well-contained in the SD-433 hexagon (dotted blue line). In both cases, the events verify  $E_{\text{rec}} < 10^{17}$  eV and  $\theta_{\text{rec}} < 45^\circ$ .

uncertainty in  $S_{\text{opt}}$  is shown in Fig. 7.5 for the subset of events with  $E_{\text{rec}} < 10^{17}$  eV and  $\theta_{\text{rec}} < 45^\circ$ .

Therefore, the event reconstruction is strongly dependent on the position of the shower core, from the point of view of the available signals to obtain the LDF (i.e. the number of triggered WCDs) but also in relation with the reconstruction procedure itself. Nevertheless, the discrimination procedure reaches its best performance for events whose shower core is located close enough to the central position of the SD-433, for which a proper reconstruction is possible.

### 7.1.2 Impact of the quality cuts

Apart from the requirement that the events themselves must be reconstructed with the SD-433, the  $M_b + Q$  discrimination procedure has been tuned under certain conditions, as explained in the previous chapters. These selection and quality cuts would then affect the energy and zenith distribution of the events accepted for the photon search. In Tab. 7.1, the number of available events in the selected data period, and the action of each quality cut, is displayed.

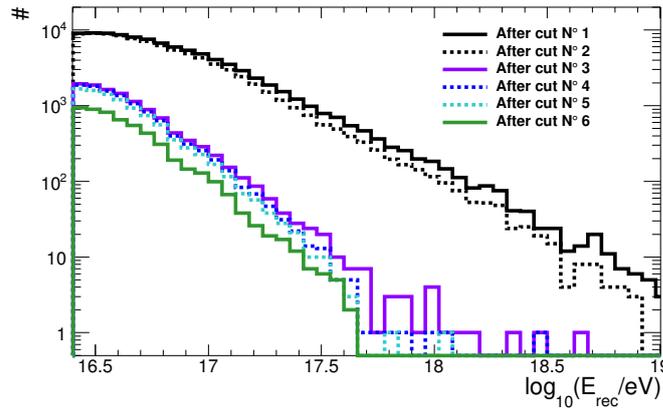
The first three cuts are strictly selection cuts which define energy, zenith and core location of the events to be used for the photon search. The next two are the quality cuts described in Chapter 5. The last cut is exclusively related to the data analysis. While in the MC studies the MD stations are ideal, in reality they are affected by downtimes as in the case of the WCDs. Thus, the condition that the MD station must be active is imposed. This topic is described more extensively in Sec. 7.2.

After the quality cuts, 5359 events are left for the photon search. Each cut alters the distribution of the observed energies and zenith incidence of the events. This is represented in Fig. 7.6, where the most populated distribution (black line) represents the one shown before in Fig. 7.2, left. As already deduced from Tab. 7.1, the requirement about the shower core position being in the elemental hexagon (purple solid line) is the one that depletes most of the event sample. The rest of the cuts can be thought of as rescalers to the distribution after that cut. The non-saturation condition (dark blue dotted line) depletes most of the high energy events, as expected. No events above  $E_{\text{rec}} = 10^{17.7}$  eV are present in the final sample.

The quality cuts do not alter considerably the shape of the zenith angle distribution, as seen

Selection/Quality cut	# of events	Fraction wrt previous cut
Raw Auger events	4349936	-
Reconstructed by SD-433	134683	3%
$\log_{10}(E_{\text{rec}}/\text{eV}) > 16.4$ and $\theta_{\text{rec}} < 45^\circ$	89265	66%
Nearest WCD is Id 1764	11488	12%
Unsaturated event	10729	93%
$r_{\text{MD}} < 200$ m	9484	88%
Active Id 1764 MD station	5359	56%

**Table 7.1:** The selection and quality cuts applied to the Auger data for the  $M_b + Q$  photon search between the GPS  $t_0 = 1129291205$  and  $t_{\text{last}} = 1160740813$ .



**Figure 7.6:** The distribution of the reconstructed energy for the surviving events after each selection/quality cut, according to Tab. 7.1.

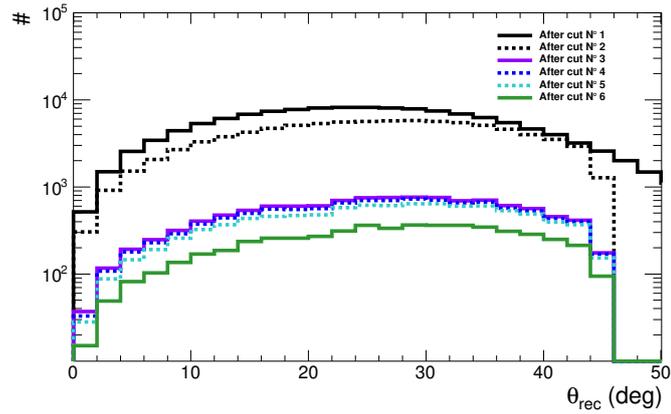
in Fig. 7.7. The median of the final distribution is  $\bar{\theta}_{\text{rec}} = 25^\circ$  and changes less than 10% with the application of each cut. The position of the shower cores of the accepted events are depicted in Fig. 7.8. Several structures are identified as accumulation of shower cores (black dots). These are typically low energy events (i.e.  $E_{\text{rec}} < 10^{16.4}$  eV) that are badly reconstructed due to the scarce number of triggered WCDs. The center of each triangle is overpopulated since in this region the T3 probability is maximal. Indeed, if a low energy shower core impinges the ground near one of the WCD, the other two forming the triangle would mostly not trigger due to the low signal generated by the shower front.

## 7.2 Exposure calculation

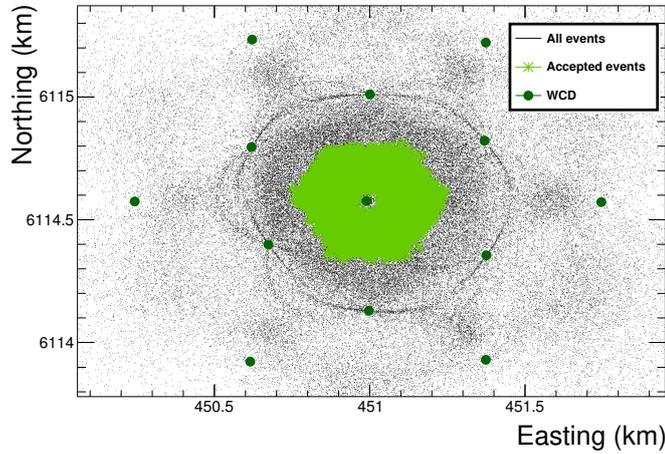
In order to perform any calculation regarding the cosmic-ray flux, it is mandatory to estimate the uptime of the detector and consider all the necessary correction factors. The elemental hexagon requirement for the discrimination procedure greatly simplifies the calculation of the *acceptance*, i.e. the collection area for the detection of the cosmic-ray flux, as depicted in Fig. 7.9. Given that the WCDs are separated by a distance  $d$ , the elemental hexagon has an area  $A_{\text{cell}}$  defined by:

$$A_{\text{cell}} = \frac{\sqrt{3}}{2} \times d^2 \quad (7.1)$$

Thus,  $A_{\text{cell}} \simeq 0.162 \text{ km}^2$  in the case of the SD-433 array. In order to calculate the effective



**Figure 7.7:** The distribution of the reconstructed zenith angle for the surviving events after each selection/quality cut, according to Tab. 7.1. The color coding is the same as in Fig. 7.6.



**Figure 7.8:** The geographic location of the reconstructed core positions of all events (tiny black dots) and those accepted for the photon search (light green stars). The position of the WCDs of the SD-433 are shown with dark green dots. The positions are represented in UTM coordinates.

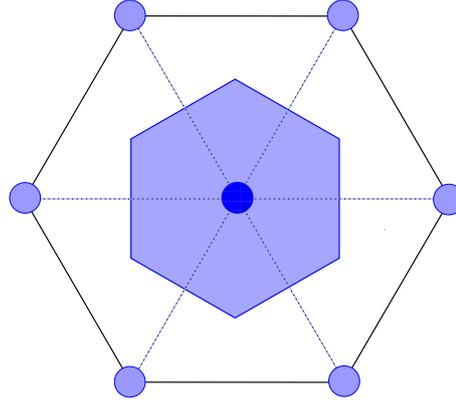
sensitive area  $A_{6T5}$  of a 6T5 cell (see Sec. 3.3.3), the zenith angle has to be taken into account, as the relevant area is obtained by projecting the cell area into the shower plane as in Eq. 7.2.

$$dA_{6T5} = A_{\text{cell}} \times \cos\theta \, d\Omega \quad (7.2)$$

with the solid angle element  $d\Omega = \sin\theta \, d\theta \, d\varphi$ . The acceptance can be integrated over a period of observation to obtain the *exposure* of the detector. In principle, the acceptance defined in Eq. does not contain explicitly any temporal dependence.

As already mentioned, the WCDs could suffer from failures impeding the stable data acquisition. When one of the WCDs is not working, it can be easily seen that the shaded area of Fig. 7.9 is reduced by a factor of  $4/6$ . In this way, the *uptime* of each WCDs affect the exposure. If the chance for a 6T5 and 5T5 configuration in an observation time  $dt$  is defined as  $P_{6T5}$  and  $P_{5T5}$  respectively, the exposure  $\varepsilon$  verifies that:

$$d\varepsilon \propto A_{\text{cell}} \times \cos\theta \times \left( P_{6T5}(t) + \frac{4}{6} \times P_{5T5}(t) \right) d\Omega dt \quad (7.3)$$



**Figure 7.9:** Schematic view of a hexagonal cell of WCDs. The shaded area around the central station defines the acceptance of the hexagon.

Similarly, trigger configurations such as 4T5 may also be possible. However, as it is described in Sec. 7.2.1, it is statistically uncommon. Moreover, events detected with two faulty WCDs in the first crown may be wrongly reconstructed. Thus, by not considering the 4T5 configurations, a conservative estimation of the exposure is obtained.

The Id 1764 MD station at the center of the SD-433 may experience downtime as well. Then, a correction factor  $P_{MD}$ , which represents the chance of the MD station being active in a given observation period of time  $dt$ , must be included in the exposure calculation:

$$d\varepsilon \propto A_{\text{cell}} \times \cos\theta \times \left( P_{6T5}(t) + \frac{4}{6} \times P_{5T5}(t) \right) \times P_{MD}(t) d\Omega dt \quad (7.4)$$

This condition is included explicitly as a cut in Tab. 7.1. In the same fashion, the remaining three quality cuts (being the non-saturation and the maximum distance requirements) can also be included as time-dependent correction factors  $P_{QC}(t)$ :

$$d\varepsilon \propto A_{\text{cell}} \times \cos\theta \times \left( P_{6T5}(t) + \frac{4}{6} \times P_{5T5}(t) \right) \times P_{MD}(t) \times P_{QC}(t) d\Omega dt \quad (7.5)$$

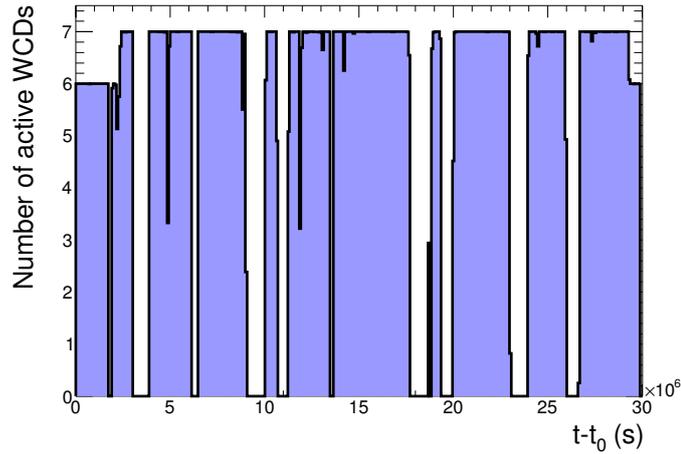
Lastly, the SD-433 is not fully efficient above  $E_0 = 10^{16.4}$  eV for  $\theta < \theta_{\text{max}} = 45^\circ$ , as discussed in Sec. 4.3.1. Thus, the observed flux is actually underestimated at the lowest energies of interest due to this effect. Thus, the exposure must be also corrected by the trigger efficiency  $P_{\text{trig}}(E, \theta)$ . Contrarily to the rest of the correction factors, this one introduces the energy dependence to the exposure calculation. Therefore, the energy spectrum of the cosmic-ray flux has to be included as well.

Wrapping up all the correction factors, the final expression for the exposure is obtained in Eq. 7.6:

$$\frac{d\varepsilon}{d\Omega dt dE} = A_{\text{cell}} \times \cos\theta \times \left( P_{6T5}(t) + \frac{4}{6} \times P_{5T5}(t) \right) \times P_{MD}(t) \times P_{QC}(t) \times \left( c \cdot \left( \frac{E}{\text{eV}} \right)^{-\gamma} \times P_{\text{trig}}(E, \theta) \right) \quad (7.6)$$

where the constant  $c$  is a scaling factor such as:

$$\int_{E_0}^{\infty} dE c \cdot E^{-\gamma} = 1 \quad (7.7)$$



**Figure 7.10:** The average number of active WCDs in the first crown of the SD-433 (including the central WCD). The bin width represents one day.

The exposure is then integrated in the observation time, the energy and the solid angle as:

$$\varepsilon = \int_{t_0}^t \int_{E_0}^{\infty} \int_0^{2\pi} \int_0^{\theta_{\max}} d\varepsilon(E, \theta, t) \quad (7.8)$$

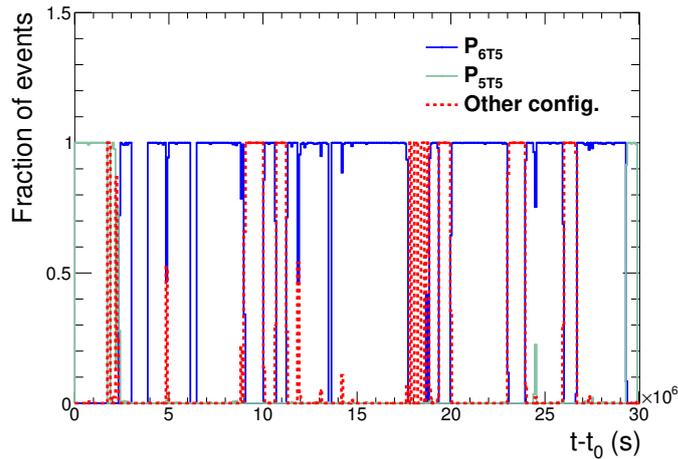
The ingredients for the exposure calculation are described and obtained in the following subsections.

### 7.2.1 Uptime of the detectors

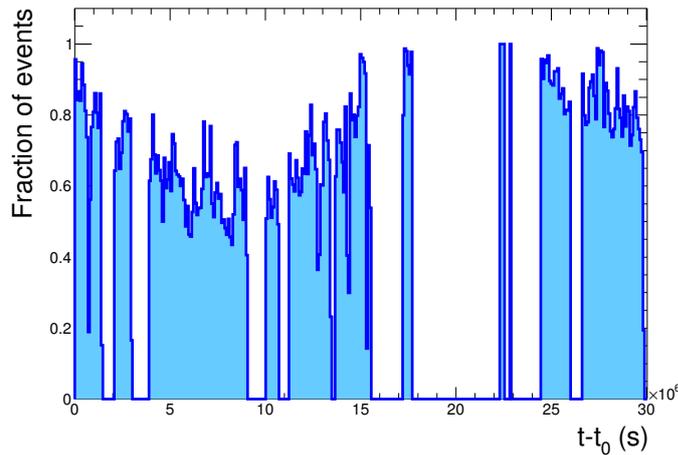
The statuses of all WCDs of the Auger Observatory is monitored and sent to the CDAS with every T3 request. In this way, the non-functional WCDs can be promptly identified in a nearly real-time basis. In the period under study, no significant behaviour is seen in any individual WCDs from the SD-433. This can be inferred from Fig. 7.10, which shows the average number of active WCDs in the first crown of the SD-433 on a daily basis. Apart from isolated gaps, no critical downtimes are identified.

As stated in Sec. 7.1.1, there was a lack of events between May 11 to July 28, 2016, which is translated in  $t - t_0 \simeq 18 \times 10^6$  s to  $t - t_0 \simeq 24 \times 10^6$  s. In this time interval, the SD-433 hexagon was active. Therefore, this loss of events must have been due to a failure further up the chain of data acquisition, which is accounted for in the exposure calculation as described in Sec. 7.13.

The correction factors  $P_{6T5}$  and  $P_{5T5}$  are calculated on a daily basis. For any given day, the number of events with the Id 1764 WCD and its first crown completely functional is compared to the total number of events. It is worth noting that this comparison does not need the events to be reconstructed by the SD-433, since the status of all WCDs in Auger are transmitted with each T3 request, independently of the intervening WCDs. In analogy, the 5T5 events can be compared to the total number of events on a daily basis as well. Both ratios are shown in Fig. 7.11. The 6T5 configuration was possible during most of the time, while only at the beginning and at the end of the data period, the acceptance was reduced to the 5T5 configuration due to one non-functional WCD, Id 97 and Id 13 respectively. The time intervals without either configuration are correlated



**Figure 7.11:** The fraction of events in a daily basis with a completely active first crown of the SD-433 (blue lines) and with six active WCDs (light blue). The remaining possible configurations, which are not considered for the exposure calculation, are symbolized by the red dotted histogram.

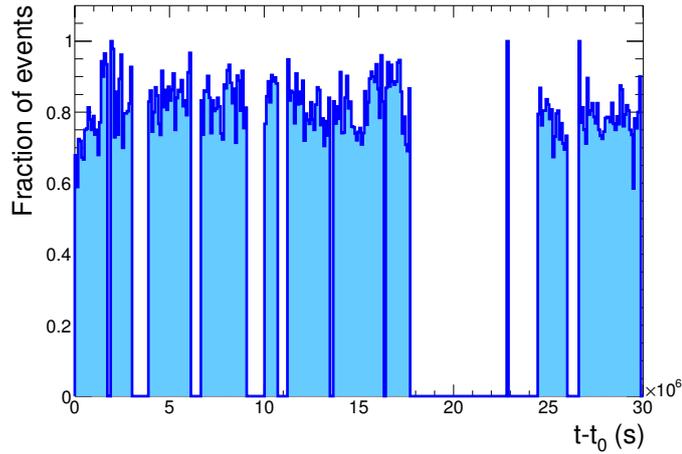


**Figure 7.12:** The fraction of events in which the Id 1764 MD station was active compared to the total number of events, averaged in a daily basis.

with a sustained lack of events, in view of Fig. 7.10, stressing the negligible presence of 4T5 events.

The MD stations are not independent from their associated WCD, since they receive the first level trigger from them. Thus, the status of a MD is constrained to the pre-requisite that the associated WCD must be active. In a similar fashion as the uptime monitoring for the SD, each event sent to the CDAS carries information about the status of all the modules of the MD stations. As in the case of the SD uptime, the fraction of events with an active Id 1764 MD station with respect to the number of total events in daily basis,  $P_{MD}$ , is displayed in Fig. 7.12. The inactive time periods are mostly correlated to the same offline periods seen in the SD-433 and in the daily rate of reconstructed events of Fig. 7.1. After an unstable period during the first few months (until  $t - t_0 \sim 15 \times 10^6$  s), the MD station reached an average daily uptime of  $\sim 89.5\%$ .

After the long downtime around  $t - t_0 \simeq 20 \times 10^6$  s, one of the  $10 \text{ m}^2$  modules was not operational, thus reducing the nominal area of the MD station to  $50 \text{ m}^2$ . Since the optic and electronics system operates at a single module level, a failure in one module does not necessarily alter the data acquisition of the MD station, as seen in Fig. 7.12. While this issue does not affect the measured



**Figure 7.13:** The fraction between the number of events passing the two quality cuts with respect to the number of events only passing the rest of the selection cuts.

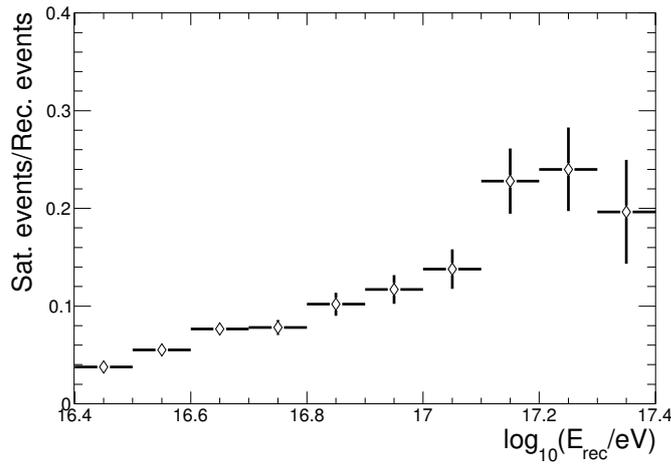
muon density, it produces an increase in the fluctuations of the observable at a statistical level, enhanced by the fact that the background events contaminating a hypothetical photon sample are those for which few muons are measured. The impact of the reduced sensitive area due to a malfunctioning module must thus be quantified in future studies.

## 7.2.2 Correction by quality cuts

The two quality cuts necessary for the discrimination procedure, namely the non-saturation condition and the maximum allowed distance from the MD station, are considered for the exposure calculation as explained before. The non-saturation constraint mainly affects high-energy events, while the distance cut is independent on the features of the primary particle.

The correction to the exposure due to the quality cuts,  $P_{QC}$ , is included as another time-dependent factor, such as the detector uptimes explained before. Following the same procedure, the fraction of events passing the two quality cuts with respect to the reconstructed events passing the rest of the selection cuts is shown in Fig. 7.13. The mean fraction of events passing the quality cuts is stable at  $\sim 85\%$ .

The energy distribution of the events passing the quality cuts is defined by the saturation condition, as the maximum distance cut impose an energy-independent rescaling. The fraction of saturated events to reconstructed events, verifying the elemental hexagon and the maximum distance conditions, is seen in Fig. 7.14. Because of the attenuation of the atmosphere, saturated events tend to have nearly-vertical incidence. It is worth noting that the saturation probability is similar to the estimations done with photon primaries in Sec. 4.4, which indicates that this quality cuts does not introduce a strong bias (favorable or not) related to the possible existence of photon primaries in the observed flux.



**Figure 7.14:** The fraction of saturated to reconstructed events in terms of the reconstructed energy. All events verify the elemental hexagon and the maximum allowed distance constraints. The events have zenith angles  $\theta_{\text{rec}} < 45^\circ$ .

### 7.2.3 Correction by trigger efficiency

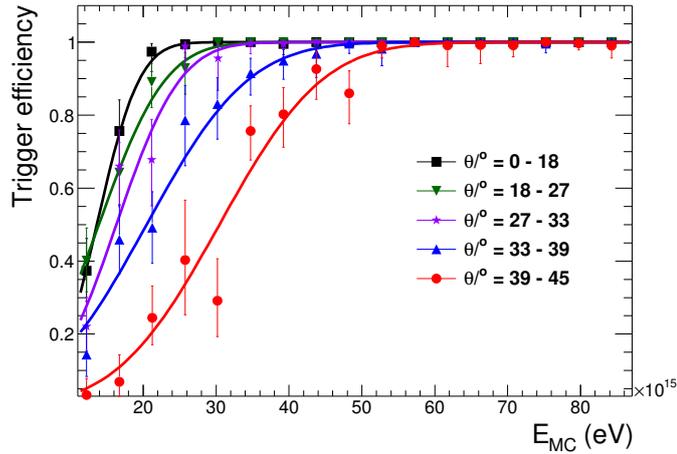
The cosmic-ray flux is partially observed at sufficiently low energies due to the trigger efficiency limitations of the SD-433, as discussed previously in Sec. 4.3.1. The showers with a non-vertical incidence have a smaller probability of generating a T3 trigger in an SD array as the primary energy decreases. The mass composition of the observed cosmic-ray flux is unknown, and thus an exact correction for the trigger efficiency can not be attained. However, a conservative scenario would be to consider that the flux has a pure proton composition (heavier primaries would generate an array trigger with a higher probability). On the other hand, the trigger efficiency for photon primaries may be different from the ones estimated under the hadronic primary hypothesis. However, the impact of using the same trigger efficiency as in the case of proton primaries is expected to be negligible in light of the discussion of Sec. 4.3.1, particular comparing Fig. 4.14 and Fig. 4.16, bottom.

The trigger efficiency is defined as the ratio between the number of events generating an array trigger and the total number of events. Additionally to the analysis performed with simulated events, in this case the trigger efficiency must be calculated while asking for the shower core to be located, at most, 200 m away from the central station. Also, a mathematical expression is needed for  $P_{\text{trig}}(E, \theta)$  to include in Eq. 7.6. The ansatz shown in Eq. 7.9 is applied, where  $a$  and  $b$  are free parameters and  $\text{Erf}(x)$  is the error function.

$$P_{\text{trig}}(E_{\text{MC}}, \theta) = \frac{1}{2} \times \left( \text{Erf} \left( a \times \frac{E_{\text{MC}}}{10^{16} \text{ eV}} + b \right) + 1 \right) \quad (7.9)$$

In Fig. 7.15, the trigger efficiency in different zenith bins is shown in terms of the simulated energy. The fitted parameters  $a$  and  $b$  for each profile are shown in Tab. 7.15.

The trigger efficiency correction is based on MC studies and, particularly, it is described in terms of the MC primary energy. However, the energy bias for proton primaries at the energies of interest is well contained within the intrinsic SD-433 energy resolution, as showed before in Fig. 4.27. Therefore, no significant bias is expected if the same fit parameters and expression are used for  $P_{\text{trig}}(E_{\text{rec}}, \theta)$ .



**Figure 7.15:** The trigger efficiency in terms of the simulated energy for proton events with different zenith angles, represented by the marker styles. The model of Eq. 7.9 is fitted to each profile.

Zenith interval ( $^{\circ}$ )	$a$	$b$
0 – 18	$(1.53 \pm 0.27)$	$(-2.09 \pm 0.44)$
18 – 27	$(0.980 \pm 0.234)$	$(-1.36 \pm 0.41)$
27 – 33	$(1.03 \pm 0.14)$	$(-1.67 \pm 0.28)$
33 – 39	$(0.642 \pm 0.088)$	$(-1.31 \pm 0.22)$
39 – 45	$(0.633 \pm 0.086)$	$(-1.93 \pm 0.26)$

**Table 7.2:** The parameters  $a$  and  $b$  of Eq. 7.9 fitted to events generated by primary protons with energies between  $E_{MC} = 10^{16}$  eV and  $10^{17}$  eV.

### 7.2.4 Exposure estimation

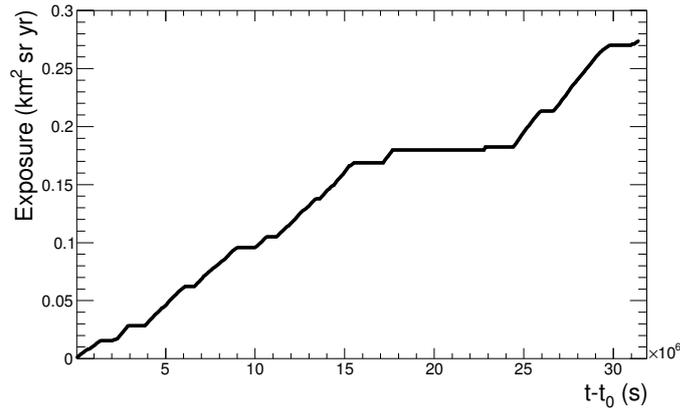
All the necessary ingredients for the exposure calculation have been obtained in the previous subsections. Therefore, the integral of Eq. 7.8 can be solved. The azimuthal integral leads to a factor of  $2\pi$ . The integral on the energy and the zenith angle are linked by  $P_{\text{trig}}(E_{\text{rec}}, \theta)$ . This double integral leads to:

$$\int_0^{\theta_{\text{max}}} d\theta \left( \int_{E_0}^{\infty} dE c \cdot \left( \frac{E}{\text{eV}} \right)^{-\gamma} \times P_{\text{trig}}(E, \theta) \right) \times \sin\theta \cos\theta = 0.901 \quad (7.10)$$

where  $c = 5.01 \times 10^{34}$  and  $\gamma = -3.06$ . The energy integral is actually performed up to  $10^{20}$  eV, which is a sufficiently high energy for the SD-433. The factor  $\sin\theta$  comes from the solid angle differential element while the  $\cos\theta$  comes from the acceptance expression. A correction of the order of 90% comes from the trigger inefficiency of the array.

The time integration takes into account the rest of the corrections and it is numerically calculated. It is an increasing function of the observation time  $t - t_0$ , i.e. the elapsed time since the first event of the data period. The exposure is thus an increasing function of  $t - t_0$  as well.

By adding the multiplicative factor  $A_{\text{cell}}$ , the exposure in terms of the observation time, as shown in Fig. 7.16, is obtained. The total accumulated exposure in the selected period comprising 364 net days is then  $0.274 \text{ km}^2 \text{ sr yr}$ .



**Figure 7.16:** The exposure of the SD-433 for the photon search with the observables  $M_b + Q$  for  $E_{\text{rec}} > 10^{16.4}$  eV and  $\theta < 45^\circ$  between the GPS 1129291205 and GPS 1160740813, comprising a net period of 364 days.

Background rejection $\rho$	Signal efficiency $\xi$
0.9990	0.897
0.9995	0.815
0.9997	0.793

**Table 7.3:** The signal efficiency at different fixed background rejection levels that could be attained with the  $M_b + Q$  discrimination procedure for events with  $E_{\text{rec}} > 10^{16.4}$  eV and  $\theta < 45^\circ$ , following a spectrum  $\sim E^{-\gamma}$  with  $\gamma = -3.06$ .

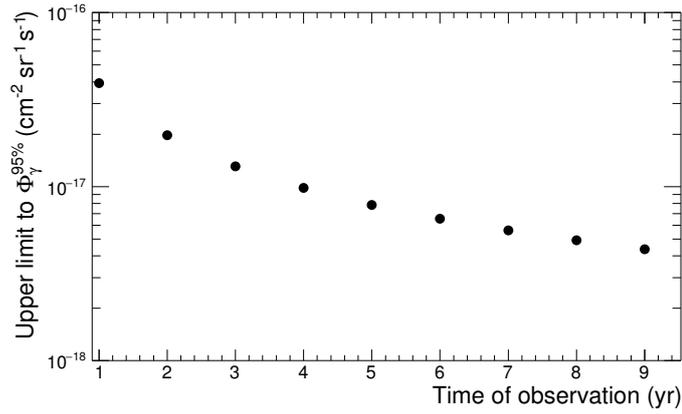
### 7.3 Sensitivity and upper limits to the photon flux

The total number of events of the data period available for this photon search is  $N = 5359$  with an accumulated exposure of  $\varepsilon = 0.274 \text{ km}^2 \text{ sr yr}$ . The discrimination procedure could be applied directly to the data to obtain the Fisher distribution and spot possible outlier events compatible with a photon origin. However, two important quantities can be studied without *unblinding* the data, this is, without having access to the observables  $M_b$  and  $Q$  of the  $N$  events.

The discrimination procedure has been applied to events following a continuous energy and zenith distribution in the ranges of interest in Sec. 6.3, particularly Fig. 6.22. The background rejection  $\rho$  and signal efficiency  $\xi$  that could be reached, provided that a suitable parametrization of the Fisher coefficients is applied, is summarized in Tab. 7.3. The fact that the signal efficiency is not exactly 100% in any of the quoted cases imposes a constraint to the estimation of the photon flux. Indeed, if no events survive the background rejection criterium of the procedure (i.e. no event has a Fisher parameter verifying  $F > F_{\text{cut}}$ ), it would indicate that the procedure is not sensitive enough to the presence of a tiny hypothetical photon component in the observed flux. Therefore, only an *upper limit to the photon flux* can be imposed.

The non-observation of photon candidate events,  $N_{\text{obs}} = 0$ , could be thought of as a statistical fluctuation from a photon flux which could have produced, in the selected observation period, a mean number  $N_\gamma$  of photon events. The probability  $P(N_{\text{obs}} | N_\gamma)$  to observe  $N_{\text{obs}}$  when the expected average is  $N_\gamma$  can be calculated by means of the standard Feldman–Cousins method for the Poisson distribution [257] described in the following.

Since this phenomenon of small counts can be modelled by Poissonian statistics, the mathematical expression for  $P(N_{\text{obs}} | N_\gamma)$  is:



**Figure 7.17:** The estimation of the upper limit to the photon flux at 95% CL and  $E > 10^{16.4}$  eV in terms of the time of observation.

$$P(N_{\text{obs}} | N_\gamma) = \frac{e^{-N_\gamma} N_\gamma^{N_{\text{obs}}}}{N_{\text{obs}}!} \quad (7.11)$$

A conservative estimation for  $N_\gamma$  can be obtained from the measurement of  $N_{\text{obs}}$  by asking that  $P < 0.05$ . In this way,  $N_\gamma$  is estimated at a 95% *confidence level* (CL). For the particular case of  $N_{\text{obs}} = 0$ ,  $N_\gamma^{95\%} \simeq 3$ . Therefore, an upper limit to the integral photon flux  $\Phi_\gamma$  above an energy  $E_0$  at 95% CL can be calculated as:

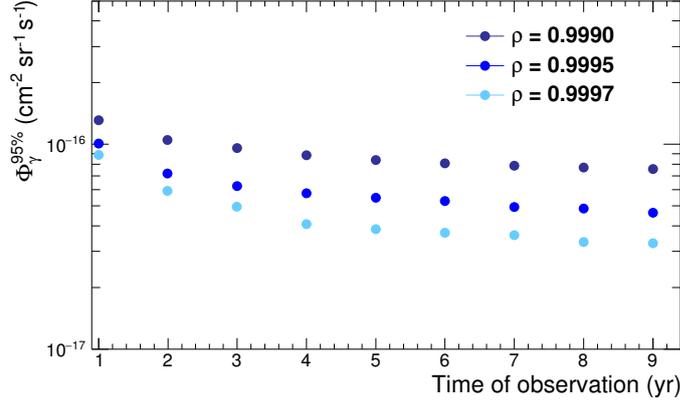
$$\Phi_\gamma^{95\%}(E_0) < \frac{N_\gamma^{95\%}}{\varepsilon(E_0)} \times \frac{1}{\xi(E_0)} \quad (7.12)$$

where the exposure (hence, the upper limit) depends on the minimum energy threshold applied in the analysis. The signal efficiency  $\xi$  must be included as well to account for the non-ideal separation power of the procedure. The upper limit for the selected data period can be obtained by replacing the parameters in Eq. 7.12. Assuming that for the benchmark rejection of  $\varrho = 0.9990$  and  $E_0 = 10^{16.4}$  eV no photon candidate events are seen in the selected data period, the upper limit to the photon flux obtained with the  $M_b + Q$  discrimination method at a 95% CL is:

$$\Phi_\gamma^{95\%}(E > 10^{16.4} \text{ eV}) < 3.87 \times 10^{-17} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} \quad (7.13)$$

The selected period corresponds to a net 364 days of data. Although several downtime periods affected the SD-433 and the Id 1764 MD station, a conservative extrapolation of the exposure for any given observation period can be obtained. The upper limit as a function of the observation time is shown in Fig. 7.17, where the first point corresponds to the estimation calculated before and the last point corresponds to the end of 2024 at the last year of the planned operation of the Observatory. As the observation time increases, the upper limits that could be attained are more stringent. In particular, after nine years of data, it is possible to enhance the current estimation by an order of magnitude, reaching  $4.29 \times 10^{-18} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$ .

The previous analysis was based on the fact that no photon candidate events were spotted in the selected data period. However, it could be also possible to detect a photon flux if it is prominent enough as to not be compatible with the expected background noise. At the same time, the existence of the photon flux is constrained by upper limits, in the energy domain of interest, imposed by other experiments such as KASCADE-Grande and EAS-MSU (recall Sec. 1.5.3).



**Figure 7.18:** The minimum photon flux not compatible with the background noise expectation at 95% CL in terms of the observation time. The markers represent three rejection criteria according to the  $M_b + Q$  procedure (see Tab. 7.3).

The expected number of background events surviving the discrimination method in the selected data period is  $N_{\text{noise}} = N \times \rho$ . Thus, the same argument as in the upper limits calculation can be summoned. Given a number of  $N_{\text{obs}}$  photon candidate events, the average number  $N_\gamma$  of photon-initiated events in the observed flux must be larger than  $N_{\text{noise}}$  as to reject the hypothesis that the  $N_{\text{obs}}$  events are all due to the background contamination at 95% CL. The number of expected noise events in the selected data period is  $N_{\text{noise}} = 5, 3$  and  $2$  at the background rejection levels of Tab. 7.3. Therefore, the minimum number of photons not compatible with noise at 95% CL is then 10, 8, 6, which can be translated to a hypothetical integral photon flux as in Eq. 7.12. If the number of photon candidate events in the selected data period is less than the mentioned numbers, then any detection can be explained as background fluctuations.

The requirement of a minimum number of photon events (thus, a minimum detected photon flux) to make a detection claim defines the *sensitivity* of the discrimination procedure. Considering the three possible scenarios for rejection and efficiency showed in Tab. 7.3, the minimum integral photon flux that can be detected with the  $M_b + Q$  discrimination method at a 95% CL is then:

$$\Phi_\gamma^{95\%} (E > 10^{16.4} \text{ eV}, \rho = 99.90\%) > 1.34 \times 10^{-16} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} \quad (7.14)$$

$$\Phi_\gamma^{95\%} (E > 10^{16.4} \text{ eV}, \rho = 99.95\%) > 1.09 \times 10^{-16} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} \quad (7.15)$$

$$\Phi_\gamma^{95\%} (E > 10^{16.4} \text{ eV}, \rho = 99.97\%) > 8.18 \times 10^{-17} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1} \quad (7.16)$$

In the same fashion as with the upper limits, the minimum integral photon flux after a more extended observation time is estimated in Fig. 7.18, where the three quoted flux estimations correspond to the first point of each scenario. By the end of 2024, an integral photon flux of  $3.32 \times 10^{-17} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$  above  $10^{16.4} \text{ eV}$ , not compatible with background expectation at 95% CL, could be detected.

The sensitivity estimation is based on the presence of a diffuse photon flux. Indeed, the number of background events  $N_{\text{noise}}$  are assumed to be isotropically distributed in the preceding calculations. Since the arrival directions of the cosmic rays are distributed isotropically below  $E_{\text{rec}} = 10^{18.5} \text{ eV}$  (see Sec. 1.3), an accumulation of events compatible with a photon primary (i.e.

$E_0$	$\xi \mid \varrho = 99.90\%$	$\xi \mid \varrho = 99.95\%$	$\xi \mid \varrho = 99.97\%$
$10^{16.4}$ eV	0.90	0.82	0.80
$10^{16.5}$ eV	0.93	0.92	0.91
$10^{16.6}$ eV	0.93	0.91	0.90
$10^{16.7}$ eV	0.94	0.92	0.90
$10^{16.8}$ eV	0.95	0.95	0.94
$10^{16.9}$ eV	0.97	0.97	0.97

**Table 7.4:** The signal efficiency  $\xi$  for three representative background rejection  $\varrho$  considering events with  $E_{\text{rec}} > E_0$  and  $\theta < 45^\circ$ . For each value of  $E_0$ , the parameters are estimated with dedicated Fisher runs as described in Sec. 6.2.2.

with a Fisher parameter verifying  $F > F_{\text{cut}}$ ) coming from a certain direction would be difficult to explain solely based on the expectation for  $N_{\text{noise}}$ . For example, considering that the angular resolution for the SD-433 has been estimated to be  $\sim 1.1^\circ$  [141], the probability that two uncorrelated events composing the sample  $N_{\text{noise}}$  arrive from the same direction is  $\sim 10^{-4}$  at  $\theta = 30^\circ$ , which is equivalent to  $> 3\sigma$ . Thus, a complementary criterium based on the arrival direction of the events may be applied to further refine the sensitivity of the  $M_b + Q$  method.

### 7.3.1 Comparison with other experiments

The energy domain above  $E > 10^{16.4}$  eV has been studied by a few experiments, as discussed previously in Sec. 1.5.3. The estimations of the upper limit and sensitivity to the photon flux presented before can also be extended to other values of  $E_0$ , which are only limited by the lack of statistics as deduced from Fig. 7.6.

The signal efficiency  $\xi$  at the benchmark rejection must also be estimated for each  $E_0$ . In order to do so, dedicated MVA runs increasing the value of  $E_0$  from  $10^{16.5}$  eV to  $10^{16.9}$  eV for  $\theta < 45^\circ$  are performed following the same procedure as in Sec. 6.2.2. The signal efficiency at three representative rejection levels for different values of  $E_0$  are presented in Tab. 7.4. For completeness, the values presented before in Tab. 7.3 are also included.

The estimated upper limits and sensitivity to the photon flux are shown in Tab. 7.5 for the selected data period ( $\sim 1$  year) and extrapolating the exposure until 2024. In the case of the upper limits, the background rejection is set to the benchmark of 99.90% as it is customary for photon searches. On the contrary, a 99.97% background rejection is set for the estimation of the minimum detectable flux. In all cases, the results are given at a 95% CL.

In Fig. 7.19, the upper limits under the mentioned conditions are compared to the results from other experiments extracted from Fig. 1.22. The upper limits estimations using the selected data period from the MD engineering array are in the same order of magnitude as the results from other experiments, with the advantage of the increased CL. However, the data accumulated until the end of the Auger operation would allow to obtain upper limits, if no photon candidate events are found in the data, nearly one order of magnitude below all the current estimations, making the Auger photon search the best one in this energy domain. Particularly, these stringent upper limits would defy some of the astrophysical models explaining the observed neutrino excess in the IceCube measurements.

$E_0$	Upper limit to $\Phi_\gamma$ ( $\text{cm}^{-2} \text{sr}^{-1} \text{s}^{-1}$ )		Minimum detectable $\Phi_\gamma$ ( $\text{cm}^{-2} \text{sr}^{-1} \text{s}^{-1}$ )	
	$\sim 1$ year	9 years	$\sim 1$ year	9 years
$10^{16.4}$ eV	$3.87 \times 10^{-17}$	$4.29 \times 10^{-18}$	$8.18 \times 10^{-17}$	$3.32 \times 10^{-17}$
$10^{16.5}$ eV	$3.54 \times 10^{-17}$	$3.93 \times 10^{-18}$	$6.23 \times 10^{-17}$	$2.06 \times 10^{-17}$
$10^{16.6}$ eV	$3.50 \times 10^{-17}$	$3.88 \times 10^{-18}$	$5.72 \times 10^{-17}$	$1.57 \times 10^{-17}$
$10^{16.7}$ eV	$3.49 \times 10^{-17}$	$3.88 \times 10^{-18}$	$5.38 \times 10^{-17}$	$1.22 \times 10^{-17}$
$10^{16.8}$ eV	$3.56 \times 10^{-17}$	$3.96 \times 10^{-18}$	$5.10 \times 10^{-17}$	$8.38 \times 10^{-18}$
$10^{16.9}$ eV	$3.72 \times 10^{-17}$	$4.13 \times 10^{-18}$	$5.20 \times 10^{-17}$	$7.54 \times 10^{-18}$

**Table 7.5:** The estimated upper limits to the photon flux and the minimum detectable flux considering different energy thresholds  $E_0$  and the quoted observation periods. The results are given at 95% CL. In the case of the upper limits, the background rejection is set to 99.90% while for the detectable flux it is set to 99.97%.

The minimum detectable flux against the threshold energy is shown in Fig. 7.20. The sensitivity would not be enough to detect a photon flux below the upper limits already imposed by other experiments with the current exposure. However, considering the extrapolated exposure until the end of the planned operation of Auger, it may be possible to detect a photon flux below the upper limits imposed by all the experiments in the energy range between  $10^{16.5}$  eV and  $10^{17}$  eV. It is worth stressing that a detection claim would be possible only if the photon flux is above the sensitivity profile and below previous upper limits, i.e. if  $0.8 \text{ cm}^{-2} \text{sr}^{-1} \text{s}^{-1} \lesssim \Phi_\gamma^{95\%} \times 10^{17} \lesssim 3 \text{ cm}^{-2} \text{sr}^{-1} \text{s}^{-1}$  between  $10^{16.5}$  eV and  $10^{17}$  eV.

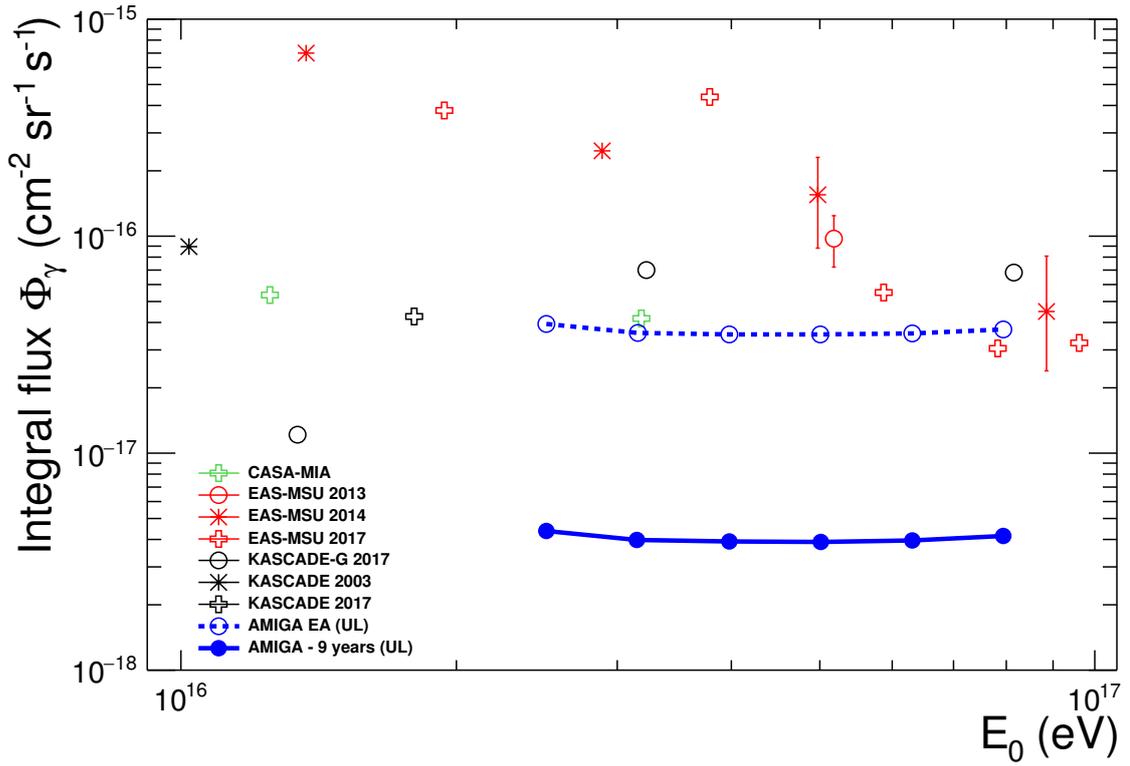
The sensitivity of the procedure can be enhanced by increasing the separation power of the procedure, which could be accomplished by using a more advanced multivariate technique, as explained before. However, an extension of the SD-433 and its MD counterpart, may also enhance the detection method developed in this work. Since July 2017, three WCDs positions were added in the SD-433, completing three hexagons in its second crown. Thus, if MD stations were deployed in the central positions of each hexagon, the acceptance of the photon search would be multiplied by a factor four. Supposing that the three MD stations are functional by the end of 2018, the accumulated exposure by 2024 would be suitable to detect an integral photon flux above  $1.11 \times 10^{-17} \text{ cm}^{-2} \text{sr}^{-1} \text{s}^{-1}$  for  $E_0 = 10^{16.4}$  eV or above  $2.81 \times 10^{-18} \text{ cm}^{-2} \text{sr}^{-1} \text{s}^{-1}$  for  $E_0 = 10^{16.8}$  eV, as shown in Fig. 7.21, leading to an increase by a factor 3 of the photon sensitivity between  $10^{16.5}$  eV and  $10^{17}$  eV.

## 7.4 Outlook and future prospects for the photon search

The upper limits and sensitivity estimations presented in this thesis are based only on statistical assumptions, i.e. without applying the discrimination procedure with the composition-sensitive observables of the events. This approach is usually named *blind search*.

The mass composition of the cosmic-ray flux below  $10^{17}$  eV has not been explored in detail by Auger. Particularly, the photon searches have been conducted in the energy domain above  $10^{18}$  eV by employing SD- and FD-related observables. Thus, this study leads to the first photon search below  $10^{18}$  eV, offering a unique possibility to impose the best upper limits to the photon flux or even detect a minuscule photon flux given enough years of observation.

The method itself is based on robust and well-understood foundations. However, some im-

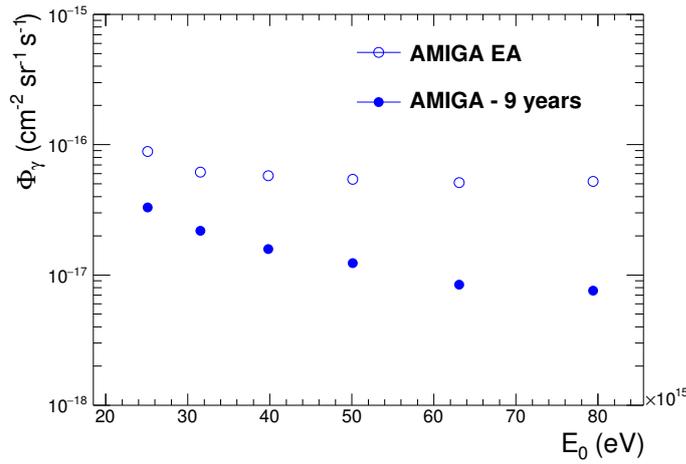


**Figure 7.19:** The upper limits to the integral photon flux in terms of the threshold energy. The estimations from this thesis are represented by blue markers. The calculation with the selected data period is represented by empty markers and that performed by extrapolating the exposure until 2024 by filled markers. The results are obtained with a 95% CL. The upper limits (detection claims) from EAS-MSU are represented by red markers without (with) error bars [121, 122, 123], KASCADE and KASCADE-Grande by black markers [258, 125, 126] and CASA-MIA by green markers [259]. For these experiments, the upper limits are given with a 90% CL. The points were extracted from [126].

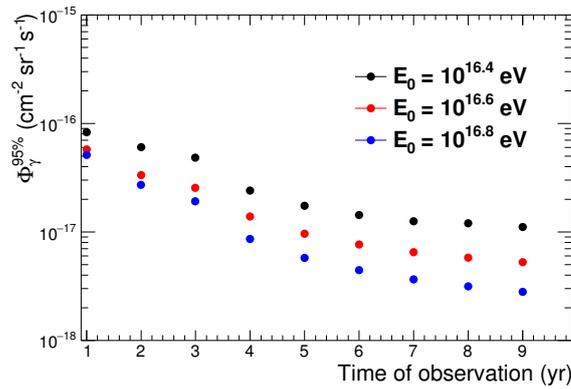
improvements can be made. For example, the usage of the linear Fisher analysis may well be outperformed by a non-linear analysis such as Boosted Decision Trees, which in turn would yield a better background rejection and signal efficiency in the energy range of interest. The sensitivity can be further enhanced by including the arrival direction of the events in the MVA discrimination method. From the point of view of the statistics in the MC studies, the lack of a sufficiently numerous CORSIKA shower libraries pointed out the need for complex sampling procedures such as bootstrap and spectral shuffle, which may be easily avoided by a production of dedicated libraries with a suitable energy and zenith angle distributions.

One crucial aspect that is left for future studies is the proper assessment of the systematic effects discussed in Sec. 6.5. While most of effects may indicate that the current analysis is performed in one of the worst case scenarios, a systematic undercounting in the MD stations could lead to the detection of false positive events. However their occurrence would be reduced by the additional information provided by the observable  $Q$ .

The analyses included in this thesis have been presented several times in the Auger Collaborations meetings for the past years, where the author received valuable comments, corrections and ideas from the Auger experts. The data *unblinding* is foreseen to be carried out once the systematics effects mentioned before are properly quantified. As a consequence, the Photon Analysis Group



**Figure 7.20:** The minimum integral photon flux not compatible to the expected background contamination at a 95% CL in terms of the threshold energy. The profiles correspond to the accumulated exposure in the selected data period (empty markers) and to the extrapolated until 2024 (filled markers).



**Figure 7.21:** The minimum integral photon flux that can be detected against the time of observation considering three different energy thresholds. The exposure is calculated as explained before until the end of 2018 (i.e. for the first three years of data). From the fourth year, an increased acceptance is considered (see text for details).

has strongly suggested that this study, applied to the data, will be the basis of a new full-author list paper of the Auger Collaboration.

The MD base design is the basis of the upcoming UMD in the framework of the AugerPrime upgrade, which will be deployed in the positions of the SD-750. Thus, it will be able to apply the  $M_b + Q$  in the energy domain above  $10^{17.5}$  eV for which the SD-750 is fully efficient (see Sec. 3.3.5). In this way, the study presented in this thesis offers a new tool for the photon searches that could be extended, with suitable changes, at a higher energy domain.

## Chapter 8

# Conclusions

The search for ultra-high energy photons above  $10^{16.5}$  eV was the main motivation for this thesis, extending in more than order of magnitude the current photon studies in the Auger Collaboration. The analyses included in this thesis represent one of the few mass composition studies at this energy domain and the only one employing the AMIGA muon detector.

Two new composition-sensitive observables were designed and extensively studied in this thesis, specially efficient in exploiting the expected differences in the EAS development between the photon and the hadronic primaries. Both of the observables were particularly tuned to be used with the SD-433 and the MD station at its center. On one hand, the observable  $M_b$  combines the muon density measured by a MD station with its distance to the shower axis as to profit from the large difference in terms of the produced muon content between photon and hadronic primaries. On the other hand, the observable  $Q$  is based on the well-known statistical test of the likelihood ratio between two models for the LDF measured by the SD, being the standard data-oriented model and a dedicated one, also designed in this thesis, for the photon-initiated EAS.

The observable  $M_b$  was optimized to obtain the best separation between photon and proton events for events with  $E_{\text{rec}} > 10^{16.4}$  eV and  $\theta < 45^\circ$ . By selecting unsaturated events sufficiently close to the MD station, the choice of  $b = 1.5$ , representative of the slope of the lateral distribution of muons, led to the best performance in terms of discrimination with a merit factor between photon and proton primaries rising from 4 to 7. The combination with the observable  $Q$  allowed for a composite background rejection suited to further enhance the signal efficiency and the stability of the discrimination method. By performing a multivariate analysis based on the Fisher discriminant, a signal efficiency of at least 89% above  $E_{\text{rec}} = 10^{16.4}$  eV while rejecting 99.90% is easily attained, thus reaching an unprecedented separation power and detection efficiency in the photon search studies in the Auger Collaboration.

The development of the discrimination procedure has been done in one of the most dis-favourable scenarios in terms of systematic effects. A pure-proton sample of simulated EAS, which are known to have a muon deficit with respect to data, has been employed as a background for photon simulated EAS. Furthermore, the parameters of the muon counting in the MD station may have produced an underestimation in the measured muon density as well. Even in this conservative scenario, the overall performance of the research presented in this thesis has attracted much attention in the Auger Collaboration and it is foreseen to be the basis of the next photon search full-author list paper.

The calculation of the exposure using one year of data reconstructed by the SD-433 allowed

to estimate the potential of the designed discrimination method. A blind analysis indicated that, if no events compatible with the photon origin are found in data, an upper limit to the integral photon flux of  $3.87 \times 10^{-17} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$  could be placed above  $10^{16.4} \text{ eV}$  at a 95% CL, which is a limit already competitive with results from other experiments. By the end of 2024, the upper limit could be improved by nearly one order of magnitude with respect to all the other cosmic-ray experiments, being  $4.29 \times 10^{-18} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$ . This upper limit would already be enough to rule out some of the IceCube neutrino excess model that are based on nearby Galactic sources.

The  $M_b + Q$  discrimination method represents a unique tool that, given its outstanding performance in terms of background rejection and signal efficiency, would allow to detect an integral photon flux not compatible with background noise fluctuations at 95% CL and below the most stringent upper limits by all the cosmic-ray experiments. After nine years of data, an integral photon flux of  $3.32 \times 10^{-17} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$  and  $7.54 \times 10^{-18} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$  above  $10^{16.4} \text{ eV}$  and  $10^{16.9} \text{ eV}$ , respectively, could be measured. This is the first time that a method is technically capable of such a feat within the Auger Collaboration.

A clear identification of events compatible with the photon primary hypothesis at the studied energies could shed light over several open problems in Astroparticle Physics, while being a milestone in the Gamma-ray Astronomy. Consequently, the  $M_b + Q$  method opens the possibility for the search for ultra-high energy photons in an unexplored energy domain in the Auger Observatory with an outstanding discovery potential at the doorsteps of the AugerPrime era.

# Bibliography

- [1] D. Pacini. La radiazione penetrante alla superficie ed in seno alle acque. *Nuovo Cim.*, VI/3:93–100, 1912.
- [2] R. A. Millikan and G. H. Cameron. The origin of the cosmic rays. *Phys. Rev.*, 32:533–557, 1928.
- [3] A. H. Compton. A geographic study of cosmic rays. *Phys. Rev.*, 43:387–403, 1933.
- [4] C. Anderson. The Positive Electron. *Phys. Rev.*, 43(491), March 1933.
- [5] S. Neddermeyer and C. Anderson. Note on the Nature of Cosmic-Ray Particles. *Phys. Rev.*, 51(884), 1937.
- [6] P. Auger et al. Extensive cosmic-ray showers. *Rev. Mod. Phys.*, 11:288–291, 1939.
- [7] B. Falkenburg and W. Rhode. *From Ultra Rays to Astroparticles*. Springer Netherlands, 2012.
- [8] E. Fermi. On the origin of the cosmic radiation. *Phys. Rev.*, 75:1169–1174, 1949.
- [9] E. Fermi. Galactic magnetic fields and the origin of cosmic radiation. *Astrophys. J.*, 119:1–6, 1954.
- [10] G. W. Clark et al. Cosmic-ray air showers at sea level. *Phys. Rev.*, 122:637–654, 1961.
- [11] J. Linsley. Evidence for a primary cosmic-ray particle with energy  $10^{20}$  eV. *Phys. Rev. Lett.*, 10:146–148, 1963.
- [12] A. A. Penzias and R. W. Wilson. A measurement of excess antenna temperature at 4080 Mc/s. *Astrophys. J.*, 142:419–421, 1965.
- [13] K. Greisen. End to the Cosmic-Ray Spectrum? *Phys. Rev. Lett.*, 16(748), April 1966.
- [14] D. J. Bird et al. Detection of a cosmic ray with measured energy well beyond the expected spectral cutoff due to cosmic microwave radiation. *Astrophys. J.*, 441(1):144–150, March 1995.
- [15] The Pierre Auger Collaboration. Observation of a large-scale anisotropy in the arrival directions of cosmic rays above  $8 \times 10^{18}$  eV. *Science*, 357:1266–1270, 2017.
- [16] W. F. G. Swann. A Mechanism of Acquirement of Cosmic-Ray Energies by Electrons. *Phys. Rev.*, 43(4):217, February 1933.
- [17] K. S. Cheng et al. Energetic radiation from rapidly spinning pulsars. II. Vela and Crab. *Astrophys. J.*, 300:522–539, January 1986.

- [18] P. Blasi, R. I. Epstein, and A. V. Olinto. Ultra-High-Energy Cosmic Rays from Young Neutron Star Winds. *Astrophys. J. Lett.*, 533(2), April 2000.
- [19] A. Venkatesan, M. Coleman Miller, and A. Olinto. Constraints on the production of ultra-high-energy cosmic rays by isolated neutron stars. *Astrophys. J.*, 484:323–328, 1997.
- [20] A. Neronov, D. Semikoz, and I. Tkachev. Ultra-high energy cosmic ray production in the polar cap regions of black hole magnetospheres. *New J. Phys.*, 11, June 2009.
- [21] P. Bhattacharjee and G. Sigl. Origin and Propagation of Extremely High Energy Cosmic Rays. *Phys. Rept.*, 327(3):109–247, April 2000.
- [22] L. Anchordoqui, T. Paul, S. Reucroft, and J. Swain. Ultrahigh Energy Cosmic Rays: The state of the art before the Auger Observatory. *Int. J. Mod. Phys. A*, 18(13):2229, May 2003.
- [23] W. Axford. The acceleration of cosmic rays by shock waves. *Annals of the New York Academy of Sciences*, 375(1):297–313, December 1981.
- [24] L. O’C Drury. An introduction to the theory of diffusive shock acceleration of energetic particles in tenuous plasmas. *Rep. Prog. Phys.*, 46(8):973, 1983.
- [25] S. A. Colgate. Acceleration in astrophysics. *Physica Scripta*, 1994(T52), 1994.
- [26] A. M. Hillas. The Origin of Ultra-High-Energy Cosmic Rays. *Ann. Rev. Astron. Astrophys.*, 22:425–444, 1984.
- [27] D. Torres and et al. Supernova remnants and gamma-ray sources. *Phys. Rept.*, 382(6):303–380, August 2003.
- [28] M. Nagano and A. A. Watson. Observations and implications of the ultra-high energy cosmic rays. *Rev. Mod. Phys.*, 72:689–732, 2000.
- [29] V. L. Ginzburg and S. I. Syrovatskii. *The Origin of Cosmic Rays*. Pergamon, 1964.
- [30] The Fermi-LAT Collaboration. Detection of the Characteristic Pion-Decay Signature in Supernova Remnants. *Science*, 339:807–811, 2013.
- [31] HESS Collaboration. Acceleration of petaelectronvolt protons in the Galactic Centre. *Nature*, 531:476–479, 2016.
- [32] P. Biermann. The origin of the highest energy cosmic rays. *J. Phys. G: Nucl. Part. Phys.*, 23(1), 1997.
- [33] A. W. Graham. Populating the Galaxy Velocity Dispersion: Supermassive Black Hole Mass Diagram, A Catalogue of (Mbh,  $\sigma$ ) Values. *Publications of the Astronomical Society of Australia*, 25(4):167–175, 2008.
- [34] F. Halzen and E. Zas. Neutrino fluxes from active galaxies: a model-independent estimate. *Astrophys. J.*, 488:669–674, 1997.
- [35] G. Sigl, M. Lemoine, and P. Biermann. Ultra-high energy cosmic ray propagation in the local supercluster. *Astropart. Phys.*, 10(2-3):141–156, April 1999.
- [36] P. P. Kronberg. Extragalactic magnetic fields. *Reports on Progress in Physics*, 57(4):325–382, 1994.

- [37] Robert Nemiroff (MTU) & Jerry Bonnell (UMCP). A service of: ASD at NASA / GSFC & Michigan Tech. U., January 2015. Image Credit: X-ray: NASA/CXC/SAO; Optical: NASA/STScI; Radio: NSF/NRAO/AUI/VLA.
- [38] P. Blasi and A. V. Olinto. Magnetized local supercluster and the origin of the highest energy cosmic rays. *Phys. Rev. D*, 59(023001), December 1998.
- [39] G. J. Fishman and C. A. Meegan. Gamma-Ray Bursts. *Annual Review of Astronomy and Astrophysics*, 33:415–458, September 1995.
- [40] C. A. Meegan et al. Spatial distribution of gamma-ray bursts observed by BATSE. *Nature*, 355:143–145, January 1992.
- [41] E. Waxman. Cosmological Gamma-Ray Bursts and the Highest Energy Cosmic Rays. *Phys. Rev. Lett.*, 75(386), July 1995.
- [42] A. V. Olinto. Ultra high energy cosmic rays: the theoretical challenge. *Phys. Rept.*, 333:329–348, 2000.
- [43] G. Lemaitre. The Primeval Atom - an Essay on Cosmogony. *D. Van Nostrand Co*, 1950.
- [44] M. Takeda et al. Extension of the Cosmic-Ray Energy Spectrum beyond the Predicted Greisen-Zatsepin-Kuz'min Cutoff. *Phys. Rev. Lett.*, 81(6):1163, August 1998.
- [45] L. Gonzalez-Mestre. Observing air showers from cosmic superluminal particles. In *AIP Conference Proceedings*, volume 433, page 418, 1998.
- [46] T. Stanev. *Ultra High Energy Cosmic Rays*. SLAC Summer Institute on Particle Physics, 2004.
- [47] R. Aloisio, V. Berezhinsky, and M. Kachelriess. Fragmentation functions in supersymmetric QCD and ultrahigh energy cosmic ray spectra produced in top-down models. *Phys. Rev. D*, 69(094023), 2004.
- [48] T. A. Clark, L. W. Brown, and J. K. Alexander. Spectrum of the Extra-galactic Background Radiation at Low Radio Frequencies. *Nature*, 228:847–849, November 1970.
- [49] R. Aloisio, S. Matarrese, and A. V. Olinto. Super Heavy Dark Matter in light of BICEP2, Planck and Ultra High Energy Cosmic Rays Observations. *Journal of Cosmology and Astropart. Phys.*, 2015, August 2015.
- [50] O.K. Kalashev and M. Kuznetsov. Constraining heavy decaying dark matter with the high energy gamma-ray limits. *Phys. Rev. D*, 94(063535), 2016.
- [51] M. Kachelriess. Particle physics solutions to the UHECR puzzle. *The European Physical Journal C*, 33(1):950–952, July 2004.
- [52] P. Sreekumar et al. EGRET Observations of the Extragalactic Gamma-Ray Emission. *Astrophys. J.*, 494(2), 1998.
- [53] G. Sigl. Ultrahigh-Energy Cosmic Rays: Physics and Astrophysics at Extreme Energies. *Science*, 291(73), 2001.
- [54] T. J. Weiler. Cosmic-ray neutrino annihilation on relic neutrinos revisited: a mechanism for generating air showers above the Greisen-Zatsepin-Kuzmin cutoff. *Astropart. Phys.*, 11(3):303–316, July 1999.

- [55] K. Kotera and A. Olinto. The Astrophysics of Ultrahigh Energy Cosmic Rays. *Annual Review of Astronomy and Astrophysics*, 49:119–153, 2011.
- [56] R. Jansson et. al. Constraining models of the large scale Galactic magnetic field with WMAP5 polarization data and extragalactic Rotation Measure sources. *Journal of Cosmology and Astropart. Phys.*, 2009, 2009.
- [57] T. K. Gaisser. *Cosmic Rays and Particle Physics*. Cambridge University Press, 1990.
- [58] F. W. Stecker, M. A. Malkan, and S. T. Scully. Intergalactic Photon Spectra from the Far-IR to the UV Lyman Limit for  $0 < z < 6$  and the Optical Depth of the Universe to High-Energy Gamma Rays. *Astrophys. J.*, 648(2), 2006.
- [59] J. W. Cronin. The highest-energy cosmic rays. *Nucl. Phys. B*, 138:465–491, 2005.
- [60] M. S. Longair. *High Energy Astrophysics: Stars, the Galaxy and the Interstellar Medium*. Cambridge University Press, 1994.
- [61] J. R. Hörandel. Cosmic-ray abundances and energy spectra at high energies: Measurements with TRACER and KASCADE. *Advances in Space Research*, 38(7):1549–1557, 2006.
- [62] J. A. Simpson. Elemental and Isotopic Composition of the Galactic Cosmic Rays. *Ann. Rev. Nucl. Part. Sci.*, 33:323–381, 1983.
- [63] J. Blümer, R. Engel, and J. R. Hörandel. Cosmic rays from the knee to the highest energies. *Progress in Particle and Nuclear Physics*, 63(2):293–338, 2009.
- [64] K. Lodders. Solar System Abundances and Condensation Temperatures of the Elements. *Astrophys. J.*, 591(2), 2003.
- [65] V. Verzi. Cosmic Rays: air showers from low to high energies - Rapporteur Report. In *34<sup>th</sup> International Cosmic Ray Conference*, 2015.
- [66] G. V. Kulikov and G. B. Khristiansen. On the Size Spectrum of Extensive Air Showers. *Sov. Phys. JETP*, 8(441):441, 1959.
- [67] K. Kampert for The KASCADE Collaboration. Cosmic ray energy spectra and mass composition at the knee - recent results from KASCADE. *Nucl. Phys. B. (Proc. Suppl.)*, 136:273–281, 2004.
- [68] The KASCADE-Grande Collaboration. Ankle-like feature in the energy spectrum of light elements of cosmic rays observed with KASCADE-Grande. *Phys. Rev. D*, 87(081101), 2013.
- [69] S. Knurenko, I. Petrov, Z. Petrov, and I. Slepsov. The Spectrum of Cosmic Rays in the Energy Range  $10^{16}$  -  $10^{18}$  eV According to the Small Cherenkov Array in Yakutsk. In *34<sup>th</sup> International Cosmic Ray Conference*, 2015.
- [70] The IceCube Collaboration. Cosmic Ray Physics with the IceCube Observatory. In *34<sup>th</sup> International Cosmic Ray Conference*, 2015.
- [71] C. C. H. Jui for the Telescope Array Collaboration. Summary of Results from the Telescope Array Experiment. In *34<sup>th</sup> International Cosmic Ray Conference*, 2015.
- [72] The High Resolution Fly’s Eye Collaboration. Cosmic Rays Composition Measurements by the HiRes Stereo . In *30<sup>th</sup> International Cosmic Ray Conference*, 2007.

- [73] M. Nagano et al. Energy spectrum of primary cosmic rays between  $10^{14.5}$  and  $10^{18}$  eV. *J. Phys. G: Nucl. Phys.*, 19:1295–1310, 1984.
- [74] W. D. Apel for The KASCADE Collaboration. Kneelike structure in the spectrum of the heavy component of cosmic rays observed with KASCADE-Grande. *Phys. Rev. Lett.*, 107:171104, 2011.
- [75] C. Patrignani et al. The Review of Particle Physics. *Chin. Phys. C*, 40(100001), 2016.
- [76] P. Ghia for the Pierre Auger Collaboration. Highlights from the Pierre Auger Observatory. In *34<sup>th</sup> International Cosmic Ray Conference*, 2015.
- [77] F. Fenu for the Pierre Auger Collaboration. The cosmic ray energy spectrum measured using the Pierre Auger Observatory. In *35<sup>th</sup> International Cosmic Ray Conference*, 2017.
- [78] R. J. Protheroe and P. L. Biermann. A new estimate of the extragalactic radio background and implications for ultra-high energy gamma-ray propagation. *Astropart. Phys.*, 6(1):45–54, 1996.
- [79] R. C. Gilmore et. al. Semi-analytic modelling of the extragalactic background light and consequences for extragalactic gamma-ray spectra. *Mon. Not. R. Astron. Soc.*, 422:3189–3207, 2012.
- [80] F. W. Stecker and M. H. Salamon. Photodisintegration of ultrahigh energy cosmic rays: A new determination. *Astrophys. J.*, 512:521–526, 1999.
- [81] D. Hooper, A. Taylor, and S. Sarkar. Cosmogenic photons as a test of ultra-high energy cosmic ray composition. *Astropart. Phys.*, 34(6), January 2011.
- [82] G. Gelmini et al. GZK photons as ultra-high-energy cosmic rays. *J. Exp. Theo. Phys.*, 106:1061, June 2008.
- [83] V. Berezhinsky, M. Kachelriess and A. Vilenkin. Ultrahigh Energy Cosmic Rays without Greisen-Zatsepin-Kuzmin Cutoff. *Phys. Rev. Lett.*, 79(4302), December 1997.
- [84] O. E. Kalashev, G. I. Rubtsov, and S. V. Troitsky. Sensitivity of cosmic-ray experiments to ultrahigh-energy photons: Reconstruction of the spectrum and limits on the superheavy dark matter. *Phys. Rev. D*, 80(103006), November 2009.
- [85] A. Kusenko, J. Schissel, and F. Stecker. Interactions of ultrahigh-energy cosmic rays with photons in the galactic center. *Astropart. Phys.*, 25, 2006.
- [86] Gopal-Krishna et. al. Ultra-high energy cosmic rays from Centaurus A: Jet interaction with gaseous shells. *Astrophys. J.*, 720(2), 2010.
- [87] M. Lemoine-Goumard. Status of ground-based gamma-ray astronomy. In *34<sup>th</sup> International Cosmic Ray Conference*, 2015.
- [88] M. Holler et al. Photon Reconstruction for H.E.S.S. Using a Semi-Analytical Shower Model. In *34<sup>th</sup> International Cosmic Ray Conference*, 2015.
- [89] W. Hofmann. HESS highlights. In *29<sup>th</sup> International Cosmic Ray Conference*, 2005.
- [90] W. Benbow. Cosmic ray background rejection and spectral analysis techniques. In *Towards a Network of Atmospheric Cherenkov Detectors VII Palaiseau*, 2005.

- [91] M. Beilicke. HESS observations of extragalactic objects. *Astrophys. Space Sci.*, 309:139–145, 2007.
- [92] D. Mazin for the MAGIC Collaboration. Upgrade of the MAGIC telescopes. In *33<sup>rd</sup> International Cosmic Ray Conference*, 2013.
- [93] M. Teshima et al. Tau neutrino search with the MAGIC telescope. In *30<sup>th</sup> International Cosmic Ray Conference*, 2007.
- [94] N. Sidro. Discovery of the microquasar LSI +61 303 at VHE gamma-rays with MAGIC. In *Proceedings of the VI Microquasar Workshop: Microquasars and Beyond*, 2006.
- [95] The MAGIC Collaboration. Discovery of Very High Energy Gamma Radiation from IC 443 with the MAGIC Telescope. *Astrophys. J. Lett.*, 664(2), 2007.
- [96] The MAGIC Collaboration. Observation of Pulsed  $\gamma$ -Rays Above 25 GeV from the Crab Pulsar with MAGIC. *Science*, 322(5905):1221–1224, 2008.
- [97] M. Spurio. *Particles and Astrophysics: A Multi-Messenger Approach*. Springer, 2015.
- [98] S. Funk. Ground- and Space-Based Gamma-Ray Astronomy. *Annu. Rev. Nucl. Part. Sci.*, 65:245–277, 2015.
- [99] L. Tibaldo for the Fermi-LAT Collaboration. A tale of cosmic rays narrated in gamma rays by Fermi. *Brazilian Journal of Physics*, 44(5):600–608, 2013.
- [100] M. Ahlers and K. Murase. Probing the Galactic origin of the IceCube excess with gamma rays. *Phys. Rev. D*, 90(023010), July 2014.
- [101] M. Ackermann et al. The spectrum of isotropic diffuse gamma-ray emission between 100 MeV and 820 GeV. *Astrophys. J.*, 799(1), 2015.
- [102] K. Murase, M. Ahlers, and B. C. Lacki. Testing the hadronuclear origin of PeV neutrinos observed with IceCube. *Phys. Rev. D*, 88(121301), 2013.
- [103] F. Halzen. High-energy neutrino astrophysics. *Nature*, 13, 2017.
- [104] M. Aartsen et al. Atmospheric and astrophysical neutrinos above 1 TeV interacting in IceCube. *Phys. Rev. D*, 91(022001), 2014.
- [105] J. V. Jelley. High-Energy gamma-ray absorption in space by a 3.5 K Microwave Field. *Phys. Rev. Lett.*, 16(11):479, March 1966.
- [106] S. Sarkar. New physics from ultrahigh-energy cosmic rays. In *Acta Phys. Polon.*, volume B35, pages 351–364, 2004.
- [107] M. Settimo and M. De Domenico. Propagation of extragalactic photons at ultra-high energy with the EleCa code. *Astropart. Phys.*, 62:92–99, March 2015.
- [108] T. Stanev. *High Energy Cosmic Rays*. 2366-0082. Springer-Verlag Berlin Heidelberg, second edition, 2010.
- [109] R. J. Protheroe and P. A. Johnson. Propagation of ultra high energy protons and gamma rays over cosmological distances and implications for topological defect models. *Astropart. Phys.*, 5(2):215, August 1996.

- [110] A. A. Zdziarski. Saturated pair-photon cascades on isotropic background photons. *Astrophys. J.*, 335:786–802, December 1988.
- [111] A. Chen, J. Dwyer, and P. Kaaret. The Inverse Compton and Extragalactic Components of Diffuse Gamma-Ray Emission. *Astrophys. J.*, 463:169, May 1996.
- [112] The HESS Collaboration. Primary particle acceleration above 100 TeV in the shell-type Supernova Remnant RX J1713.7-3946 with deep H.E.S.S. observations. *Astron. Astrophys.*, 464:235–234, 2007.
- [113] The IceCube Collaboration, The Pierre Auger Collaboration, and The Telescope Array Collaboration. Report from the Multi-Messenger Working Group at UHECR-2014 Conference. *JPS Conf. Proc.*, 9(010021), 2016.
- [114] The Pierre Auger Collaboration. Search for photons above  $10^{18}$  eV using the hybrid detector of the Pierre Auger Observatory. *Journal of Cosmology and Astropart. Phys.*, 2017(009), April 2017.
- [115] C. Bleve. Updates on the neutrino and photon limits from the Pierre Auger Observatory. In Proceedings of Science, editor, *The 34th International Cosmic Ray Conference*, 2015.
- [116] The Pierre Auger Collaboration. The Pierre Auger Observatory III: Other Astrophysical Observations. In *Contributions to the 32<sup>nd</sup> International Cosmic Ray Conference*, 2011.
- [117] The Telescope Array Collaboration. Telescope Array search for photons and neutrinos with the surface detector data. In *Proceedings of The 34th International Cosmic Ray Conference*, volume 331. The Hague, The Netherlands, July 2016.
- [118] A. V. Glushkov et al. Constraints on the flux of primary cosmic-ray photons at energies  $E > 10^{18}$  eV from Yakutsk muon data. *Phys. Rev. D*, 82(4), August 2010.
- [119] M. Ave et al. New constraints from Haverah Park data on the photon and iron fluxes of UHE cosmic rays. *Phys. Rev. Lett.*, 85:2244, 2000.
- [120] J. R. Ellis, V. E. Mayes, and D. V. Nanopoulos. UHECR particle spectra from crypton decays. *Phys. Rev. D*, 74(115003), 2006.
- [121] Y. Formin, N. Kalmykov, and G. Kulikov. Estimate of the fraction of primary photons of the fraction of primary photons in the cosmic-ray flux at energies  $10^{17}$  eV from the EAS-MSU experiment data. *JETP*, 117(6):1011–1023, 12 2013.
- [122] Y. Formin et al. Estimates of the cosmic gamma-ray flux at PeV to EeV energies from the EAS-MSU experiment data. *JETP Letters*, 100(11):699–702, February 2015.
- [123] Y. Formin et al. Constraints on the flux of  $10^{16}$  -  $10^{17.5}$  eV cosmic photons from the EAS-MSU muon data. *Phys. Rev. D*, 95(12), June 2017.
- [124] D. Kang et al. KASCADE-Grande. In *34<sup>th</sup> International Cosmic Ray Conference*, 2015.
- [125] D. Kang et al. A limit on the diffuse gamma-rays measured with KASCADE-Grande. *Journal of Physics: Conference Series*, 632(012013), 2015.
- [126] The KASCADE-Grande Collaboration. KASCADE-Grande Limits on the Isotropic Diffuse Gamma-Ray Flux between 100 TeV and 1 EeV. *Astrophys. J.*, 848(1), October 2017.

- [127] M. Risse et al. Search for ultra-high energy photons using air showers. *Mod. Phys. Lett. A*, 22(11), April 2007.
- [128] J. Bellido for the Pierre Auger Collaboration. Depth of maximum of air-shower profiles at the Pierre Auger Observatory Measurements above  $10^{17.2}$  eV and Composition Implications. In *35th International Cosmic Rays Conference, Busan, Korea, 2017*.
- [129] M. Fairbairn, T. Rashba and S. Troitsky. Photon-axion mixing and ultra-high-energy cosmic rays from BL Lac type objects - Shining light through the Universe. *Phys. Rev. D*, 84(125019), December 2011.
- [130] M. Galaverni and G. Sigl. Lorentz Violation for Photons and Ultra-High Energy Cosmic Rays. *Phys. Rev. Lett.*, 100(021102), January 2008.
- [131] The IceCube Collaboration. Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector. *Science*, 342(6161), 2013.
- [132] C. Grupen and I. Buvat. *Handbook of Particle Detection and Imaging*, chapter Indirect Detection of Cosmic Rays by R. Engel, pages 594–632. Springer-Verlag, 2012.
- [133] D. Zavrtnik. Ultra high energy cosmic rays. *Contemporary Physics*, 51(6):513–529, 2010.
- [134] A. A. Watson. The search for the highest energy particles in nature. In *Proc. de la Escuela Mexicana de Astrofísica*, 2002.
- [135] R. Engel, D. Heck, and T. Pierog. Extensive Air Showers and Hadronic Interactions at High Energy. *Ann. Rev. Nucl. Part. Sci.*, 61:467–489, November 2011.
- [136] S. Andringa et al. The muonic longitudinal shower profiles at production. *Astropart. Phys.*, 35:821–827, 2012.
- [137] T. Pierog et al. Latest Results of Air Shower Simulation Programs CORSIKA and CONEX. In *30<sup>rd</sup> International Cosmic Ray Conference*, volume 4, pages 625–628, 2007.
- [138] M. Bahmanabadi and S. Mortazavi. Simulation of the time structure of Extensive Air Showers with CORSIKA initiated by various primary particles at Alborz-I observatory level. *New Astronomy*, 61:5–13, 2018.
- [139] D. Heck and T. Pierog. Extensive air shower simulation with CORSIKA: A User's guide. Institut für Kernphysik, Forschungszentrum Karlsruhe GmbH, Karlsruhe, Germany.
- [140] H-J. Drescher and G. R. Farrar. Dominant Contributions to Lateral Distribution Functions in Ultra-High Energy Cosmic Ray Air Showers. *Astropart. Phys.*, 19:235–244, 2003.
- [141] S. Messina. *Extension to lower energies of the cosmic-ray energy window at the Pierre Auger Observatory*. PhD thesis, KVI-CART, University of Groningen, 2016. Submitted as GAP Note 2017-044.
- [142] W. Heitler. *The Quantum Theory of Radiation*. Oxford University Press, 1944.
- [143] N. N. Kalmykov et al. One-dimensional hybrid simulation of EAS using cascade equations. *Proc. 28<sup>th</sup> International Cosmic Ray Conference, Tsukuba, Japan*, 2:511–514, 2003.

- [144] T. K. Gaisser and A. M. Hillas. Reliability of the method of constant intensity cuts for reconstructing the average development of vertical showers. In *15<sup>rd</sup> International Cosmic Ray Conference, 1977*.
- [145] K. Kamata and J. Nishimura. The Lateral and the Angular Structure Functions of Electron Showers. *Progr. Theor. Phys. Suppl.*, 6:93–155, 1958.
- [146] K. Greisen. Cosmic Ray Showers. *Ann. Rev. Nucl. Sci.*, 10:63–108, 1960.
- [147] J. Matthews. A Heitler model of extensive air showers. *Astropart. Phys.*, 22(5-6):387–397, January 2005.
- [148] The Pierre Auger Collaboration. Upper limit on the cosmic-ray photon flux above  $10^{19}$  eV using the surface detector of the Pierre Auger Observatory. *Astropart. Phys.*, 29(4):243–256, May 2008.
- [149] M. Settimo for the Pierre Auger Collaboration. Search for ultra-high energy photons with the Pierre Auger Observatory. In *Photon 2013: International Conference on the Structure and the Interactions of the Photon*. International Conference on the Structure and the Interactions of the Photon, May 2013.
- [150] P. K. F. Grieder. *Extensive Air Showers*. Springer-Verlag Berlin Heidelberg, first edition, 2010.
- [151] T. Erber. High-Energy Electromagnetic Conversion Processes in Intense Magnetic Fields. *Reviews of Modern Physics*, 38(4):626–659, 1966.
- [152] The High Resolution Fly’s Eye Collaboration. Measurement of the flux of ultrahigh energy cosmic rays from monocular observations by the High Resolution Fly’s Eye experiment. *Phys. Rev. Lett.*, 92:151101, 2004.
- [153] A. Etchegoyen for The Pierre Auger Collaboration. Properties of the site and layout of the Pierre Auger Observatory. *Proc. 27<sup>th</sup> International Cosmic Ray Conference, Hamburg, Germany, HE1:703–706*, 2001.
- [154] The Pierre Auger Collaboration. Pierre Auger Project Design Report. Technical report, Fermi Laboratory, 2nd Edition, March 1997.
- [155] I. Allekotte et al. Site survey for the Pierre Auger observatory. *J. Phys. G: Nucl. Part. Phys.*, 28(6), May 2002.
- [156] B. Dawson for the Pierre Auger Collaboration. Hybrid Performance of the Pierre Auger Observatory. In *30<sup>th</sup> International Cosmic Ray Conference, 2007*.
- [157] The Pierre Auger Collaboration. Properties and performance of the prototype instrument for the Pierre Auger Observatory. *Nucl. Instrum. and Meth.*, A523:50–95, 2004.
- [158] R. Conceicao for the Pierre Auger Collaboration. The Pierre Auger Observatory: results on the highest energy particles. In *Time and Matter, 2013*.
- [159] D. Ravignani for the Pierre Auger Collaboration. Measurement of the energy spectrum of cosmic rays above  $3 \times 10^{17}$  eV using the AMIGA 750 m surface detector array of the Pierre Auger Observatory. *33<sup>rd</sup> International Cosmic Ray Conference, 2013*.

- [160] M. Nagano et. al. New measurement on photon yields from air and the application to the energy estimation of primary cosmic rays. *Astropart. Phys.*, 22(3-4):235–248, 2004.
- [161] J. F. Debatin. Investigation of optical properties of the fluorescence telescopes of the Pierre Auger Observatory. Master's thesis, Karlsruhe Institute of Technology, 2015.
- [162] The Pierre Auger Collaboration. The Pierre Auger Cosmic Ray Observatory. *Nucl. Instr. Meth.*, 798:172–213, 2015.
- [163] The Pierre Auger Collaboration. A Study of the Effect of Molecular and Aerosol Conditions in the Atmosphere on Air Fluorescence Measurements at the Pierre Auger Observatory. *Astropart. Phys.*, 33(2):108–129, 2010.
- [164] The Pierre Auger Collaboration. The Rapid Atmospheric Monitoring System of the Pierre Auger Observatory. *JINST*, 7(P09001), 2012.
- [165] B. Fick for the Pierre Auger Collaboration. Techniques for measuring aerosol attenuation using the Central Laser Facility at the Pierre Auger Observatory. *JINST*, 8, 2013.
- [166] P. van Bodegom et. al. Cloud Camera. In *Pierre Auger November Meeting*, 2016.
- [167] The Pierre Auger Collaboration. Description of Atmospheric Conditions at the Pierre Auger Observatory using the Global Data Assimilation System (GDAS). *Astropart. Phys.*, 35:591–607, 2012.
- [168] B. Dawson. Amplitude Dynamic Range in Auger Fluorescence Electronics: Update for the Four Eyes System at Nihuil. GAP Note 1999-038.
- [169] The Pierre Auger Collaboration. The fluorescence detector of the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, 620:227–251, 2010.
- [170] C. Bonifazi for the Pierre Auger Collaboration. The angular resolution of the Pierre Auger Observatory. *Nucl. Phys. Proc. Suppl.*, 190:20–25, 2009.
- [171] M. Mostafa et al. Hybrid Activities of the Pierre Auger Observatory. *Nucl. Phys. Proc. Suppl.*, 165:50, 2007.
- [172] D. Kuempel, K-H Kampert, and M. Risse. Geometry reconstruction of fluorescence detectors revisited. *Astropart. Phys.*, 30(4):167–174, 2008.
- [173] M. Tueros. Estimate of the non-calorimetric energy of showers observed with the fluorescence and surface detectors of the Pierre Auger Observatory. In *33<sup>rd</sup> International Cosmic Ray Conference*, 2013.
- [174] V. Verzi for the Pierre Auger Collaboration. The Energy Scale of the Pierre Auger Observatory. In *33<sup>rd</sup> International Cosmic Ray Conference*, 2013.
- [175] M. Unger et al. Reconstruction of longitudinal profiles of ultra-high energy cosmic ray showers from fluorescence and Cherenkov light measurements. *Nucl. Instrum. Meth. A*, 588(3):433–441, 2008.
- [176] I. Allekotte et al. The surface detector system of the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 586(3), March 2008.

- [177] R. Sato. Long Term Performance of the Surface Detectors of the Pierre Auger Observatory. In *32<sup>th</sup> International Cosmic Ray Conference*, 2011.
- [178] The Pierre Auger Collaboration. Calibration of the surface array of the Pierre Auger Observatory. *Nucl. Instrum. Meth. A.*, 568(2):839–846, December 2006.
- [179] M. Aglietta for the Pierre Auger Collaboration. Response of the Pierre Auger Observatory water Cherenkov detectors to muons. In *29<sup>rd</sup> International Cosmic Ray Conference*, 2005.
- [180] P. Billoir. Proposition to improve the local trigger of the Surface Detector for low energy showers. GAP Note 2009-179.
- [181] P. Billoir, P. Ghia, D. Nitz, and R. Sato. First results of the ToTd trigger in the test hexagon. GAP Note 2011-032.
- [182] P. Billoir. New proposal to improve the local trigger of the Surface Detector. GAP Note 2011-089.
- [183] The Pierre Auger Collaboration. Trigger and aperture of the surface detector array of the Pierre Auger Observatory. *Nucl. Instrum. Meth., A*(613):29–39, 2010.
- [184] V. Van Elewyck. Detection of Inclined and Horizontal Showers in the Pierre Auger Observatory. *AIP Conference Proceedings*, 809(187), 2006.
- [185] B. Zamorano Garcia. *Inferring the mass composition of ultra-high energy cosmic rays through the characterisation of the muon production profile*. PhD thesis, Universidad de Granada, 2014.
- [186] D. Allard for the Pierre Auger Collaboration. The trigger system of the Pierre Auger Surface Detector: operation, efficiency and stability. In *29<sup>th</sup> International Cosmic Ray Conference*, volume 7, pages 287–290, 2005.
- [187] D. Veberic and M. Roth. SD reconstruction manual. GAP Note 2005-035.
- [188] I. C. Maris. Measurement of the Ultra High Energy Cosmic Ray Flux using Data of the Pierre Auger Observatory. GAP Note 2008-026.
- [189] M. Horvata and D. Veberic. On Shower-Front Start-Time Variance. GAP Note 2007-057.
- [190] D. Barnhill for the Pierre Auger Collaboration. Measurement of the lateral distribution function of UHECR air showers with the Auger Observatory. In *29<sup>th</sup> International Cosmic Ray Conference*, 2005.
- [191] T. Schmidt, I. C. Maris, and M. Roth. Fine tuning of the LDF parametrization and the influence on S1000. GAP Note 2006-070.
- [192] D. Newton, J. Knapp, and A. A. Watson. The Optimum Distance at which to Determine the Size of a Giant Air Shower. GAP Note 2006-045.
- [193] M. Ave et al. The Accuracy Of Signal Measurement With The Water Cherenkov Detectors Of The Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 578(180), 2007.
- [194] M. Settimo. *Hybrid detection of Ultra High Energy Cosmic Rays with the Pierre Auger Observatory*. PhD thesis, Universita del Salento, 2010.
- [195] M. Aglietta et al. Recovery of saturated signals of the surface detector. GAP Note 2008-030.

- [196] M. Aglietta et al. SD signal saturation: Tests with low gain PMTs. GAP Note 2011-059.
- [197] R. Pesce. Energy calibration of data recorded with the surface detectors of the Pierre Auger Observatory: an update. In *32<sup>nd</sup> International Cosmic Ray Conference*, 2011.
- [198] J. Hersil et al. Observations of Extensive Air Showers near the Maximum of Their Longitudinal Development. *Phys. Rev. Lett.*, 6:22–23, 1961.
- [199] M. Unger for the Pierre Auger Collaboration. Study of the Cosmic Ray Composition above 0.4 EeV using the Longitudinal Profiles of Showers observed at the Pierre Auger Observatory. In *30<sup>rd</sup> International Cosmic Ray Conference*, 2007.
- [200] The Pierre Auger Collaboration. Observation of the suppression of the flux of cosmic rays above  $4 \times 10^{19}$  eV. *Phys. Rev. Lett.*, 101(061101), 2008.
- [201] M. Nagano et al. Energy spectrum of primary cosmic rays above  $10^{17}$  eV determined from extensive air shower experiments at Akeno. *J. Phys. G: Nucl. Part. Phys.*, 18(2):423, 1992.
- [202] D. J. Bird et al. The cosmic-ray energy spectrum observed by the Fly’s Eye. *Astrophys. J.*, 424(1):491–502, 1994.
- [203] T. Abu-Zayyad et al. Measurement of the Cosmic-Ray Energy Spectrum and Composition from  $10^{17}$  to  $10^{18.3}$  eV Using a Hybrid Technique. *Astrophys. J.*, 557:689–699, 2001.
- [204] D. Fuhrmann et al. KASCADE-Grande measurements of energy spectra for elemental groups of cosmic rays. In *32<sup>rd</sup> International Cosmic Ray Conference*, 2011.
- [205] S. Messina and Ad M. van den Berg. Reconstruction comparison between Infill and AERA array. GAP Note 2013-049, 2013.
- [206] S. Messina et al. Energy calibration for the 433 m SD Infill Array. GAP Note 2014-094, July 2014.
- [207] S. Messina et al. Reconstruction of AERA-triggered air showers. GAP Note 2013-115, 2013.
- [208] The MINOS Collaboration. The MINOS detectors technical design report. Technical report, Fermi National Accelerator Laboratory, 1998.
- [209] B. Wundheiler. *Rayos Còsmicos con Energías entre  $10^{17}$  y  $10^{19}$  eV*. PhD thesis, Universidad de Buenos Aires, 2013.
- [210] The Pierre Auger Collaboration. Prototype muon detectors for the AMIGA component of the Pierre Auger Observatory. *JINST*, 11, February 2016.
- [211] B. Wundheiler. The AMIGA Muon Counters of the Pierre Auger Observatory: Performance and Studies of the Lateral Distribution Function. In *International Cosmic Ray Conference (ICRC)*, 2015.
- [212] A. Etchegoyen for the Pierre Auger Collaboration. AMIGA, Auger Muons and Infill for the Ground Array. In *In Proc. 30<sup>th</sup> International Cosmic Rays Conference*, volume 1307, 2007.
- [213] J. M. Figueira. Simulations of AMIGA shielding: electromagnetic contamination, transmission and energy threshold. In *Pierre Auger March Meeting*, 2016.

- [214] J. M. Figueira. An improved reconstruction method for the AMIGA detectors. In *35<sup>th</sup> International Cosmic Ray Conference*, 2017.
- [215] M. Platino for The Pierre Auger Collaboration. AMIGA - Auger Muons and Infill for the Ground Array of the Pierre Auger Observatory. In *31<sup>st</sup> International Cosmic Ray Conference*, 2009.
- [216] B. Wundheiler for the Pierre Auger Collaboration. The AMIGA muon counters of the Pierre Auger Observatory: performance and first data. In *32<sup>nd</sup> International Cosmic Ray Conference*, 2011.
- [217] A. Almela et al. Design and implementation of an embedded system for particle detectors. In *33<sup>rd</sup> International Cosmic Ray Conference*, 2013.
- [218] O. Wainberg et al. Digital electronics for the Pierre Auger Observatory AMIGA muon counters. *JINST*, 9, April 2014.
- [219] M. Maur. Improving the Composition Sensitivity of the Pierre Auger Observatory with Additional Scintillators. Master's thesis, Karlsruhe Institute of Technology, 2012.
- [220] J. van Buren for the KASCADE Collaboration. Muon size spectrum measured by KASCADE-Grande. *26<sup>rd</sup> International Cosmic Ray Conference*, 6:301–304, 2005.
- [221] A. D. Supanitsky et al. Underground muon counters as a tool for composition analyses. *Astropart. Phys.*, 29:461–470, 2008.
- [222] D. Ravnani, A. D. Supanitsky, and D. Melo. Reconstruction of air shower muon densities using segmented counters with time resolution. *Astropart. Phys.*, 82:108–116, 2016.
- [223] H. Falckle et al. Detection and imaging of atmospheric radio flashes from cosmic ray air showers. *Nature*, 435(7040):313–316, 2005.
- [224] P. Schellart et al. Detecting cosmic rays with the LOFAR radio telescope. *Astronomy and Astrophysics*, 560(A98):14, 2013.
- [225] A. Aminaei for the Pierre Auger Collaboration. An Introduction to the Auger Engineering Radio Array (AERA). *Proceedings of Science*, 2010.
- [226] A. Etchegoyen for the Pierre Auger Collaboration. The Pierre Auger Project and Enhancements. In *VIII Latin American Symposium on Nuclear Physics and Applications*, 2009.
- [227] S. Falk et al. A First Look at HEAT Data. GAP Note 2010-123.
- [228] The Pierre Auger Collaboration. The Pierre Auger Observatory Upgrade: AugerPrime (Preliminary Design Review). *arXiv:1604.03637*, 2016.
- [229] S. Argiro et al. The Offline software framework of the Pierre Auger Observatory. *Nucl. Instrum. Meth. A.*, 580:1485–1496, 2007.
- [230] D. Heck et al. CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers. *Report FZKA*, 6019, 1998.
- [231] W. R. Nelson et al. The EGS4 Code System. Technical Report Report No. SLAC-265, Stanford Linear Accelerator Center, 1985.

- [232] The Pierre Auger Collaboration. Testing Hadronic Interactions at Ultrahigh Energies with Air Showers Measured by the Pierre Auger Observatory. *Phys. Rev. Lett.*, 117(192001), 2016.
- [233] S. Ostapchenko. On the re-summation of enhanced pomeron diagrams. *Phys. Lett. B*, 636(1):40–45, April 2006.
- [234] A. Ferrari, P. Sala, A. Fasso, and J. Ranft. FLUKA: A multi-particle transport code. *CERN-2005-010, SLAC-R-773, INFN-TC-05-11*, 2005.
- [235] F. Sánchez et al. The AMIGA extension of Auger Offline Software. GAP Note 2012-135, 2012.
- [236] N. Gonzalez et al. Comparison between simulated and measured single photoelectron pulses. GAP Note 2014-049.
- [237] M. Settimo et al. Trigger probability for single stations and air-showers with the TOTd and MoPS algorithms. GAP Note 2013-114, 2013.
- [238] M. Niechciol. *A New Window to the Universe? Searching for Ultra-High-Energy Photons at the Pierre Auger Observatory*. PhD thesis, University of Siegen, 2015. Submitted as GAP Note 2015-067.
- [239] C. J. Clopper and E. S. Pearson. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4):404–413, 1934.
- [240] B. Ravenu and P. Billoir. Notions of Geodesy and the UTM Coordinate System. Proposition of a Local Cartesian Coordinate System for the Southern Auger Site. GAP Note 2001-038.
- [241] The Pierre Auger Collaboration. An upper limit to the photon fraction in cosmic rays above  $10^{19}$  eV from the Pierre Auger Observatory. *Astropart. Phys.*, 27:155–168, 2007.
- [242] S. Messina. Private communication. March 2016.
- [243] The Pierre Auger Collaboration. Muons in air showers at the Pierre Auger Observatory: Mean number in highly inclined events. *Phys. Rev. D*, 91(032003), 2015.
- [244] G. Ros et al. A new surface parameter for composition studies at high energies. *Proc. of the International Cosmic Ray Conference*, 2009.
- [245] The Pierre Auger Collaboration. A search for point sources of EeV photons. *Astrophys. J.*, 789(2), June 2014.
- [246] A.D. Supanitsky et al. Underground muon counters as a tool for composition analyses. *Astropart. Phys.*, 29(6):461–470, July 2008.
- [247] A. Etchegoyen for the Pierre Auger Collaboration. AMIGA Design Report. Technical report, Instituto de Tecnologias en Deteccion y Astroparticulas, 2006.
- [248] G. Ros et al.  $S_b$  for photon-hadron discrimination. GAP Note 2010-052.
- [249] S. Müller and M. Roth. A CORSIKA study on the influence of muon detector thresholds on the separability of primary cosmic rays at highest energies. *Proc. of the International Cosmic Ray Conference*, 2015.
- [250] B. Efron. Bootstrap methods: another look at the jackknife. *The Annals of Statistics* 7, pages pp. 1–26, 1979.

- [251] C. Bérat, F. Montanet and M. Settimo. Study of proton showers with leading  $\pi^0$ . GAP Note 2016-052, 2016.
- [252] M. Niechciol et al.  $F_\gamma$ : a new observable for photon-hadron discrimination in hybrid air shower events. *Astropart. Phys.*, 97:85–95, 2018.
- [253] T. W. Anderson et. al. Review: Contemporary Textbooks on Multivariate Statistical Analysis: A Panoramic Appraisal and Critique. *Journal of the American Statistical Association*, 81(394), June 1986.
- [254] R. A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Human Genetics*, 7(2):179–188, 1936.
- [255] A. Hoecker et al. TMVA - Toolkit for Multivariate Data Analysis. *PoS ACAT*, 040(CERN-OPEN-2007-007), 2009.
- [256] S. Ostapchenko. QGSJET-II: towards reliable description of very high energy hadronic interactions. *Nucl. Phys. B*, 151:143–146, 2006.
- [257] G. J. Feldman and R. D. Cousins. A Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57(3873), 1998.
- [258] G. Schatz et al. Search for extremely high energy gamma rays with the KASKADE experiment. In *28<sup>th</sup> International Cosmic Ray Conference*, volume 4, 2003.
- [259] M. C. Chantell et al. Limits on the Isotropic Diffuse Flux of Ultrahigh Energy Gamma Radiation. *Phys. Rev. Lett.*, 79(1805), 1997.
- [260] M. Ave, M. Roth, and A. Schulz. A universal description of temporal and lateral distributions of ground particles in extensive air showers. *Proc. 34th Int. Cosmic Ray Conf.*, 378, 2015. The Hague, Netherlands.



# Appendix A

## Sequences of Offline modules

### Simulation module sequence

```
<sequenceFile xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation='@SCHEMALOCATION@/ModuleSequence.xsd'>
  <enableTiming/>
  <moduleControl>
    <loop numTimes="1" pushEventToStack="yes">
      <module> EventFileReaderOG </module>
    <loop numTimes="10" pushEventToStack="yes">
      <module> EventGeneratorOG </module>
      <loop numTimes="unbounded" pushEventToStack="no">
        <module> CachedXShowerRegeneratorAG </module>
        <module> G4XTankSimulatorAG </module>
      </loop>
      <module> SdSimulationCalibrationFillerOG </module>
      <module> SdPMTSimulatorOG </module>
      <module> SdFilterFADCSimulatorMTU </module>
      <module> SdBaselineSimulatorOG </module>
      <module> TankTriggerSimulatorOG </module>
      <module> TankGPSSimulatorOG </module>
      <module> UnderGrdInjectorAG </module>
      <module> EdepSimulatorAG </module>
      <module> MdCounterSimulatorAG </module>
      <module> ClearParticleLists </module>
      <try>
        <module> CentralTriggerSimulatorXb </module>
        <module> CentralTriggerEventBuilderOG </module>
        <module> EventBuilderOG </module>
      </try>
      <module> EventFileExporterOG </module>
    </loop>
  </loop>
</moduleControl>
</sequenceFile>
```

## MC reconstruction module sequence

```

<sequenceFile xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation='@SCHEMALOCATION@/ModuleSequence.xsd'>
  <enableTiming/>
  <moduleControl>
    <loop numTimes="unbounded" pushEventToStack="yes">
      <module> EventFileReaderOG </module>
      <module> EventCheckerOG </module>
      <module> SdPMTQualityCheckerKG </module>
      <module> SdCalibratorOG </module>
      <module> SdSignalRecoveryKLT </module>
      <module> PositionRejector </module>
      <try>
        <module> SdEventSelectorOG </module>
        <module> SdMonteCarloEventSelectorOG </module>
        <module> SdPlaneFitOG </module>
        <module> LDFFinderKG </module>
        <module> DLECorrectionWG </module>
        <module> SdEventPosteriorSelectorOG </module>
        <module> MdMuonCounterAG </module>
        <module> MdoduleRejectorAG </module>
        <module> MdBiasCorrecterAG </module>
        <module> MdEventSelectorAG </module>
        <module> MdLDFFinderAG </module>
      </try>
      <module> EventSqueezer </module>
    </loop>
  </moduleControl>
</sequenceFile>

```

## Data reconstruction module sequence

```

<sequenceFile xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation="@SCHEMALOCATION@/ModuleSequence.xsd">
  <enableTiming/>
  <moduleControl>
    <loop numTimes="unbounded" pushEventToStack="yes">
      <module> EventFileReaderOG      </module>
      <module> EventCheckerOG        </module>
      <module> SdPMTQualityCheckerKG  </module>
      <module> TriggerTimeCorrection  </module>
      <module> SdCalibratorOG         </module>
      <module> SdStationPositionCorrection </module>
      <module> SdBadStationRejectorKG </module>
      <module> SdSignalRecoveryKLT    </module>
    <try>
      <module> SdEventSelectorOG      </module>
      <module> SdPlaneFitOG           </module>
      <module> LDFFinderKG            </module>
      <module> DLECorrectionWG        </module>
      <module> SdEventPosteriorSelectorOG </module>
      <module> MdMuonCounterAG        </module>
      <module> MdoduleRejectorAG      </module>
      <module> MdBiasCorrecterAG      </module>
      <module> MdEventSelectorAG      </module>
      <module> MdLDFFinderAG          </module>
    </try>
    <module> EventSqueezer </module>
  </loop>
</moduleControl>
</sequenceFile>

```



## Appendix B

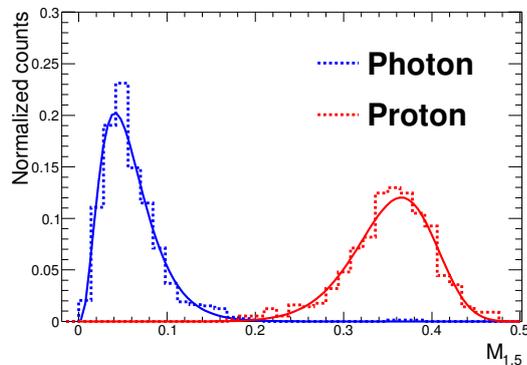
# Model for the $M_b$ distribution

All analyses presented in this thesis are aimed towards the rejection of nearly all the background events, while keeping most of the signal events. Since in practice the hadronic background will be hugely dominant over any ultra-high energy photon component in the cosmic-ray flux, the proper description of its  $M_b$  distribution (particularly, the tail towards the photon distribution) is a key aspect.

Both distributions can be described by a Weibull model, as shown in Eq. B.1, with four free parameters  $p_0$ ,  $x_0$ ,  $\alpha$  and  $\beta$ .

$$P(x) = p_0 \times \left(\frac{x}{x_0}\right)^\alpha \times e^{-\left(\frac{x}{x_0}\right)^\beta} \quad (\text{B.1})$$

The original and the modelled distributions are shown in Fig. B.1 for a subset of events, where the free parameters are estimated by an unbinned likelihood fit. The shape of the distributions are fairly well described, although the peak is usually underestimated by the model. Although the region around  $M_{1.5} \sim 0.2$  needs to be modelled more accurately, the model provided by Eq. B.1 may be a starting point of a more elaborated iterative fitting procedure.



**Figure B.1:** The  $M_{1.5}$  normalized distributions for photon (dotted blue) and proton (dotted red) events with the corresponding fitted models (in solid lines) from Eq. B.1. The events are selected according to their reconstructed energy between  $E_{\text{rec}} = 10^{16.9}$  eV and  $E_{\text{rec}} = 10^{17}$  eV and  $\theta_{\text{MC}} = 0^\circ$ .



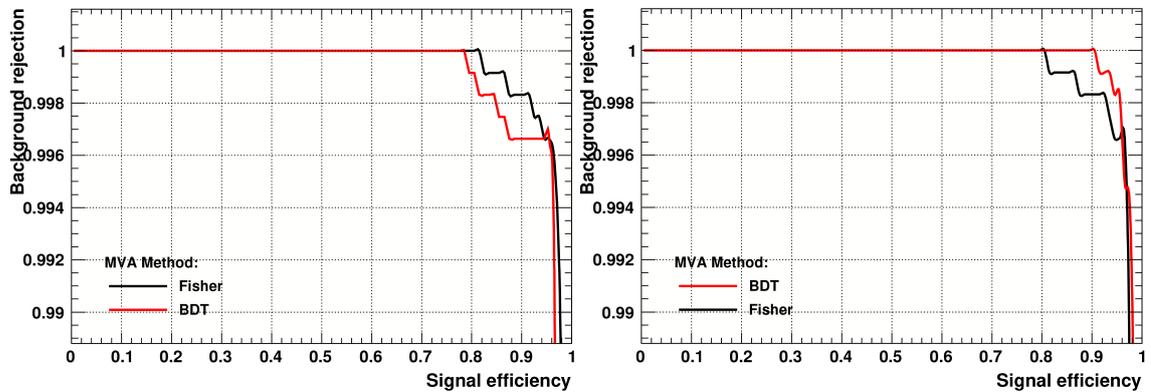
## Appendix C

# A thought experiment: adding $X_{\max}$ to the MVA

The discrimination procedure is based on the observables  $M_b$  and  $Q$ , measured with the MD and SD respectively. The photon-initiated showers are expected to develop deeper than their hadronic counterparts, and thus,  $X_{\max}$  differs between each type of primary. This approach has been exploited in photon searches at energies above  $10^{18}$  eV.

The HEAT telescopes may be used to estimate  $X_{\max}$  under specific atmospheric conditions. Thus, a thought experiment in which this observable is added to the MVA machinery is presented in this section. Fig. C.1 displays the background rejection in terms of the signal efficiency for subset of events between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{16.6}$  eV with a fixed zenith angle of  $\theta_{\text{MC}} = 0^\circ$  when the standard  $M_b + Q$  approach is used and when  $X_{\max}$  is added. The linear Fisher analysis (black line) barely changes with the addition of  $X_{\max}$ . However, a great improvement is seen in the case of the non-linear BDT analysis, which may lead to a complete background rejection at  $\sim 90\%$  signal efficiency. The additional information from the longitudinal profile of the shower has then a noticeable impact only if the discrimination frontier between the two classes (signal and background) is considered as a curved surface in the multiparameter phase space, which is why the Fisher discriminant is unchanged.

As mentioned before, the direct measurement of the longitudinal development may be chal-



**Figure C.1:** The background rejection against the signal efficiency using  $M_b$  and  $Q$  as discriminating observables (left) and adding  $X_{\max}$  to (right). The black (red) profile represents the results obtained with a Fisher (BDT) run. The events are selected between  $E_{\text{rec}} = 10^{16.4}$  eV and  $10^{16.6}$  eV with a fixed zenith angle of  $\theta_{\text{MC}} = 0^\circ$ .

lenging at these low energies. However, it may be possible to reconstruct it using the temporal structure of the SD signals through an alternative approach called *Universality* [260]. Originally designed for  $E > 10^{18.5}$  eV, an extension to lower energies is currently under development.